BLED eCONFERENCE 2026

UNI BIZ TECH GOV

University of Maribor Press

39th Bled eConference

CO-CREATING HUMAN-CENTRED AND RESPONSIBLE DIGITAL FUTURES CONFERENCE PROCEEDINGS

June 7 - June 10, 2026 | Bled, Slovenia

EDITORS

Doroteja **VIDMAR**

Andreja **PUCIHAR**

Mirjana **KLJAJIĆ BORŠTNAR**

Roger W. H. **BONS**

Matthias **GLOWATZ**

Hans-Dieter **ZIMMERMANN**

UNI
BIZ
TECH
GOV

University of Maribor
Faculty of Organizational Sciences

39th Bled eConference

Co-Creating Human-Centred and Responsible Digital Futures Conference Proceedings

Editors
Doroteja Vidmar
Andreja Pucihar
Mirjana Kljajić Borštnar
Roger W. H. Bons
Matthias Glowatz
Hans-Dieter Zimmermann

June 2026

Title 39th Bled eConference

Subtitle Co-Creating Human-Centred and Responsible Digital Futures; Conference Proceedings

Editors Doroteja Vidmar
(University of Maribor, Faculty of Organizational Sciences, Slovenia)

Andreja Pucihar
(University of Maribor, Faculty of Organizational Sciences, Slovenia)

Mirjana Kljajić Borštnar
(University of Maribor, Faculty of Organizational Sciences, Slovenia)

Roger W. H. Bons
(FOM Uniersity of Applied Sciences, Essen, Germany)

Matthias Glowatz
(University College Dublin, Ireland)

Hans-Dieter Zimmermann
(Eastern Switzerland University of Applied Sciences, Campus St. Gallen, Switzerland)

Review Hamed Abbaszadeh Avilagh (Åbo Akademi University, Finland), Satu Aksovaara (Jamk University of Applied Sciences, Finland), Art Alishani (Tartu University, Estonia), Selma Ayuanshari (Talinn university of technology, Estonia), Alenka Baggia (University of Maribor, Faculty or Organizational Sciences, Slovenia), Dusan Barac (University of Belgrade, Faculty of organizational sciences, Serbia), Valmir Bekiri (Eastern Switzerland University of Applied Sciences, Switzerland), Fatima Bilal (Åbo Akademi University, Finland), Staša Blatnik (University of Maribor, Faculty or Organizational Sciences, Slovenia), Diana Boekman (HU University of Applied Sciences Utrecht, The Netherlands), Roger W. H. Bons (Open Universiteit, the Netherlands), Mladen Borovič (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Alenka Brezavšček (University of Maribor, Faculty or Organizational Sciences, Slovenia), Helen Cripps (Edith Cowan University, Australia), Nejc Čelik (University of Maribor, Faculty or Organizational Sciences, Slovenia), Mitja Dečman (University of Ljubljana, Faculty of public administration, Slovenia), Andreas Emrich (German Research Center for Artificial Intelligence (DFKI), Germany), Kristina Eriksson-Backa (Åbo Akademi University, Finland), Esther Federspiel (Eastern Switzerland University of Applied Sciences, School of Management, Switzerland), Matt Glowatz (University College Dublin, Ireland), Nerko Hadziarapovic (HU University of Applied Sciences Utrecht, The Netherlands), Lucas Harms (HU University of Applied Sciences Utrecht, The Netherlands), Sami Heikkinen (LAB University of Applied Sciences, Faculty of Business and Hospitality Management, Finland), Vincentius Homburg (University of Tartu, Estonia), Štefan Horvat (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Rok Hržica (University of Ljubljana, Faculty of public administration, Slovenia), Ina Jäger (Technical University of Munich, Germany), Mercy Jemursoi Koech (Griffith College Dublin, Ireland), Eva Jereb (University of Maribor, Faculty or Organizational Sciences, Slovenia), David Jesenko (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Henrik Johan Nyman (Åbo Akademi University, Finland), Tina Jukić (University of Ljubljana, Faculty of Public Administration, Slovenia), Nexhat Kapidani (Maritime Safety Administration (MSA), Montenegro), Domen Kavran (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Štefan Kohek (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Alinda Kokkinou (Breda University of Applied

Sciences, The Netherlands), Simon Kolmanič (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Tanja Korhonen (Kajaani University of Applied Sciences, Finland), Eva Krhač (University of Maribor, Faculty or Organizational Sciences, Slovenia), Luka Lah (University of Maribor, Faculty or Organizational Sciences, Slovenia), Gregor Lenart (University of Maribor, Faculty or Organizational Sciences, Slovenia), Robert Leskovar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Luka Lukač (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Markus Makkonen (Tampere University, Finland), Matjaž Maletič (University of Maribor, Faculty or Organizational Sciences, Slovenia), Arjen Maris (University of Twente, The Netherlands), Marjeta Marolt (University of Maribor, Faculty or Organizational Sciences, Slovenia), Dušan Mežnar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Ondrej Mitas (Breda University of Applied Sciences, the Netherlands), Paul Morsch (HU University of Applied Sciences Utrecht, The Netherlands), Andrej Nerat (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Katrin Oettmeier (University Liechtenstein, Liechtenstein), Guido Ongena (HU University of Applied Sciences Utrecht, The Netherlands), Luuk P.A. Simons (Delft University of Technology, Faculty of Computer Science, the Netherlands), Zdravko Paladin (Maritime Safety Administration (MSA), Montenegro), Primož Pevcin (University of Ljubljana, Faculty of public administration, Slovenia), Andreja Pucihar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Pascal Ravesteijn (HU University of Applied Sciences Utrecht, The Netherlands), Dominik Rebholz (Fraunhofer FIT & University of Bayreuth, Germany), Christopher Reichstein (Baden-Wuerttemberg Cooperative State University, Germany), Blaž Rodič (University of Novo mesto, Faculty of information studies, Slovenia), Bogdan Romanov (Tartu University, Estonia), Alexander Ruh (Independent), Paulo Rupino da Cunha (University of Coimbra, CISUC/LASI, DEI, Portugal), Zaid Sako (La Trobe University, Australia), Robert Schmelzer (Chemnitz University of Technology, Germany), Juergen Seitz (Baden-Wuerttemberg Cooperative State University, Germany), Navin Sewberath Misser (HU University of Applied Sciences Utrecht, The Netherlands), Anand Sheombar (HU University of Applied Sciences Utrecht, The Netherlands), Muhammad Shuakat (La Trobe University, Australia), Koen Smit (HU University of Applied Sciences Utrecht, The Netherlands), Vjeran Strahonja (University of Zagreb Faculty of Organization and Informatics, Croatia), Damjan Strnad (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Andrej Škraba (University of Maribor, Faculty or Organizational Sciences, Slovenia), Ayesha Thanthrige (La Trobe University, Australia), Beat Tödtli (Eastern Switzerland University of Applied Sciences, Switzerland), Martina Tomicic Furjan (University of Zagreb, Faculty of organization and informatics, Croatia), Nevena Toporova (Technical University of Munich, Germany), Nalika Ulapane (La Trobe University, Australia), Annemae van de Hoef (HU University of Applied Sciences Utrecht, The Netherlands), Esther van der Stappen (Avans University of Applied Sciences, Research Group Digital Education, the Netherlands), Stan van Ginkel (HU University of Applied Sciences Utrecht, The Netherlands), John van Meerten (HU University of Applied Sciences Utrecht, The Netherlands), Johan Versendaal (HU University of Applied Sciences Utrecht and Open University of the Netherlands, The Netherlands), Doroteja Vidmar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Mirko Vintar (University of Ljubljana, Faculty of public administration, Slovenia), Doug Vogel (Harbin Institute of Technology, China), Sanja Vrbek (University of Ljubljana, Faculty of Public Administration, Slovenia), Thorsten Weber (FOM university of applied sciences, Germany), Tze Wei Liew (Multimedia University, Malaysia), Alessa Wein (German Research Center for Artificial Intelligence (DFKI), Germany), Borut Werber (University of Maribor, Faculty or Organizational Sciences, Slovenia), Nilmini Wickramasinghe (La Trobe University, Australia), John Zelcer (La Trobe University, Australia), Hans-Dieter Zimmermann (Eastern Switzerland University of Applied Sciences, Switzerland), Jasmina Žnidaršič (University of Maribor, Faculty or Organizational Sciences, Slovenia)

Language editing Barrak AlBabtain, Alenka Baggia, Matthias Baldauf, João Barata, Kevin Bieber, Fatima Bilal, Roger W. H. Bons, Annelies Botjes, Marie-Claude Boudreau, Alenka Brezavšček, Graham Burns, Jos van Capelleveen, Asma Chiha, Anastasiia Dardykina, Suzana Dodić, Kenneth Dosanjh, Mohamed Eledeisy, Ann-Kathrin Ermisch, Lorena García Ramón, Ahmad Ghazawneh, Eefje Gijzen, Swarupa Hardikar, Milena Head, Eija Heikkinen, Tim Huygh, Tina Jukić, Kristina Kaldon, Heikki Karjaluoto, Tiina Kemppainen, Christopher Killingley, Mirjana Kljajić Borštnar, Alinda Kokkinou, Anoushka Kothari, Raechel Kratochvit, Eva Krhač Andrašec, Eveliina Lakka, Sam Leewis, Evert Lindquist, Pamela Minch, Ondrej Mitas, Eve Mnai, Raghava Mukkamala, Andrew Mulley, Juha Munnukka, Youssouf Nabé, Aleksanteri Negru-Vode, Tobey Nelson, Vesna Novak, Tiina Nyman, Dragana Paladin, Yianna Paris, Andteas Peter, Hanne Petersen, Marija Pokos Lukinec, Pascal Ravesteijn, Lloyd Rutledge, Zaid Sako, Florian Schmitt, Tennessee Schrage, Michael Schöppe, Luuk Simons, Simon Sprott, Ousmane Sy, Eva Štupica, Perce Thrower, Valtteri Tomminen, Nalika Ulapane, Benjamin Urh, Petra Vujković, Franziska Wenger, Markus Westner, Sanaz Yassaee

Technical editors Aljaž Murko
(University of Maribor, Faculty of Organizational Sciences)

Jan Perša
(University of Maribor, University of Maribor Press)

Marina Bajić
(University of Maribor, University of Maribor Press)

Cover designer Jan Perša
(University of Maribor, University of Maribor Press)

Cover graphics Blejski otok, photo: Neven Krcmarek, unsplash.com, 2016

Graphics material Sources are own unless otherwise noted.
Authors of proceedings and Pucihar, Borštnar Kljajić, Vidmar, Bons, Glowatz, Zimmermann (editors), 2026

Conference 39th Bled eConference – Co-Creating Human-Centred and Responsible Digital Futures

Location and date Bled, Slovenia, June 7 – 10, 2026

Programme committee Art Alishani (Tartu University, Estonia), Alenka Baggia (University of Maribor, Faculty or Organizational Sciences, Slovenia), Matthias Baldauf (Eastern Switzerland University of Applied Sciences, Switzerland), Roger W. H. Bons (Open University, The Netherlands), Alenka Brezavšček (University of Maribor, Faculty or Organizational Sciences, Slovenia), Helen Cripps (Edith Cowan University, Australia), Matt Glowatz (University College Dublin, Ireland), Nerko Hadžiarapović (HU University of Applied Sciences Utrecht, The Netherlands), Marikka Heikkila (University of Turku, Finland), Tina Jukić (University of Ljubljana, Faculty of Public Administration, Slovenia), Mirjana Kljajić Borštnar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Sam Leewis (HU University of Applied Sciences Utrecht, The Netherlands), Gregor Lenart (University of Maribor, Faculty or Organizational Sciences, Slovenia), Xander Lub (HU University of Applied Sciences Utrecht, The Netherlands), Marjeta Marolt (University of Maribor, Faculty or Organizational Sciences, Slovenia), Axel Minten (FOM University of Applied Sciences, Germany), Domen Mongus (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Guido Ongena (HU University of Applied Sciences Utrecht, The Netherlands), Andreja Pucihar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Pascal Ravesteijn (HU University of Applied Sciences Utrecht, The Netherlands), Juergen Seitz (Baden-Wuerttemberg Cooperative State University, Germany), Anand Sheombar (HU University of Applied

Sciences Utrecht, The Netherlands), Koen Smit (HU University of Applied Sciences Utrecht, The Netherlands), Damjan Strnad (University of Maribor, Faculty or electrical engineering and computer science, Slovenia), Andrej Škraba (University of Maribor, Faculty or Organizational Sciences, Slovenia), Esther van der Stappen (Avans University of Applied Sciences, The Netherlands), Stan van Ginkel (HU University of Applied Sciences Utrecht, The Netherlands), Johan Versendaal (HU University of Applied Sciences Utrecht and Open University of the Netherlands, The Netherlands), Doroteja Vidmar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Doug Vogel (Harbin Institute of Technology, China), Nilmini Wickramasinghe (La Trobe University, Australia), Hans-Dieter Zimmerman (Eastern Switzerland University of Applied Sciences, Switzerland)

Organizing committee Petra Gorjanc (University of Maribor, Faculty or Organizational Sciences, Slovenia), Mirjana Kljajić Borštnar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Klara Knific (University of Maribor, Faculty or Organizational Sciences, Slovenia), Ana Kranner Porenta (University of Maribor, Faculty or Organizational Sciences, Slovenia), Gregor Lenart (University of Maribor, Faculty or Organizational Sciences, Slovenia), Maruša Luštrik (University of Maribor, Faculty or Organizational Sciences, Slovenia), Marjeta Marolt (University of Maribor, Faculty or Organizational Sciences, Slovenia), Aljaž Murko (University of Maribor, Faculty or Organizational Sciences, Slovenia), Andreja Pucihar (University of Maribor, Faculty or Organizational Sciences, Slovenia), Doroteja Vidmar (University of Maribor, Faculty or Organizational Sciences, Slovenia),

Published by **University of Maribor**
University of Maribor Press
Slomškov trg 15, 2000 Maribor, Slovenia
https://press.um.si, zalozba@um.si

Issued by **University of Maribor**
Faculty of Organizational Sciences
Kidričeva cesta 55a, Kranj, Slovenia
http://www.fov.um.si, dekanat.fov@um.si

Edition 1ST

Publication type E-book

Available at https://press.um.si/index.php/ump/sl/catalog/book/1128

Published Maribor, Slovenia, June 2026

© University of Maribor, University Press
/Univerza v Mariboru, Univerzitetna založba

Text © The authors of proceedings & Vidmar, Pucihar, Borštnar Kljajić, Bons, Glowatz, Zimmermann (editors), 2026

This book is published under a Creative Commons 4.0 International licence (CC BY 4.0). This license lets others remix, tweak, and build upon your work even for commercial purposes, as long as they credit you and license their new creations under the identical terms. This license is often compared to "copyleft" free and open source software licenses.

Any third-party material in this book is published under the book's Creative Commons licence unless indicated otherwise in the credit line to the material. If you would like to reuse any third-party material not covered by the book's Creative Commons licence, you will need to obtain permission directly from the copyright holder.

https://creativecommons.org/licenses/by /4.0/

CIP - Kataložni zapis o publikaciji
Univerzitetna knjižnica Maribor

005:004(082)(0.034.2)

BLED eConference Co-Creating Human-Centred and Responsible Digital Futures (39 ; 2026 ; Bled)
39th Bled eConference Co-Creating Human-Centred and Responsible Digital Futures [Elektronski vir] : conference proceedings : June 7 - June 10, 2026, Bled, Slovenia / editors Doroteja Vidmar ... [et al.]. - 1st ed. - E-zbornik. - Maribor : University of Maribor, University of Maribor Press, 2026

Način dostopa (URL): https://press.um.si/index.php/ump/catalog/book/1128

ISBN 978-961-299-152-4 (PDF)
doi: 10.18690/um.fov.4.2026
COBISS.SI-ID 280300547

ISBN 978-961-299-152-4 (pdf)

DOI https://doi.org/10.18690/um.fov.4.2026

Price Free copy

For publisher Prof. Dr. Zdravko Kačič,
Rector of University of Maribor

Attribution Vidmar, D., Pucihar, A., Borštnar Kljajić, M., Bons, R. W. H., Glowatz, M., Zimmermann, H. D. (Eds.) (2026). *39th Bled eConference Co-Creating Human-Centred and Responsible Digital Futures: Conference Proceedings*. University of Maribor, University of Maribor Press. doi: 10.18690/um.fov.4.2026

University of Maribor Press

Table of Contents

1	**Exploring Productivity Gains from GenAI-Enabled Virtual Assistants: A Configurational Study of a Multifaceted Phenomenon** Niklas Rieger	1
2	**Trust and Student Adoption of AI-Enhanced Augmented Reality in Daily Life: A PLS-SEM and fsQCA Investigation** Bad Re Alam, Anssi Öörni, Jozsef Mezei, Fatima Bilal	19
3	**Embodied Finance: A Conceptual Framework for Agency, Value, and Trust** Simone Chiappino	43
4	**Green Behavior and Well-being: A Bibliometric and Network Analysis of Research Trends** Hamed Abbaszadeh Avilagh, Parinaz Pazhuhanrad, Kristina Eriksson-Backa	59
5	**Designing a Layered Learning Infrastructure with a Digital Twin Component** Satu Aksovaara, Hannu Haapala, Minna Silvennoinen	81
6	**Pedagogical Data Production as the Generative Foundation of Meaningful Learning Analytics** Satu Aksovaara, Minna Silvennoinen	95
7	**Leveraging Transactional Business Data to predict Employee Workload Satisfaction in Operations: An Empirical Study – Part 1** Valmir Bekiri, Stefan Stöckler, Katrin Oettmeier	109
8	**The AI Implementation Paradox in Danish SMEs: Why Positive Intentions Fail to Translate into Sustained Use** Kenneth Dosanjh	123
9	**AI Adoption and Capability Gaps in Swiss Public Administration** Claudia Pedron, Hans-Dieter Zimmermann, Matthias Baldauf	143

10	**Fifty Shades of Orange: Construal Levels, Cognitive Biases, and The Amber Dilemma in Cybersecurity Risk Communication** Gulet Barre, Dinh Khoi Nguyen, Tim Huygh, Arno Nuijten	163
11	**Exploring the Relation between Information Security Policy Compliance and Cyber Resilience** Abigaïl de Rijp, Tim Huygh, Clara Maathuis, Rogier van de Wetering, Harald Vranken	177
12	**Students' Exploration of Elective Courses in Flexible Higher Education: Findings From a Digital Simulation Task** Suzan van Brussel, Johan Smarius, Sami Heikkinen, Esther van der Stappen	193
13	**Structural Governance Shortcomings in Business Model Canvas Use** Jos van Capelleveen, Guy Janssens, Rogier van de Wetering, Natasha Alechina, Gideon Maillette de Buy Wenniger	211
14	**An Identity-Oriented Framework for AI-Supported mHealth** Esther Federspiel	227
15	**Review of Disruptive Digital Solutions in the Function of Maritime Transport and Safety Matters** Zdravko Paladin, Mira Šorović, Nexhat Kapidani, Žarko Lukšić, Rasim Mujalović	243
16	**Psychologically Grounded Chatbots in Online Furniture Retail** Claudia Pedron, Esther Federspiel	259
17	**Exploring AI Adoption for Manufacturing SMEs in Industry 5.0** Lars-Olof Johansson	279
18	**Organisational Readiness for Open Government Data-Driven Co-Creation** Petra Vujković, Tina Jukić, Sanja Vrbek	295
19	**Hypertension eHealth Intervention National Upscaling** Luuk P.A. Simons	311
20	**Strategic Embedding of Responsible AI: Governance Architecture and Hybrid Human–AI Value Creation** Valtteri Tomminen, Matti Leppäniemi, Tiina Kemppainen, Lauri Frank	325
21	**From Algorithm to Doorstep: Consumer Acceptance of Chatbots in Last-mile Logistics** Alinda Kokkinou, Hans Quak, Ondrej Mitas	339

22	**Developing People-centric Data Governance with Service Design** Jaana Helsing, Tanja Korhonen	355
23	**Four Gen AI-Enabled Value Logics in Industrial B2B Markets** Eveliina Lakka, Joel Mero, Matti Leppäniemi, Heikki Karjaluoto	371
24	**Applying Antifragility in Design Principles for a Cyber Resilience Maturity Model** Edzo A. Botjes, Tim Huygh, Laury Bollen, Remko W. Helms	387
25	**Understanding Trust in GenAI: The Roles of Reliability, Explanation Quality, and Confidence in Judgment** Fatima Bilal, Anssi Öörni, Jozsef Mezei	407
26	**When Security Meets Intelligence: Trust in AI, Blockchain Security, and Investor Behavior** Fatima Bilal, Anssi Öörni, Raghava Rao Mukkamala, Bad Re Alam	429
27	**Bridging the Gap: Teaching Data Analytics to Business Students using R** Alinda Kokkinou, Reyhaneh Mazinani, Hans van Gils	447
28	**Why Do Consumers Purchase from a Controversial Online Store? Reasons for Purchasing from Temu** Tiina Kaleton, Tiina Kemppainen, Lauri Frank	463
29	**Impact of Digital Literacy on Turnover Intention Among Slovenian Healthcare Professionals** Nejc Bernik, Polona Šprajc	479
30	**Inertial Motion Capture and Digital Human Modeling for Ergonomic Risk Assessment of Accordion Performance** Tilen Medved, Branka Jarc Kovačič, Janez Kalan, Zvone Balantič	489
31	**Bridging the Theory-to-Design Gap in Digital Wellness: A Design Science Approach to Motivational and Compassionate System Design** Ayesha Thanthrige, Nilmini Wickramasinghe	505
32	**The Impact of Generative AI on B2B Marketing Processes: Evidence from Industrial Firms** Mikko Vesterinen, Joel Mero, Mika Skippari, Heikki Karjaluoto	523
33	**Structure of Data-Driven Business Models for the Circular Automotive Sector: A Taxonomy** Hannah Scheerer, Hendrik van der Valk	537

34	**Towards Pedagogically Structured Human–AI Engagement in a Peer Learning Context** Minna Silvennoinen, Sanna Paananen, Sami Määttä, Satu Aksovaara	555
35	**Automatic Transformation of Textual Content Using LLM-Technology: a Feasibility Study** Koen Smit, Sam Leewis, Annemae van de Hoef, Duain Castro	571
36	**(Open) Data Stewardship in the Public Sector: Evidence from a Survey of Slovenian Public Administration** Tina Jukić, Sanja Vrbek, Damijana Keržič	587
37	**Integrating AI into Small Firm Knowledge Work: A Qualitative Case Study** Tanja Korhonen, Heli Kemppainen	601
38	**Location-Aware Machine Learning to Forecast Urban Foot Traffic** Katayoon Pourmahdi, Bahareh Naseri, Ilia Gugenishvili, Jozsef Mezei, Anna-Greta Nyström	617
39	**Towards Agentic Personal Health and Organizational Domain Coordination in Occupational Nutrition** Kevin Nils Röhl, Rainer Alt, Jan Wirsam	631
40	**SME Production Resilience: A Four-Layer AI Use Case Selection Approach** Klaus Müller, Michael Schöppe, Horst Wildemann	647
41	**Big Five Personality Traits and Security Awareness Training Preferences** Isabel K. Glückert, Tobias Fertig, Ricard Bibaj, Sebastian Biedermann	663
42	**Analysis of the Impact of Advanced Technologies on Business Processes** Eva Krhač Andrašec, Tilen Medved, Tomaž Kern	681
43	**Efficient Explainable and Evidence Based Precision and Personalisation of Treatment** Nilmini Wickramasinghe, Nalika Ulapane, Anamika Ranaut, Neale Cohen	697
44	**When Text Is Not Enough: Structural Limits of Text-Only Transformer-Based Emotion Classification** Szymon Chirowski, Maciej Czerniak, Ondrej Mitas, Maks Burchard	705

45 **Future of Work in the Gastronomy Sector: Challenges of Adopting Artificial Intelligence in a Service Industry** 721
Abdullah Sardogan, Jan Wirsam, Rainer Alt

46 **User Frictions in Process Intelligence Dashboards: Implications for AI Assistants** 737
Candan Cetin, Teodora Lata

47 **Being There: Using Combined Experience Data Sources to Capture Tourist Interactions** 757
Ondrej Mitas, Loïs Camp, Ruben Raidt, Kyle Woosnam, Joël Pardaan, Simone Moretti, Jeroen Klijs, Sait Durgun

48 **Process Model-guided Post-filtering for Wearable Activity Recognition in Ergonomic Work Processes** 771
Andreas Emrich, Alessa Wein, Janaki Viswanathan, Michael Frey, Peter Loos

49 **The Role of Supplier Trust in Enabling the Twin Transition** 787
Alinda Kokkinou, Jan van Elderen

50 **The Need for Deployment Norms in Value Sensitive Data-Sharing Infrastructures** 803
Christel van de Wal, Vanessa Dirksen, Roger Bons, Johan Versendaal

51 **Tenth Anniversary of eHealth at the National Institute of Public Health** 819
Živa Rant, Eva Štupica, Jure Janet, Miha Knavs, Lucija Tepej Jočić, Matjaž Malešič, Eva Hafner, Boško Simeunović, Dijana Čančarević, Simon Indihar, Hajdi Kosednar

52 **A Recent Review on Artificial Intelligence in Ports** 839
Vítor Ribeiro, João Barata, Paulo Rupino da Cunha, Jacinto Estima

53 **Negotiating Usability and Security in Digital Transformation: A Qualitative Study of Everyday Work in Finnish Vocational Education** 855
Anna-Liisa Ojala, Tuomo Sipola, Karo Saharinen

54 **Sales and Operations Planning – Augmenting Existing Maturity Models with New Practices** 869
Henrik J. Nyman, Fabian Tschirch

55 **It Takes Two to Co-Create - Digital Twin Mechanism for Value Co-Creation** 885
Robert Schmelzer

56 **Helpful Enterprise Architecture Elements for Organizational Networks** 901
Arjen Maris, Pascal Ravesteijn, Stijn Hoppenbrouwers, Jos van Hillegersberg

57 **A Framework for Evaluating the Sustainability of IoT Smart City Projects** 917
Hans-Dieter Zimmermann

58 **A Cross-Layer Integration Framework for the Development of a Platform Business Model, Applications and Services** 933
Vjeran Strahonja, Dijana Oreški

59 **Effects of Personality on Perceived Immersion in Low-Immersive Virtual Reality Stores** 949
Markus Makkonen, Lauri Frank, Tiina Kemppainen, Matilda Holkkola, Tiina Kaleton

60 **Expert-Driven Criteria for Assessing SME Readiness for Open Data: Insights from an AI-Supported Brainstorming Workshop** 965
Andrej Škraba, Andreja Pucihar, Staša Blatnik, Nejc Čelik, Franc Lavrič, Radovan Stojanović, Mirjana Kljajić Borštnar

RESEARCH IN PROGRESS

61 **Design and Evaluation of a Digital Transformation Canvas for SME** 983
Lukas Neidlinger, Jürgen Jung

62 **Training Oral Communication-Skills using an Avatar** 993
Nanda van der Stap, Emanuel van Dongen, Stan van Ginkel

63 **Designing Evidence-Based Managerial Decision Support Tools for Older Adult Managers** 1001
Khaled Hassanein, Milena Head

64 **Teaching a Systems Approach to Manage Smart Sustainable Cities** 1009
Quan Zhu, Steven Haveman

65 **The Impact of Data Sharing in Product Recalls: A Cost-Based Approach** 1017
Maarten de Mildt, Sofie Verbrugge, Didier Colle

66 **Constructing a Cost–Benefit Analysis Framework for Digital Twin–Enabled Spatial Planning: Articulating (Hidden) Costs and (Implicit) Benefits in Governmental Organizations** 1025
Sam Leewis, Koen Smit, John van Meerten, Ronald Buijsse

67	**Evaluating the Feasibility of LLM-Based Automation of Manual WCAG Compliance Testing** Annemae van de Hoef, Koen Smit, Sam Leewis, Duain Castro, Fabian Hartman, Juliana Todorova, Nick Kuiper	1035
68	**Subjective Knowledge and AI Privacy Risks** Mari Paalimäki, Juha Munnukka	1043
69	**The Digital Path to Sustainability: A Configurational Investigation** Rohit Nishant, Marie-Claude Boudreau	1051
70	**Frugal Digital Transformation for Resilient Food Banks: A Multiple-Case Study of Distributed Nonprofits** Ousmane Sy, Elaine Mosconi, Marie-Claude Boudreau	1059
71	**Evaluating an AI Act Regulatory Sandbox Framework Against Pilot Cases: Towards a Methodologically Grounded Three-Round Process Framework** Jukka Heikkilä, Altti Lagstedt, Seppo Heikura, Anastasiia Dardykina	1069

DOCTORAL CONSORTIUM

72	**Strategic Adoption of AI-Enabled Decision-Making Systems: Designing for Human Agency in Data-Driven Organizations** Barrak Albabtain	1077
73	**Creating Business Value in Unpredictable Times Aiming for Antifragility** Edzo A. Botjes, Tim Huygh, Laury Bollen, Remko W. Helms	1085
74	**Concept Mining for Data Management** Swarupa Hardikar	1097
75	**Value-Sensitive Design for Responsible Data Sharing in Governmental Ecosystems** Christel van de Wal	1109
76	**Business System Long-Term Viability Across Life-Cycle Phases** Sara Kremsar, Eva Krhač Andrašec, Tomaž Kern	1121
77	**Graph Neural Networks and Deep Reinforcement Learning for Adaptive Warehouse Optimization** Nejc Čelik, Andrej Škraba	1135

78 **Who Will Keep Us Safe? Rethinking Cybersecurity Competencies for AI-Transformed Work in SMEs and Small States** 1141
Andreja Markun

79 **Designing Strategic Planning for Higher Education to Collectively Shape Digital Transformation** 1155
Gül Akcaova

80 **Co-Creating Human-Centred and Responsible AI Gover-Nance in Consulting SMEs** 1169
Ludmila Jiříčková

81 **An Integrated Decision Support Framework for Multi-Phase Flood Event Management** 1181
Vesna Vidmar, Mirjana Kljajić Borštnar

82 **Closing the Loop in AI-Driven Athletic Training Prescription** 1193
Luka Lah, Mirjana Kljajić Borštnar

Exploring Productivity Gains from GenAI-Enabled Virtual Assistants: A Configurational Study of a Multifaceted Phenomenon

Niklas Rieger (born Preiss)
Sheffield Hallam University, Sheffield, United Kingdom of Great Britain and Northern Ireland
niklas.j.preiss@student.shu.ac.uk

Enterprises increasingly embed generative AI–enabled virtual assistants (VA) into enterprise productivity software to improve digital work and performance. Evidence indicates that productivity gains (PG) are a multifaceted phenomenon of organized complexity shaped by interdependent organizational, technological, and human factors. Drawing on established IS lenses, this paper develops a theory-driven model of when and how VA-enabled enterprise productivity software yields PG. This paper deductively derives a research model specifying four IS theory-based conjunctural factors, and articulates equifinal pathways that enable or inhibit PG. To evaluate necessity and sufficiency, the paper outlines an expert interview-based fuzzy-set Qualitative Comparative Analysis research design. The contributions are threefold: (i) defining VA-enabled PG within enterprise productivity software as a multifaceted phenomenon; (ii) introducing a configurational research model; and (iii) extending IS business-value theorizing to AI-augmented digital work through set-theoretic reasoning and theory-driven propositions for empirical validation. The model clarifies when VA enhance enterprise performance and offers a testable roadmap for research on human–AI collaboration.

DOI
https://doi.org/10.18690/um.fov.4.2026.1

ISBN
978-961-299-152-4

Keywords:
generative AI, virtual assistants, enterprise productivity software, productivity gains, configurational theory

University of Maribor Press

1 Introduction

Generative artificial intelligence (GenAI)–powered virtual assistants (VA) are rapidly becoming integral components of enterprise productivity software and are expected to transform the nature of knowledge work (Brynjolfsson et al., 2025; Cambon et al., 2023; Edelman et al., 2023). Commercially available examples, such as Microsoft 365 Copilot and Google Workspace Gemini, exemplify this new class of artificial intelligence (AI)-augmented productivity tools. These VA integrate large language models (LLMs) and multimodal AI capabilities, spanning text, voice, and image, to extend the functionality of traditional productivity applications. Embedding advanced generative and reasoning features directly into core applications (e.g., Word, Excel, PowerPoint) enables content generation, data analysis, summarization, and communication within the respective enterprise productivity software (Bano et al., 2025; Microsoft Corporation, 2023; Vorvoreanu et al., 2025). Early empirical evidence indicates that these tools can yield substantial productivity benefits (Bano et al., 2025; Vorvoreanu et al., 2025). Recent advances in GenAI, and applications built on them, have been hypothesized to deliver step-changes in labor productivity at historically significant scale (Brynjolfsson et al., 2025; Edelman et al., 2023). Initial experiments and field studies corroborate these expectations: GenAI users complete tasks faster and with higher quality than control groups (Cambon et al., 2023; Noy & Zhang, 2023). Despite this promise, realizing productivity gains (PG) from VA cannot be assumed. IS research has long grappled with understanding how IT adoption translates into PG, a debate known as the IT productivity paradox (Brynjolfsson, 1993). This paradox raises the research question:

Under what combinations of organizational, functional, and human factors do GenAI-enabled VA embedded in enterprise productivity software yield PG?

This overarching question is addressed through three sub-questions: (I) How can VA-enabled PG be conceptualized as a multifaceted phenomenon? (II) Which theory-grounded organizational, technological, and human factors combine to produce or inhibit PG? (III) What configurational propositions can guide empirical fsQCA validation? The systematic review by Preiss & Westner (2025a) highlights those gaps, showing that while digital assistant technologies (DAT) hold strong potential to reshape organizational performance and strategy, empirical research remains scarce. DAT are primarily conversational systems, such as chatbots, voice

agents, and intelligent assistants, that leverage machine learning, natural language processing or GenAI, to retrieve information, perform routine tasks, and support decision-making (Preiss & Westner, 2025b). Existing studies emphasize short-term, task-level effects, neglect longitudinal adoption trajectories, and rarely employ established IS theories, resulting in limited theoretical integration and ambiguous findings (Preiss & Westner, 2025a). This gap is particularly salient given the IT productivity paradox, which highlights that technological potential does not automatically translate into realized organizational performance (Brynjolfsson, 1993). Addressing it in the GenAI context requires research that considers not only direct productivity effects but also the configurational interplay of organizational, technological, and human factors over time, thereby capturing equifinal pathways to IT-enabled innovation (Coreynen et al., 2017; Parida et al., 2019; Sjödin et al., 2019). To address this gap, I propose a theory-driven configurational model explaining how GenAI-enabled VA embedded in enterprise productivity software generate PG (hereafter, for readability, "VA-enabled PG"; the full term is used when needed for clarity). The remainder of this paper (1) conceptualizes VA-enabled PG and motivates a configurational perspective; (2) specifies a research model grounded in the core building blocks of configurational theorizing, theoretical multiplicity, factorial logic, and combinatorial logic (Park et al., 2020) and develops initial propositions; and (3) discusses expected contributions and outlines an empirical fsQCA.

2 Theoretical Background: PG as a Configurational Phenomenon

Previous research shows that organizations' ability to realize PG from IT adoption is an important antecedent of business-value creation, competitive performance and broader business transformation (Acharya, 2023; Melville et al., 2004; Storey et al., 2016). In recent years, enterprises have increasingly integrated GenAI-enabled VA into their enterprise productivity software to streamline operations and enhance service delivery (Brynjolfsson et al., 2025; Peng et al., 2023a). VA already play a central role in administrative work (adopted by 37.7% of businesses) and nearly 10% of organizations use them to improve operational workflows (Dutta, 2025). Explaining the mechanisms behind such PG has therefore become highly relevant. Prior IS research suggests that PG accomplished through GenAI-enabled VA constitutes a multifaceted phenomenon that must be analyzed from organizational, technological, and human perspectives (Maharaj et al., 2024; Schmidhuber et al.,

2021; Todericiu, 2025). This multifaceted nature underscores the importance of applying insights from established IS theories to maximize the gains of new technologies. As the review by Preiss & Westner (2025a) found, prior studies on DAT have drawn on frameworks such as the Technology Acceptance Model (TAM), the Unified Theory of Acceptance and Use of Technology (UTAUT), Task–Technology Fit (TTF) or the IS Success Model. Recent work reflects a growing recognition that aligning technology with user needs, task requirements, and organizational context, guided by established IS theories, is essential for realizing the full productivity benefits of VA (Amarilli et al., 2023; Chan & Reich, 2007). Accordingly, contemporary studies on DAT increasingly emphasize real-world PG and tangible outcomes as key measures of success (Afroogh et al., 2025; Celis et al., 2025; Song et al., 2025). Beyond efficiency, quality, and output, additional benefits, including reduced cognitive load, improved well-being, accelerated workplace learning, and enhanced decision quality also sustain PG (Preiss & Westner, 2025a; Tarafdar et al., 2019). PG refers to measurable, attributable gains, such as shorter cycle times, lower error rates, higher throughput, or lower cost in the efficiency, quality, or output arising from the use of VA in enterprise productivity software that leverages expertise, data, collaboration, and AI (Dwivedi et al., 2021; Fang et al., 2025; Noy & Zhang, 2023). These gains are typically contingent on combinations of factors rather than single causes (Park et al., 2020; Ragin & Fiss, 2008). Prior IS research shows that the effects of IT capabilities differ in combination compared to isolation (Chen et al., 2020), and that the joint influence of user acceptance (UTAUT/TAM), TTF (Goodhue & Thompson, 1995), and information/system quality (DeLone & McLean, 2014) conditions whether VA yield PG when integrated into productivity software. This points to conjunctural causation, where specific combinations of organizational, technological, and human factors generate gains (Coreynen et al., 2017; Ragin & Fiss, 2008). IS business-value research identifies multiple paths to achieving PG (Melville et al., 2004; Rai et al., 2012). Early empirical studies on GenAI-enabled tools reinforce this. Noy & Zhang (2023) find that productivity effects vary by task type, while Cambon et al. (2023) show that contextual factors such as collaboration mode and data integration moderate VA effectiveness. These findings underscore equifinality and motivate a configurational, set-theoretic approach (Park et al., 2020; Rihoux & Ragin, 2009). Several factors exhibit asymmetric effects. High TTF appears to function as a foundational/near-necessary enabler of PG, while its absence often precludes gains (Goodhue & Thompson, 1995; Jeyaraj, 2022; Jeyaraj et al., 2006). By contrast, managerial support

or training can amplify productivity when present, but their absence does not always fully preclude improvements (Boonstra, 2013; Thong et al., 1996). Such asymmetries further support configurational theorizing that accommodates both enabling and inhibiting mechanisms (Coreynen et al., 2017; Sjödin et al., 2019). Given these diverse outcomes, I argue that VA-enabled PG is a configurational phenomenon: specific combinations of organizational (Business–IT Alignment (BIA); Top Management Support (TMS)), behavioral (Adoption and Change Enablement (ACE)), and technological/task (TTF) factors jointly determine success (Fiss, 2011; Park et al., 2020; Ragin & Fiss, 2008).

2.1 Defining PG from VA-enabled Enterprise Productivity Software

Organizations seeking PG from DAT face a complex phenomenon that resists explanation by any single theoretical lens (Aral et al., 2012; Benbya et al., 2020; Sarker et al., 2019). Research on IT-enabled PG shows that success is best understood through multiple, interrelated perspectives and emerges when organizational, technological, and human dimensions are considered jointly (Aral et al., 2012; Brachten et al., 2021; Brosig, 2021). VA-enabled PG is defined as organizationally relevant improvements in knowledge-work performance attributable to GenAI-enabled VA over approximately 6–12 months after rollout (Freeman et al., 2024; Sirmon et al., 2007; Vahter & Vadi, 2024). For example, Fang et al. (2025, p. 28) found that a chatbot "can deliver measurable productivity improvements with comparable inputs". Organizations adopt these tools not only for experimentation but also with explicit expectations of realizing PG. Although PG arises from individual and team work practices, the configurational model is calibrated at the organizational level. Accordingly, PG is conceptualized as an organizational-level set indicating whether an organization achieves sustained, substantively meaningful PG from VA-supported knowledge work. Consistent with the organizational unit of analysis, all four factors are conceptualized as organizational-level factors. Although ACE and TTF operate at the user–task interface, they are treated as organizational factors, reflected in aggregated user experiences and workflow integration. The four factors are deliberately specified as complementary and minimally overlapping: BIA captures strategic use-case clarity and business linkage; TTF denotes the degree to which VA capabilities fit target tasks or use cases; ACE comprises user-proximal enablement activities (e.g., training, communication, process adaptation); and TMS denotes top-management legitimation, resourcing, and governance of VA initiatives.

Building on established IS theories, these factors are introduced in section 3, where their theoretical grounding, construct definitions, and indicators are developed in detail. Together, they form a focused yet sufficiently rich set of organizational factors that provides a basis for theorizing configurational "causal recipes" for PG (Park et al., 2020). Accordingly, a configurational perspective, operationalized through methods such as fsQCA, is appropriate to capture three essential ideas: equifinality (multiple viable pathways to PG), conjunctural causality (combinations of factors that jointly produce outcomes), and causal asymmetry (configurations leading to PG need not mirror those leading to its absence) (Pappas & Woodside, 2021; Park et al., 2020). This perspective is particularly relevant for GenAI-based VA, where stochastic model behavior, hallucinations, grounding in organizational knowledge, and privacy/compliance constraints jointly shape realized PG (Acharya, 2023; Bano et al., 2025; Peng et al., 2023b).

3 Research Model and Methods

The four factors draw on IS theories with extensive meta-analytic and review support in IT adoption and business-value research: strategic alignment (Chan & Reich, 2007; Coltman et al., 2015), task–technology fit (Goodhue & Thompson, 1995; Jeyaraj, 2022), technology acceptance (Balakrishnan et al., 2024; Davis, 1985; Venkatesh et al., 2003), and top management support (Boonstra, 2013; Jeyaraj et al., 2023). The structured literature review by Preiss & Westner (2025a) found that, across 61 articles, only four explicitly employed IS theories, underscoring both a gap in theoretical integration and the relevance of these established foundations for the VA context. The four constructs were deductively selected to ensure complementary coverage of organizational (BIA, TMS), behavioral (ACE), and technological/task (TTF) dimensions. VA-enabled PG depends on how organizational, technological, and human factors combine. FsQCA will be suitable for empirical testing (Brosig, 2021; Pappas & Woodside, 2021). Factors and outcomes are conceptualized as sets. Standard notation is used. Prior IS research suggests that the four factors capture complementary yet interdependent conditions influencing how organizations realize PG (Park et al., 2020; Schneider & Wagemann, 2012). Each is modeled as a graded set that combines conjuncturally with the others to explain differences in PG, representing a distinct facet of the broader phenomenon rather than a single net-effect predictor. Each factor is modeled as a condition that reflects differences

within a given variable, i.e., every condition distinguishes between two or more types of the same variable (Schneider & Wagemann, 2012).

BIA is rooted in the Strategic Alignment Model (Henderson & Venkatraman, 1993) and IS business-value research (Chan & Reich, 2007; Melville et al., 2004). BIA captures the degree to which IT initiatives are synchronized with organizational objectives, business processes, and performance metrics. High BIA is present when organizations articulate prioritized VA use cases, connect them to core processes and KPIs, and integrate VA initiatives into business planning (Chan & Reich, 2007; Coltman et al., 2015). Low BIA (~BIA) reflects ad hoc pilots or deployments disconnected from core workflows (Kopper et al., 2020). In organizations exhibiting high BIA, VA are purposefully applied to improve customer responsiveness, reduce coordination costs, or accelerate information processing, for example (Abdellatif et al., 2022; Andrade & Moazeni, 2023; Bird & Lotfi, 2024). In the context of GenAI-enabled VA, BIA is particularly critical. Empirical and practitioner reports show that many organizations are stuck in prolonged "pilot phases," with widespread GenAI experimentation but only a small fraction of use cases delivering measurable business value (Alsheiabni et al., 2019; Hansen et al., 2024). Without such alignment, VA usage tends to remain opportunistic and scattered, making it unlikely that local gains add up to organization-level PG.

TTF posits that technology delivers performance impacts only when its functionalities align with the requirements of users' tasks (Goodhue & Thompson, 1995). For GenAI-enabled VA, TTF comprises functional fit (Bano et al., 2025; Peng et al., 2023b), information quality and reliability, including hallucination risk and error patterns (Helal et al., 2025; Liu et al., 2025) or workflow integration, such as embedding into productivity software or ticketing systems (Acharya, 2023). High TTF is typically present when VA provide accurate, contextually grounded, and timely assistance for key knowledge-work tasks, within the organization's data and security constraints (Bano et al., 2025). Low TTF (~TTF) can reflect missing capabilities, poor grounding in organizational data, limited context, or misaligned integration that generates rework and distrust (Gebauer et al., 2010; Helal et al., 2025). The IS Success Model (DeLone & McLean, 1992, 2014) further serves as a complementary theoretical foundation for this factor, extending the TTF argument by explaining how fit translates into realized outcomes. LLMs are stochastic and prone to hallucinations, plausible but nonfactual content, which raise fundamental

concerns about reliability and trust (Huang et al., 2025). These properties mean that VA outputs may be helpful in some tasks but hazardous in others unless suitably constrained (Brynjolfsson et al., 2025). Emerging empirical work in IS and management shows that PG from GenAI tools are highly task contingent: productivity improvements are largest for tasks where AI knowledge and capabilities align with work demands, and much smaller or even negative when this alignment is weak (Brynjolfsson et al., 2025). Studies of human–AI collaboration likewise find that organizing work based on TTF standards is crucial for realizing PG (Przegalinska et al., 2025).

ACE reflects the behavioral and organizational mechanisms that translate technological potential into realized value. Drawing from TAM (Davis, 1985), the UTAUT (Venkatesh et al., 2003), and organizational change theory (Lapointe & Rivard, 2005), this factor captures the organization's ability to facilitate employee adoption and adapt work practices around new technologies. In this model, ACE refers to user-proximal mechanisms that support the adoption, routinization, and effective use of VA. It includes activities such as communication about VA capabilities, limitations, and appropriate use or user training and coaching (e.g., prompt design, verification strategies). High ACE means organizations actively onboard users, provide ongoing support, and adapt workflows based on feedback (Alter, 2018; Söllner et al., 2025). In contrast, low ACE (~ACE) denotes resistance to change, lack of managerial communication, or insufficient training, barriers that undermine integrating VA and attenuate PG (Lapointe & Rivard, 2005; Venkatesh et al., 2003). GenAI-enabled VA require active, skillful use, including formulating effective prompts, interpreting probabilistic outputs, and integrating those outputs into existing work routines. Empirical studies and large surveys indicate that, even when GenAI tools are technically available, adoption and value often remain limited because employees do not see clear, localized use cases, feel uncertain about how to use the tools, or fear making mistakes (Söllner et al., 2025). Recent work on the responsible use of GenAI stresses the need for training, guidelines, and human-in-the-loop oversight to ensure that workers understand both the capabilities and the limitations of GenAI tools (Söllner et al., 2025). Without user-centered enablement, organizations risk either shadow AI usage that bypasses controls or underutilization of enterprise-grade VA solutions (Silic et al., 2025).

TMS captures legitimation and resourcing of VA initiatives at the top-management level. Grounded in organizational control and IT implementation research (Boonstra, 2013; Thong et al., 1996), TMS denotes the extent to which senior executives actively champion, resource, and monitor IT and especially AI initiatives. In the case of VA, strong TMS ensures that the technology is institutionally legitimized, adequately funded, and integrated into strategic and operational priorities. High TMS is typically present when senior executives articulate a VA vision, set expectations for productivity improvements, allocate budgets and staff, and integrate VA initiatives into governance structures (Boonstra, 2013; Jeyaraj et al., 2023). Low TMS (~TMS) can occur when VA initiatives lack sponsorship, remain side projects, or are constrained by ambiguous mandates and limited resources (Korzyński et al., 2024). In the GenAI context, TMS is critical for two main reasons. First, GenAI introduces new risk, security, and compliance exposures, such as data leakage, inadvertent disclosure of personal or confidential information, and violations of regulations. Enterprise evidence shows that employees often paste sensitive data into public GenAI tools via unmanaged accounts, creating serious compliance and security risks without proper governance and technical controls (Dong et al., 2024; Korzyński et al., 2024; J. Yang et al., 2024). Second, GenAI programs usually begin as broad experimentation but must be refocused on a few high-value, feasible use cases once firms understand the technology's strengths and limits. Such strategic reprioritization, and the associated decisions on budgets, data access, and risk appetite, requires active TMS (Korzyński et al., 2024).

4 Configurational Propositions

Building on the factors defined in the previous section, this section develops the combinatorial logic that links BIA, TTF, ACE, and TMS to VA-enabled PG. In line with Park et al. (2020), this final step articulates causal recipes that explain how the selected factors, modeled as fuzzy-set conditions, jointly produce the outcome PG. Moving beyond factorial logic, causality is treated as conjunctural and equifinal: distinct configurations of conditions can yield the same outcome in complex, nonlinear settings (Ragin & Fiss, 2008). Moreover, multiple distinct combinations, or equifinal pathways, may lead to the same outcome, underscoring the complex and nonlinear nature of IT-enabled PG (Fiss, 2011; Park et al., 2020). Figure 1 presents the proposed research model, which explains how organizational-level conditions combine to produce PG. The paper advances with theoretically grounded

propositions that specify how configurations of BIA, TMS, ACE, and TTF lead to PG. Each is stated in Boolean notation and provides the groundwork for fsQCA validation (Kohtamäki et al., 2019; Shulman, 2008). Building on Brosig et al. (2022), a set-theoretic ontology is adopted, emphasizing causal asymmetry: the presence and absence of PG may arise from different configurations. Propositions are framed as truth-table configurations, conjunctions expected to be sufficient for PG, enabling empirical evaluation via fsQCA consistency and coverage (Schneider & Wagemann, 2012). This configurational approach extends IT-performance frameworks (Melville et al., 2004; Wade & Hulland, 2004) by foregrounding interactive effects among the four factors. It advances digital business-value theory toward a complexity-sensitive account of GenAI's impact, where adaptive human–AI collaboration resists explanation by isolated capabilities.

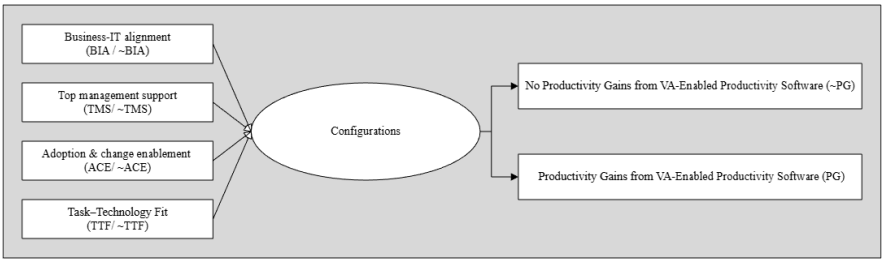

Figure 1: Preliminary research model on VA-enabled PG
Source: Own

4.1 Potential Foundational Conditions

Prior IS research suggests that certain conditions frequently behave as foundational enablers of PG, irrespective of other contextual factors (Brynjolfsson & Hitt, 2000; Melville et al., 2004). Both BIA and TTF are expected to be foundational enabling conditions for PG. In set-theoretic terms, PG cases are expected to be near-subsets of high-BIA and high-TTF; that is, BIA and TTF may behave approximately as necessary conditions. Future studies can test whether PG is consistently a subset of BIA and TTF using standard necessity consistency and coverage indicators, rather than imposing necessity ex ante (Ragin & Fiss, 2008; Schneider & Wagemann, 2012). This approach helps avoid premature claims of necessity and enhances the transparency of empirical validation. While necessary factors indicate factors that must be present for PG to occur, they do not reveal how outcomes materialize.

Configurational theorizing therefore extends to sufficiency analysis, which examines how distinct combinations of factors jointly produce PG. This step recognizes the asymmetrical nature of causality: the presence of a factor may be essential but not, by itself, sufficient. The following propositions thus specify equifinal configurations of BIA, TMS, ACE, and TTF that together form alternative, sufficient pathways to PG. These are empirical expectations, not hard constraints; necessity will be assessed via fsQCA rather than assumed ex ante.

Potential Condition C1 (PG → BIA) expects that organizations with high VA-enabled PG will almost always exhibit high BIA. In other words, PG cases are likely to be (near-)subsets of the set of BIA cases. Strategic alignment between IT and business objectives ensures that digital initiatives are directed toward organizationally relevant use cases and that technological resources are purposefully applied to value-creating activities (Chan & Reich, 2007; Coltman et al., 2015). Without such alignment, IT efforts, including VA deployment, tend to become fragmented and fail to yield sustainable productivity improvements (Kopper et al., 2020; Silic et al., 2025). Hence, high BIA is expected to act as a foundational enabling condition for realizing PG. Potential Condition C2 (PG → TTF) posits that organizations with high VA-enabled PG will almost always exhibit high TTF. Consistent with Goodhue & Thompson's (1995) foundational model, productivity improvements arise only when technological features align with the cognitive and procedural demands of users' tasks. Without sufficient fit, even well-supported and strategically aligned systems fail to translate into tangible PG (Gebauer et al., 2010; Gebauer & Ginsburg, 2006). Therefore, TTF is expected to function as an enabling condition for realizing PG.

4.2 Potential Causal Configurations for PG

Proposition P1 (BIA*TMS*TTF→PG) is characterized by (high) BIA, (strong) TMS, and (high) TTF. Indeed, meta-analytic evidence shows TMS drives IT assimilation, which in turn yields business-value. Jeyaraj et al. (2023) find that strong executive backing significantly boosts IT adoption and its ultimate impact on PG. Combined with a well-fitted VA (high TTF), these factors suggest a "leadership-driven" path to productivity. Proposition P2 (BIA*ACE*TTF→PG) is the alternative configuration of (high) BIA, (extensive) ACE, and (high) TTF suggesting a user-centered success route. Ifinedo (2008) found that business vision is a

significant factor in achieving success, which suggests that clarity of vision and business context matter for technology initiatives (part of change management). These configurations reflect equifinality: organizations may achieve PG via BIA, TMS, and TTF (P1) or via BIA, TTF, and ACE (P2).

4.3 Potential Causal Configurations for Limited PG

Proposition P3 (~BIA*TTF→~PG) states that with high TTF, organizations with low BIA are expected to fail to realize substantial PG. Without clear linkage to core processes and goals, VA use remains fragmented and its effects difficult to aggregate (Jeyaraj et al., 2023). Proposition P4 (BIA*~TTF→~PG) states that organizations with high BIA but ~TTF are expected to experience limited or absent PG. Strategically chosen use cases may nonetheless underperform if VA functionality, information quality, or grounding in organizational data are insufficient (Freeman et al., 2024). Proposition P5 (BIA*TTF*~TMS*~ACE→~PG) indicates a situation where initiatives are aligned (BIA) and VA are technically capable (TTF), but weak TMS and weak ACE are expected to prevent PG. VA use remains sporadic or shadow-like and does not translate into sustained PG.

5 Discussion and Conclusion

VA-enabled PG is conceptualized as a multifaceted outcome of organized complexity. By shifting from a variance-based logic to a configurational perspective, we contribute to IS business-value theory in three ways. First, we provide a definition of PG that accounts for the multifaceted nature of GenAI-driven impact. Second, we introduce a parsimonious model of four interdependent factors (BIA, TTF, ACE, and TMS) that captures the "causal recipes" of VA-enabled work. Finally, the proposed equifinal propositions (P1–P5) offer a robust theoretical foundation for fsQCA validation. These insights provide immediate utility for practitioners seeking to move GenAI initiatives towards more sustained organizational PG. To empirically validate the theorized propositions, a multi-stage fsQCA study is proposed. By identifying necessary and sufficient conditions through truth table analysis, this design will reveal the causal asymmetries inherent in GenAI adoption, specifically how the absence of one factor may necessitate different configurations of the remaining conditions to achieve the same level of PG. As a conceptual paper, the propositions represent testable expectations rather than confirmed findings; this

limitation is inherent to the theory-building stage of configurational research (Park et al., 2020) and is addressed through the outlined fsQCA design. By combining configurational reasoning with set-theoretic analysis, the model extends IS business-value theory to the GenAI era and offers a foundation for future empirical research on how VA factor sets jointly determine organizational performance.

References

Abdellatif, A., Badran, K., Costa, D. E., & Shihab, E. (2022). A comparison of natural language understanding platforms for chatbots in software engineering. *IEEE Transactions on Software Engineering*, *48*(8), 3087–3102. Scopus. https://doi.org/10.1109/TSE.2021.3078384

Acharya, V. V. (2023). The impact of Microsoft 365 Copilot on small and medium-sized businesses. *JETIR*, *10*(4), Article 4. https://www.jetir.org/view?paper=JETIR2304546

Afroogh, S., Varshney, K. R., & D'Cruz, J. (2025). *A task-driven human-AI collaboration: When to automate, when to collaborate, when to challenge* (arXiv:2505.18422). arXiv. https://doi.org/10.48550/arXiv.2505.18422

Alsheiabni, S., Cheung, Y., & Messom, C. (2019). Factors inhibiting the adoption of artificial intelligence at organizational-level: A preliminary investigation. *AMCIS 2019 Proceedings*, 2. https://research.monash.edu/en/publications/factors-inhibiting-the-adoption-of-artificial-intelligence-at-org/

Alter, S. (2018). A systems theory of IT innovation, adoption, and adaptation. *Research Papers*, (26). https://aisel.aisnet.org/ecis2018_rp/26

Amarilli, F., van den Hooff, B., & van Vliet, M. (2023). Business-IT alignment as a coevolution process: An empirical study. *The Journal of Strategic Information Systems*, *32*(2), 101776. https://doi.org/10.1016/j.jsis.2023.101776

Andrade, R., & Moazeni, S. (2023). Transfer rate prediction at self-service customer support platforms in insurance contact centers. *Expert Systems with Applications*, *212*. Scopus. https://doi.org/10.1016/j.eswa.2022.118701

Aral, S., Brynjolfsson, E., & Van Alstyne, M. (2012). Information, technology, and information worker productivity. *Information Systems Research*, *23*(3-part-2), 849–867. https://doi.org/10.1287/isre.1110.0408

Balakrishnan, J., Dwivedi, Y. K., Hughes, L., & Boy, F. (2024). Enablers and Inhibitors of AI-Powered Voice Assistants: A Dual-Factor Approach by Integrating the Status Quo Bias and Technology Acceptance Model. *Information Systems Frontiers*, *26*(3), 921–942. Scopus. https://doi.org/10.1007/s10796-021-10203-y

Bano, M., Zowghi, D., Whittle, J., Zhu, L., Reeson, A., Martin, R., & Parsons, J. (2025). *A qualitative study of user perception of M365 AI copilot* (arXiv:2503.17661). arXiv. https://doi.org/10.48550/arXiv.2503.17661

Benbya, H., Nan, N., Tanriverdi, H., & Yoo, Y. (2020). Complexity and information systems research in the emerging digital world. *Management Information Systems Quarterly*, *44*(1), 1–18. https://doi.org/10.25300/MISQ/2020/13304

Bird, J. J., & Lotfi, A. (2024). Customer service chatbot enhancement with attention-based transfer learning. *Knowledge-Based Systems*, *301*, 1–12. Scopus. https://doi.org/10.1016/j.knosys.2024.112293

Boonstra, A. (2013). How do top managers support strategic information system projects and why do they sometimes withhold this support? *International Journal of Project Management*, *31*(4), 498–512. https://doi.org/10.1016/j.ijproman.2012.09.013

Brachten, F., Kissmer, T., & Stieglitz, S. (2021). The acceptance of chatbots in an enterprise context – A survey study. *International Journal of Information Management*, *60*. Scopus. https://doi.org/10.1016/j.ijinfomgt.2021.102375

Brosig, C. (2021). Manufacturers' IT-enabled service innovation success as a multifaceted phenomenon: A configurational study. *ICIS 2021 Proceedings*. https://aisel.aisnet.org/icis2021/governance/governance/6

Brosig, C., Bley, K., Strahringer, S., & Westner, M. (2022). The missing piece – calibration of qualitative data for qualitative comparative analyses in IS research. *ECIS 2022 Research Papers*, (37), 1–16.

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communications of the ACM*, *36*(12), 66–77. https://doi.org/10.1145/163298.163309

Brynjolfsson, E., & Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, *14*(4), 23–48. https://doi.org/10.1257/jep.14.4.23

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, *140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Cambon, A., Hecht, B., Edelman, B., Ngwe, D., Jaffe, S., Heger, A., Vorvoreanu, M., Peng, S., Hofman, J., Farach, A., Bermejo-Cano, M., Knudsen, E., Bono, J., Sanghavi, H., Spatharioti, S., Rothschild, D., Goldstein, D. G., Kalliamvakou, E., Cihon, P., … Teevan, J. (2023). *Early LLM-based tools for enterprise information workers likely provide meaningful boosts to productivity* (MSR-TR-2023-43). Microsoft. https://www.microsoft.com/en-us/research/publication/early-llm-based-tools-for-enterprise-information-workers-likely-provide-meaningful-boosts-to-productivity/

Celis, L. E., Huang, L., & Vishnoi, N. K. (2025). *A mathematical framework for AI-Human integration in work* (arXiv:2505.23432). arXiv. https://doi.org/10.48550/arXiv.2505.23432

Chan, Y. E., & Reich, B. H. (2007). IT alignment: What have we learned? *Journal of Information Technology*, *22*(4), 297–315. https://doi.org/10.1057/palgrave.jit.2000109

Chen, M., Chen, Y., Liu, H., & Xu, H. (2020). Influence of information technology capability on service innovation in manufacturing firms. *Industrial Management & Data Systems*, *121*(2), 173–191. https://doi.org/10.1108/IMDS-04-2020-0218

Coltman, T., Tallon, P., & Sharma, R. (2015). Strategic it alignment: Twenty-five years on. *Journal of Information Technology*, *30*(2), 91–100. https://doi.org/10.1057/jit.2014.35

Coreynen, W., Matthyssens, P., & Van Bockhaven, W. (2017). Boosting servitization through digitization: Pathways and dynamic resource configurations for manufacturers. *Industrial Marketing Management*, *60*, 42–53. https://doi.org/10.1016/j.indmarman.2016.04.012

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* [Thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/15192

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95.

DeLone, W. H., & McLean, E. R. (2014). The Delone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

Dong, L., Neufeld, D., & Higgins, C. (2024). Top Management Support of Enterprise Systems Implementations. *Journal of Information Technology*, *24*(1), 55–80. https://doi.org/10.1057/jit.2008.21

Dutta, P. (2025, July 1). Virtual Assistant Statistics By Adoption and Facts (2025). *Market.Biz*. https://market.biz/virtual-assistant-statistics/

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., … Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and

agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Edelman, B. G., Ngwe, D., & Peng, S. (2023). *Measuring the impact of AI on information worker productivity* (SSRN Scholarly Paper No. 4648686). Social Science Research Network. https://doi.org/10.2139/ssrn.4648686

Fang, L., Yuan, Z., Zhang, K., Donati, D., & Sarvary, M. (2025). *Generative AI and firm productivity: Field experiments in online retail* (arXiv:2510.12049). arXiv. https://doi.org/10.48550/arXiv.2510.12049

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, *54*(2), 393–420. https://doi.org/10.5465/amj.2011.60263120

Freeman, B. S., Arriola, K., Cottell, D., Lawlor, E., Erdman, M., Sutherland, T., & Wells, B. (2024). *Evaluation of task specific productivity improvements using a generative artificial intelligence personal assistant tool* (arXiv:2409.14511). arXiv. https://doi.org/10.48550/arXiv.2409.14511

Gebauer, J., & Ginsburg, M. (2006). Exploring the black box of task-technology fit: The case of mobile information systems. *Working Papers, Working Papers*, Article 06–0120. https://ideas.repec.org//p/ecl/illbus/06-0120.html

Gebauer, J., Shaw, M. J., & Gribbins, M. L. (2010). Task-Technology fit for mobile information systems. *Journal of Information Technology*, *25*(3), 259–272. https://doi.org/10.1057/jit.2010.10

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology fit and individual performance. *MIS Quarterly*, *19*(2), 213–236. https://doi.org/10.2307/249689

Hansen, H. F., Lillesund, E., Mikalef, P., & Altwaijry, N. (2024). Understanding artificial intelligence diffusion through an AI capability maturity model. *Information Systems Frontiers*, *26*(6), 2147–2163. https://doi.org/10.1007/s10796-024-10528-4

Helal, M. Y. I., Elgendy, I. A., Albashrawi, M. A., Dwivedi, Y. K., Al-Ahmadi, M. S., & Jeon, I. (2025). The impact of generative AI on critical thinking skills: A systematic review, conceptual framework and future research directions. *Information Discovery and Delivery*. https://doi.org/10.1108/IDD-05-2025-0125

Henderson, J. C., & Venkatraman, H. (1993). Strategic alignment: Leveraging information technology for transforming organizations. *IBM Systems Journal*, *38*(32), 4–16. https://doi.org/10.1147/SJ.1999.5387096

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, *43*(2), 1–55. https://doi.org/10.1145/3703155

Ifinedo, P. (2008). Impacts of business vision, top management support, and external expertise on ERP success. *Business Process Management Journal, Business Process Management Journal. - Emerald Group Publishing Limited, ISSN 1758-4116, ZDB-ID 2014421-0. - Bd. 14.2008, 4, S. 551-568*, *14*(4).

Jeyaraj, A. (2022). A meta-regression of task-technology fit in information systems research. *International Journal of Information Management*, *65*, 102493. https://doi.org/10.1016/j.ijinfomgt.2022.102493

Jeyaraj, A., Roberts, N., & Gerow, J. (2023). Assessing the connections among top management support, IT assimilation, and business value of IT: A meta-analysis. *Journal of the Association for Information Systems Preprints*, *24*(1), 107–135. https://doi.org/10.17705/1jais.00772

Jeyaraj, A., Rottman, J., & Lacity, M. (2006). A review of the predictors, linkages, and biases in IT innovation adoption research. *Journal of Information Technology*, *21*(1), 1–23. https://doi.org/10.1057/palgrave.jit.2000056

Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., & Baines, T. (2019). Digital servitization business models in ecosystems: A theory of the firm. *Journal of Business Research*, *104*, 380–392. https://doi.org/10.1016/j.jbusres.2019.06.027

Kopper, A., Westner, M., & Strahringer, S. (2020). From shadow IT to business-managed IT: A qualitative comparative analysis to determine configurations for successful management of IT

by business entities. *Information Systems and E-Business Management*, *18*(2), 209–257. https://doi.org/10.1007/s10257-020-00472-6

Korzyński, P., Silva, S. C. e, Górska, A. M., & Mazurek, G. (2024). Trust in AI and top management support in Generative-AI adoption. *Journal of Computer Information Systems*, *0*(0), 1–15. https://doi.org/10.1080/08874417.2024.2401986

Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *MIS Quarterly*, *29*(3), 461–491. https://doi.org/10.2307/25148692

Liu, M., Okuhara, T., Shirabe, R., Nishiie, Y., Xu, Y., Okada, H., & Kiuchi, T. (2025). Evaluating the reliability and accuracy of an AI-powered search engine in providing responses on dietary supplements: Quantitative and qualitative evaluation. *JMIR AI*, *4*, e78436. https://doi.org/10.2196/78436

Maharaj, A. V., Qian, K., Bhattacharya, U., Fang, S., Galatanu, H., Garg, M., Hanessian, R., Kapoor, N., Russell, K., Vaithyanathan, S., & Li, Y. (2024). *Evaluation and Continual Improvement for an Enterprise AI Assistant* (arXiv:2407.12003). arXiv. https://doi.org/10.48550/arXiv.2407.12003

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organizational performance: an integrative model of it business value. *MIS Q.*, *28*(2), 283–322.

Microsoft Corporation. (2023). *Introducing Microsoft 365 Copilot: Your copilot for work*. https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science (New York, N.Y.)*, *381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

Pappas, I. O., & Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsQCA): Guidelines for research practice in information systems and marketing. *International Journal of Information Management*, *58*, 102310. https://doi.org/10.1016/j.ijinfomgt.2021.102310

Parida, V., Sjödin, D., & Reim, W. (2019). Reviewing literature on digitalization, business model innovation, and sustainable industry: Past achievements and future promises. *Sustainability*, *11*(2), 391. https://doi.org/10.3390/su11020391

Park, Y., Fiss, P., & Sawy, O. E. (2020). Theorizing the multiplicity of digital phenomena: The ecology of configurations, causal recipes, and guidelines for applying QCA. *Management Information Systems Quarterly*, *44*(4), 1493–1520.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023a). *The impact of AI on developer productivity: Evidence from github copilot* (arXiv:2302.06590). arXiv. https://doi.org/10.48550/arXiv.2302.06590

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023b, February 13). *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot*. arXiv.Org. https://arxiv.org/abs/2302.06590v1

Preiss, N., & Westner, M. (2025a). From agents to copilots: A systematic review of digital assistant technology adoption in proprietary productivity software. *2025 20th Conference on Computer Science and Intelligence Systems (FedCSIS)*, 565–576. https://doi.org/10.15439/2025F3271

Preiss, N., & Westner, M. (2025b). Unraveling jingle-jangle fallacies in digital assistant technologies: A comprehensive systematic review and research agenda. In X.-S. Yang, S. Sherratt, N. Dey, & A. Joshi (Eds.), *Proceedings of Tenth International Congress on Information and Communication Technology* (pp. 199–227). Springer Nature. https://doi.org/10.1007/978-981-96-6438-2_15

Przegalinska, A., Triantoro, T., Kovbasiuk, A., Ciechanowski, L., Freeman, R. B., & Sowa, K. (2025). Collaborative AI in the workplace: Enhancing organizational performance through resource-based and task-technology fit perspectives. *International Journal of Information Management*, *81*, 102853. https://doi.org/10.1016/j.ijinfomgt.2024.102853

Ragin, C., & Fiss, P. (2008). Net effects analysis versus configurational analysis: An empirical demonstration. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*, *15*(4). https://www.scirp.org/reference/referencespapers?referenceid=3884041

Rai, A., Pavlou, P. A., Im, G., & Du, S. (2012). Interfirm IT capability profiles and communications for cocreating relational value: Evidence from the logistics industry. *MIS Quarterly*, *36*(1), 233–262. https://doi.org/10.2307/41410416

Rihoux, B., & Ragin, C. (2009). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vol. 51). SAGE Publications, Inc. https://doi.org/10.4135/9781452226569

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance. *Management Information Systems Quarterly*, *43*(3), 695–719.

Schmidhuber, J., Schlögl, S., & Ploder, C. (2021). Cognitive load and productivity implications in human-chatbot interaction. *2021 IEEE 2nd International Conference on Human-Machine Systems (ICHMS)*, 1–6. https://doi.org/10.1109/ICHMS53169.2021.9582445

Schneider, C. Q., & Wagemann, C. (2012). *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*. Cambridge University Press. https://doi.org/10.1017/CBO9781139004244

Shulman, M. A. (2008). *Set theory for category theory* (arXiv:0810.1279). arXiv. https://doi.org/10.48550/arXiv.0810.1279

Silic, M., Silic, D., & Kind-Trüller, K. (2025). From shadow It to shadow AI–threats, risks and opportunities for organizations. *Strategic Change*, *n/a*(n/a). https://doi.org/10.1002/jsc.2682

Sirmon, D. G., Hitt, M. A., & Ireland, R. D. (2007). Managing firm resources in dynamic environments to create value: Looking inside the black box. *Academy of Management Review*, *32*(1), 273–292. https://doi.org/10.5465/amr.2007.23466005

Sjödin, D., Parida, V., & Kohtamäki, M. (2019). Relational governance strategies for advanced service provision: Multiple paths to superior financial performance in servitization. *Journal of Business Research*, *101*, 906–915. https://doi.org/10.1016/j.jbusres.2019.02.042

Söllner, M., Arnold, T., Benlian, A., Bretschneider, U., Knight, C., Ohly, S., Rudkowsi, L., Schreiber, G., & Wendt, D. (2025). ChatGPT and beyond: Exploring the responsible use of generative AI in the workplace. *Business & Information Systems Engineering*, *67*(2), 289–303.

Song, Y., Qiu, X., & Liu, J. (2025). The impact of artificial intelligence adoption on organizational decision-making: An empirical study based on the technology acceptance model in business management. *Systems*, *13*(8), 683. https://doi.org/10.3390/systems13080683

Storey, C., Cankurtaran, P., Papastathopoulou, P., & Hultink, E. J. (2016). Success factors for service Innovation: A meta-analysis. *Journal of Product Innovation Management*, *33*(5), 527–548. https://doi.org/10.1111/jpim.12307

Tarafdar, M., CL, C., & Stich, J.-F. (2019). The technostress trifecta - techno eustress, techno distress and design: Theoretical directions and an agenda for research. *Info Systems J.*, *29*, 6–42. https://doi.org/10.1111/isj.12169

Thong, J. Y. L., Yap, C.-S., & Raman, K. S. (1996). Top management support, external expertise and information systems implementation in small businesses. *Information Systems Research*, *7*(2), 248–267. https://doi.org/10.1287/isre.7.2.248

Todericiu, I. A. (2025). Virtual assistants: A Review of the next frontier in AI interaction. *Acta Universitatis Sapientiae, Informatica*, *17*(1), 1. https://doi.org/10.1007/s44427-025-00002-7

Vahter, P., & Vadi, M. (2024). The relationship of technological and organizational innovation with firm performance: Opening the black box of dynamic complementarities. *Technological Forecasting and Social Change*, *206*, 123516. https://doi.org/10.1016/j.techfore.2024.123516

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). *User acceptance of information technology: Toward a unified view* (SSRN Scholarly Paper No. 3375136). Social Science Research Network. https://papers.ssrn.com/abstract=3375136

Vorvoreanu, M., Graham, S., Heger, A., Dhanorkar, S., & Walker, K. (2025). *New employee Copilot usage: Insights into productivity and socialization* (MSR-TR-2025-24). Microsoft. https://www.microsoft.com/en-us/research/publication/new-employee-copilot-usage-insights-into-productivity-and-socialization/

Wade, M., & Hulland, J. (2004). The resource-based view and information systems research: Review, extension and suggestions for future research. *Management Information Systems Quarterly*, *28*(1), 107–142. https://doi.org/10.2307/25148626

Yang, J., Blount, Y., & Amrollahi, A. (2024). Artificial intelligence adoption in a professional service industry: A multiple case study. *Technological Forecasting and Social Change*, *201*, 123251. https://doi.org/10.1016/j.techfore.2024.123251

DOI
https://doi.org/
10.18690/um.fov.4.2026.2

ISBN
978-961-299-152-4

Trust and Student Adoption of AI-Enhanced Augmented Reality in Daily Life: A PLS-SEM and fsQCA Investigation

Bad Re Alam, Anssi Öörni, Jozsef Mezei,
Fatima Bilal
Åbo Akademi University, Turku, Finland
badre.alam@abo.fi, anssi.oorni@abo.fi, jozsef.mezei@abo.fi, fatima.bilal@abo.fi

Purpose: AI-powered augmented reality (AI-AR) technologies are increasingly used in everyday consumer activities, yet the factors influencing their adoption remain underexplored. This study examines how perceived usefulness, perceived risk, and system interactivity affect users' intention to use AI-AR daily, and whether trust in AI mediates these relationships. Design/methodology/approach: An integrated research model grounded in the Stimulus–Organism–Response (S-O-R) framework was proposed and tested using survey data from general technology users across multiple countries. PLS-SEM and fsqca were employed to analyze the relationships among the constructions. Findings: The results show that perceived usefulness and system interactivity positively affect the intention to use AI-AR, whereas perceived risk has an adverse effect. Furthermore, trust in AI significantly mediates the impacts of these factors on adoption intention. Originality/value: The study contributes to literature by introducing an S-O-R-based theoretical framework integrating trust and risk into technology acceptance models. The findings provide practical insights for designing AI-AR applications (in domains such as retail, health, and navigation) that foster user trust, emphasize useful interactive features, and address risk concerns to enhance consumer adoption in daily life.

Keywords:
trust,
AI,
augmented reality,
student Adoption,
AI-powered

University of Maribor Press

1 Introduction

Artificial intelligence (AI) and augmented reality (AR) are changing the lives of people in their daily routine in different fields, with applications including navigation, commerce, entertainment, design, and social interaction (Rauschnabel et al., 2022). AR is a technology that superimposes digital information, such as images, sounds, and other data, onto the real-world environment in real-time(Alkaeed et al., 2024). It enhances the user's thinking style of the physical world by overlaying computer-generated or AI sensory changes, such as graphics, audio, and haptic feedback(Flavián et al., 2019). The integration of AI-powered AR allows users to add intelligent digital information to their actual environments, facilitating seamless and engaging experiences in everyday activities (Dwivedi et al., 2023).

In previous studies, researchers have used technology acceptance models, such as TAM and UTAUT, which highlight usefulness and ease of use, although they often ignore trust and perceived risk(Kenesei et al., 2025), even though concerns such as, algorithmic bias, data leak, or over-reliance on AI increase perceived risk and may deter use(Featherman & Pavlou, 2003). Moreover, Current research has not examined how trust in AI mediates the effects of usefulness, risk, or interaction on use intentions (Nazaretsky et al., 2025). Few studies have included system interactivity (a fundamental aspect of AR); thus, we don't know how an interactive AR system could affect its perceived usefulness, trust, and intention to use (Baleghi-Zadeh et al., 2017). To fill these gaps, we are answering these questions.

This study adopts the Stimulus–Organism–Response (S-O-R) framework as its theoretical foundation, a model first introduced by Mehrabian and Russell (1974), which explains how external environmental stimuli influence individuals thinking' internal states and, ultimately, their behavioral responses of adoption (Liu et al., 2023). In the context of our study, we define stimuli as the system's technical features which include system interactivity, organism as user perceptions (trust, risk, and usefulness), and response as the intention to adopt AI-AR. Based on this framework, we propose the following research questions: *1. How do perceived usefulness, system interactivity, and perceived risk influence students' intention to adopt AI-powered augmented reality (AI–AR)? 2. Does trust in AI, as the Trust–Risk framework explains, mediate the relationships between these factors and students' intention to use AI–AR? 3. Which of the identified factors is the most influential in shaping students' intention to use AI-AR?*

Furthermore, this study adds trust in AI and perceived risk to technology adoption models. It proposes that trust in the AI system mediates the effect of perceived usefulness, risk, and interaction on intention to use in student adoption of AI-enhanced AR technology(Hmoud et al., 2025). This integrated approach addresses current recommendations for AI adoption models to prioritize trust and consider risk perceptions as possible barriers(Rana et al., 2024). The research extends AI-AR use knowledge using a multidisciplinary framework covering technological acceptance and trust-risk theories. The results will address a gap in AR research by demonstrating how usefulness, interaction, risk, and trust interact. The suggested model is experimentally evaluated using survey data from AI-AR users in various countries, and it is assessed by partial least squares structural equation modeling (PLS-SEM) and fsQca.

2 Theoretical background

The stimulus-organism-response (S-O-R) framework proposed by Mehrabian and Russell (1974) is one of the most prominent models in environmental psychology(Manthiou et al., 2017), which explains the mechanism of the influence of external stimuli on people's behavior(Jacoby, 2002). It describes a process where an external environmental factor (stimulus) influences consumers' internal state (organism), which consequently results in their approach or avoidance behaviour (response) (Wang et al., 2023). Second, an organism encompasses an individual's internal state and assessment of the stimulus, encompassing cognitive and emotional perception (Manthiou et al., 2017). Chan and Li have proposed that two types of stimuli arouse consumers' cognition and affection, namely external and internal stimuli, as well as two types of organisms, namely cognitive and affective reactions (Chan et al., 2017). Third, response encompasses the behavioral outcomes of individuals, comprising both their intentions and actual actions(Shang et al., 2023). The S-O-R framework assumes that external stimuli influence individuals' cognitive and emotional appraisal, which in turn triggers their positive or negative reaction behaviors. Researchers of late have started to employ the S-O-R framework in explaining mobile consumer behaviour (Li et al., 2012). It has been widely used in human-machine interaction as a comprehensive model for elucidating the impact of technological qualities on a person's behavior(Leung et al., 2022). Comprehending how people assess technological characteristics may significantly influence their adoption and retention of technology(Ashfaq et al., 2021). A study applied the S-O-

R paradigm to analyse the effect of mobile instant messaging on consumers' electronic word of mouth (e-WOM) towards an online retailer in China(Vazquez et al., 2017). In another study have empirically tested and validated an integrative model to enlist various drivers of mobile shoppers repurchase intention(Rodríguez-Torrico et al., 2019). In the fields of intelligent services, stimuli refer to the technological attributes of intelligent goods, and users engage with these products via their interactive functionalities to fulfill their service requirements(Basha et al., 2022).

In our study, system interactivity serves as the stimuli that influence internal perceptions. Perceived usefulness, perceived risk, and trust in AI represent the organism (users' internal states), shaping the response to the intention to use AI-AR technology. In our study AI-AR usage behaviour is the Response. Recent studies have used trust and risk as mediators. For example, research used S-O-R to illustrate how a system's characteristics (including interactive cues and privacy risk) affect users' trust, which drives adoption intentions (Yang et al., 2025). Our model understands that students will weigh the perceived usefulness of the AI-AR system against perceived risks, and that their trust in the AI will tilt the balance in favour of adoption. A recent study found that trust, risk, and usefulness must be addressed to explain user behaviour in AI-driven systems fully (Yu et al., 2025).

3 Literature review and hypotheses development

3.1 System Interactivity

System interactivity first denotes a fundamental quality of face-to-face interaction in communication science(Chen & Yen, 2004), which has evolved into an essential feature of the Internet and technology context(Horning, 2017). Interactivity has been extensively used in the fields of communication and marketing over the last decades; nevertheless, experts have yet to achieve an agreement on its definition(Shao & Chen, 2021). This construct is primarily described from two perspectives: technical attribute-based interactivity and perception-based interactivity(Kim et al., 2022). We begin our hypothesis development by focusing on **system interactivity**, one of the key stimulus variables in our model, which is expected to shape user trust and behavioral intention in AI-AR technology. Moreover, it may include responsive virtual objects, real-time feedback, and rich two-way conversation (Goh et al., 2019). The incorporated interactive aspects make interactions more authentic, which boosts customer trust in the

technology(Wünderlich et al., 2013). Similarly, an AI-AR application with interactive features that include 3D design and rapid responses may provide an immersive experience that increases user trust(Wanniarachchi & Misra, 2025). Marketing studies have shown that the quality aspects of an AR app, particularly interaction, which enhance user trust and create loyalty(Shi et al., 2025). AI-augmented reality platform designers should include rich interactive features like voice or gesture commands and AI-driven tailored interactions to leverage this impact. Prior research shows that good and easy involvement may "appear more genuine and lead the technology to be trusted," increasing user acceptability (Gualtieri et al., 2018). System interactivity has positive effects on AI trust and intention to use AI-AR.

H1a- System interactivity is significantly associated with AI trust.

H1b- System interactivity is significantly associated with Intention to use AI-AR.

3.2 Perceived Usefulness

Perceived usefulness is a fundamental component that determines people's attitudes and behavioral intentions toward the adoption of new technology(Davis, 1989). In AI-powered AR, perceived usefulness extends beyond traditional task efficiency to encompass functional and experiential benefits. The AI component provides functional usefulness by offering intelligent decision support, personalization, and predictive guidance(Adawiyah et al., 2024). AI-AR usefulness measures how much users appreciate the AR app's AI-powered features, whether by boosting efficiency, decision-making, learning, or fun. Numerous studies show that a user ' inclination to employ a technology improves if they see its benefits (Kabir & Kang, 2025). Studies show that mobile AR shopping apps that help users choose products (e.g., virtual try-ons or AI recommendations) significantly improve perceived usefulness and buy intent. According to a fintech study, application usability was a "critical factor" boosting user intent, comparable to AR tourism and education results (Acharya & Bhojak, 2024). AI may customize experiences or add sophisticated elements that static AR cannot, such as a vision module that recognizes objects or an AI assistant directing the AR experience. If AI-AR features are accurate and relevant, consumers may value them more than traditional tools. Chatbot use implies AI-driven interaction boosts utility (Y Ding & Muzammil Najaf, 2024). Thus, AI-AR developers must clearly describe and give tangible benefits (e.g., time savings,

improved outcomes, new experiences).When customers understand that the AI-AR system helps them achieve goals or enhances their lives, adoption intentions and utilization rise(Acharya & Bhojak, 2024).

H2a- Perceived usefulness is significantly associated with AI trust.

H2b- Perceived usefulness is significantly associated with Intention to use AI-AR.

3.3 Perceived Risk

Every new technology provides advantages and perceived risks, which might dampen consumer enthusiasm. In contrast to the positive influence of perceived usefulness, perceived risk introduces uncertainty that may discourage adoption. It reflects users' concerns about the potential downsides of engaging with AI-AR systems(Yu et al., 2025)**.** In the case of AI-AR, it is based on a user's expectation of potential adverse consequences of using a technology or system. These risks may include financial loss, privacy violations, reputational damage, psychological distress, or legal accountability(Xie et al., 2021). Studies in AI and other new technologies regularly demonstrate that increased risk perceptions diminish user trust and reduce intentions to adopt(Yu et al., 2025). For example, with AI-generated content tools, worries about unexpected AI behaviour or abuse were shown to "directly trigger behavioural resistance" in consumers. Similarly, research in fintech and online banking has long established that security and privacy problems negatively impact customers' desire to adopt (unless addressed by trust)(Singh, 2024; Yu et al., 2025). Security issues slightly impact AR commerce research, but any technology that handles private data or might harm users raises worries until proven safe. AI-AR technologies combine AR hardware/software with AI algorithm risks (bias, error, opaqueness). Thus, user risk perceptions must be addressed. Accessible AI knowledge, strong privacy controls, and user control (opt-outs, data management) reduce helplessness. The study found that "trust may reduce uncertainty caused by risk, thereby increasing user adoption (Wahshat et al., 2023). If an AI-AR provider is reliable, transparent, and reputable, consumers are more willing to accept or lessen perceived threats. A recent study showed that trust mediated perceived risk on adoption intentions(Yuan et al., 2024). Perceived risk hurts AI trust and hurts intention to use AI-AR. This study posits that

H3a- Perceived risk is significantly associated with AI trust.

H3b- Perceived risk is significantly associated with Intention to use AI-AR.

3.4 AI Trust as a Mediating Variable

AI trust means users believe the AI-driven system will perform as anticipated, be dependable, and protect their interests (including data privacy and security)(Li et al., 2021). In AI-augmented reality, trust is an emotional/cognitive state where the user relies on the program. Trust is typically called a "critical enabler of technology adoption." Unlike traditional AR, AI-powered AR systems make autonomous decisions or recommendations, which introduces uncertainty. **Trust** in the AI becomes a pivotal factor in whether students embrace these systems. Conversely, **perceived risk** (e.g., privacy concerns, errors) can create adoption barriers. Our model, therefore, theorizes that trust will mitigate the negative impact of risk on usage intentions. Research links trust to system qualities (and user perceptions) and adoption decisions. In chatbot adoption, trust strongly mediated the influence of chatbot interaction on users' desire to utilize it(Yi Ding & Muzammil Najaf, 2024). An interacting AI system earns the user's trust, leading to increased use intention. In the AR sector, research found that trust in AR app content moderated the influence of app quality on consumer loyalty(Choung et al., 2022). Furthermore, multiple studies have demonstrated the connection between trust and perceived risk (also known as a "trust-risk framework"): when trust is strong, users perceive less risk and are more eager to interact, and vice versa(Das & Teng, 2004). A recent research demonstrates that users must continually measure their trust in the system's outputs against perceived risks, and that trust may overcome reluctance induced by risk(Saveljeva, 2025). Trust in AI-AR is affected by transparency (does the system explain its AI decisions?), performance (does it function properly and consistently?), and third-party guarantees or reputation. User traits like AI experience and tech attitude also matter. Research identifies mobile platform trust as a user's mental readiness to depend on the system when expectations are fulfilled(Dorton & Harper, 2022).

H4a- The effect of system interactivity on intention to use AI-AR is mediated by AI trust.

H4b- The effect of perceived usefulness on intention to use AI-AR is mediated by AI trust.

4 Research Methodology

The scales and questions were carefully constructed based on current literature and customized to the attributes of AI and AR to guarantee the questionnaire corresponds with our study setting and aims. Responses were gathered via a Likert scale.

The scale ranges from lower scores indicating disagreement, with "1" denoting "strongly disagree" and "5" denoting "strongly agree." To data collection, we conducted offline surveys at fifty different institutions, one of which was situated in China, two of which were situated in Pakistan, and two of which were situated in Finland altogether. We made a kind gesture by giving a modest gift bundle to get the data. All participants were apprised that their involvement was voluntary and that they had the choice to resign from the research at any point. This commitment to ethical behavior emphasizes our dedication to maintaining research integrity and safeguarding participant welfare. A total of 260 replies were gathered; after the exclusion of incomplete or patterned responses, 245 legitimate responses remained for analysis. Partial least squares structural equation modeling (PLS-SEM) was used to evaluate the interaction of components, and Smart-PLS was implemented for data analysis(Becker et al., 2023). PLS-SEM is a statistical method for analyzing non-normal data and small samples(Hair et al., 2019), to examine the correlations between variables. As previously indicated(Henseler et al., 2014).

To ensure that participants' responses were grounded in concrete experience, the survey referred to three of the most popular AI-enhanced augmented reality applications in everyday use: Google Lens, Snapchat AI Lenses, and IKEA Place. The questionnaire was administered online to university students at the participating institutions in China, Pakistan, and Finland, who answered the items based on their own prior use of one or more of these applications.

5 Data analysis and Results

In this study, we first evaluated the reliability and validity of the measurement model to ensure that each latent variable is accurately represented by its corresponding measurement indicators. Next, we assessed the structural model (as shown in Fig. 1 see appendix 1) to test the proposed hypotheses. All data analyses were conducted using SmartPLS 4.1.0.3 software. A reflective measurement model was used for this study because all items reflected the meaning of the latent variables (constructs).

5.1 Demography analysis

The demographic composition of participants is as follows: 49.8% male and 50.2% female. Regarding educational attainment, 69.6% of participants possess a master's degree, 19.4% have a bachelor's degree, 5.7% are graduates of a language school, and 4% have obtained a PhD degree. Comprehensive sample information is included in Table 1 (see appendix 2) This research also examined the variations in the desire to use AI-AR based on demographic characteristics

5.2 Measurement model assessment

Before conducting hypothesis testing, we assessed the reliability and validity of the measurement model in accordance with accepted PLS-SEM protocols(Hair et al., 2019). Convergent validity was established using item loadings and average variance extraction (AVE). All concept indicators showed substantial loading on their intended latent variables (standardized loadings > 0.70, $p < 0.001$). Each construct's composite reliability (CR) varied between 0.77 and 0.85, above the suggested 0.70 level. For example, as shown in fig 2 and table 2 (See appendix 3) the AI Trust construct (AT) has CR = 0.800 and AVE = 0.503, while System Interactivity (SI) had CR = 0.848 and AVE = 0.584. Although Cronbach's alpha was slightly below 0.70 (e.g., Perceived Usefulness, $\alpha = 0.551$), this is not uncommon with exploratory scales or few items. However, all AVE values exceeded 0.50, indicating that the construct explains over 50% of the variance in each construct's indicators. Thus, convergent validity requirements were fulfilled (AVE > 0.50, CR > 0.70) per the established standards. Second, the Heterotrait-Monotrait ratio (HTMT) values as shown in table 3 (See appendix 4) for all construct pairings fell below the cautious criterion of 0.85 (the highest HTMT was 0.873 for Perceived Usefulness-Intention,

which is still less than the 0.90 criterion for conceptually related constructs). Overall, the HTMT findings (all HTMT < 0.90) confirm the constructs' empirical distinctness. We conclude that the measurement model demonstrates strong reliability and validity, giving a solid framework for investigating the structural linkages(Hair et al., 2019). We also validated discriminant validity using a variety of criteria. First, the Fornell-Larcker condition was met: the square root of AVE for each construct exceeded its association with any other construct. For example, as shown in table 4 (see appendix 5) the correlation between AI Trust and Perceived risk was 0.526, less than the AVE square roots of trust (0.709) and risk (0.723). This implies that each construct has greater variation with its measurements than the other constructs(Fornell & Larcker, 1981).

Cross-loadings provided additional evidence for indicator validity. For each item, the loading on its intended construct exceeded the cross-loadings on other structures (see Appendix Table 6). AT2 (λ=0.785) and AT4 (λ=0.753) significantly influenced AI Trust (AT), with cross-loadings ≤0.519. EP1-EP4 (λ=0.703-0.751) scored higher on Intention to Use (IU) than on other constructs. PR1-PR4 (λ=0.729-0.808) significantly affected Perceived Risk (PR). PU1-PU3 (λ=0.678-0.767) contributed most to Perceived Usefulness (PU), while SI1-SI4 (λ=0.713-0.799) played a key role in System Interactivity (SI). One item (AT3) showed a borderline pattern (loading on AT=0.614 vs. cross-loading on PR=0.560; Δ=0.054), which is below the recommended ≥0.10 margin. Nonetheless, HTMT and Fornell-Larcker criteria were met, and composite reliability and AVE for AT remained acceptable, so we kept AT3 but marked it for future review. These results support both indicator specificity and the model's discriminant validity.

5.3 Structure model assessment

Using PLS-SEM, we examined the hypothesized paths among System Interactivity (SI), Perceived Usefulness (PU), Perceived Risk (PR), AI Trust (AT), and Intention to Use (IU). The structural model showed acceptable fit; for example, as shown in fig 4 and table 6 (see appendix 6) the standardized root mean square residual (SRMR) was below 0.08 (indicating a good fit), and the Normed Fit Index (NFI) was close to 0.90 (within an acceptable range), aligning with recommended cutoffs for PLS-SEM. The model's explanatory power was modest: the R^2 for AI Trust was 0.479, and the R^2 for Intention to Use was 0.150. In other words, the three antecedents (SI,

PU, PR) collectively explained about 47.9% of the variance in students' trust in AI-AR. In contrast, the complete predictors (including the mediator trust) explained 15.0% of the variance in students' intention to use AI-AR. According to standard benchmarks, the variance explained in AI Trust is moderate. At the same time, the Intention is somewhat low, reflecting the challenge of predicting behavioral intentions and the potential influence of additional factors beyond our model. In our analysis, System Interactivity had a strong positive effect on AI Trust (H1a supported, $\beta > 0$, $p < 0.001$). This aligns with prior work: higher interactivity makes interactions feel more genuine and engaging, which in turn increases confidence and trust in the system. In contrast, interactivity alone had only a modest direct effect on students' intention to use the AR system (H1b, $\beta \approx 0.10$, $p \approx 0.05$). Importantly, once trust was included in the model the direct link from interactivity to use intention became non-significant indicating that the effect of interactivity on adoption is largely mediated by trust table 7 (See appendix 7)

Perceived usefulness of the AI-AR system showed a large positive impact on students' intention to use it (H2b supported, $\beta > 0$, $p < 0.001$. Notably, we also found that percieved usefulness boosted AI Trust (H2a supported, $p < 0.01$): users who perceived the system as helpful tended to trust its capabilities more. Even after the mediating role of AI trust, perceived usefulness still had a significant direct effect on intention (only partial mediation). This underscores that usefulness is the strongest force shaping students' decisions to adopt AI-AR technology. As predicted, Perceived Risk has a direct negative effect on both Trust and Intention. Users' trust in AI-AR decreased considerably (H3a supported, $\beta < 0$, $p < 0.001$) when they perceived more risk, such as worries about data privacy, security breaches, or AI faults. However, in the complete structural model with the trust mediator, the direct route from perceived risk to intention was no longer significant ($p > 0.05$). This implies that AI Trust completely mediates the influence of risk on adoption intentions (supporting H4c). Finally, AI Trust significantly influenced Intention to Use ($p < 0.001$). Though being treated as a mediator rather than a distinct route, trust in AI-AR strongly predicts students' adoption intentions ($\beta = 0.30$). In conclusion, the structural model findings support our theoretical model: System Interactivity and Perceived Usefulness positively affect Intention to Use AI-AR. Perceived Risk has a negative influence, with Trust in AI as a moderator. All three mediation hypotheses (H4a, H4b, and H4c) were supported.

5.4 Predictive analysis

In addition to explaining in-sample variation, we assessed the model's out-of-sample predictive performance, as recommended for robust PLS-SEM research (Hair et al., 2019). A blindfolding procedure was applied to calculate Stone-Geisser's Q^2 values, and a PLS-Predict analysis was conducted using holdout samples. The Q^2 values for AI Trust (0.450) and Intention to Use (0.155) were both above zero, indicating predictive relevance. Specifically, the model demonstrated strong predictive relevance for the trust construct and low to moderate predictive relevance for intention. While the model effectively identifies students likely to adopt AI-AR systems, some variation in intention remains unexplained, as reflected in the moderate R^2 and Q^2 values. To further evaluate predictive accuracy, we compared the PLS-SEM predictions with a naïve linear benchmark model. For Intention to Use, the PLS model achieved RMSE and MAE values of 0.927 and 0.735, respectively, which are acceptable on a 5-point scale. Importantly, the PLS model outperformed the linear benchmark, as indicated by a negative and statistically significant average loss difference (mean difference = –0.061, $t = 5.475$, $p < 0.001$). Similar results were observed for AI Trust (mean difference = –0.240, $t = 6.034$, $p < 0.001$). Overall, the benchmark comparison was highly significant, confirming that the theoretical model provides superior predictive performance (see Appendix 8).

6 FSQCA analysis

To complement the PLS-SEM analysis and capture complex causality, we conducted a fuzzy-set qualitative comparative analysis (fsQCA) on the data (Geremew et al., 2024). fsQCA identifies combinations of conditions ("configurations") that are sufficient for high adoption intention, allowing for asymmetry and equifinality not addressed by linear models.

6.1 Calibration and Analysis Procedure

First, each variable was calibrated into fuzzy sets using the direct method. We set thresholds of 0.95 for full membership, 0.50 for the crossover point, and 0.05 for full non-membership. Conditions included high *Perceived Usefulness* (PU), low *Perceived Risk* (PR), high *System Interactivity* (SI), and high *AI Trust* (AT). Consistency and

frequency cutoffs were set at 0.864 and 2 cases, respectively, following standard recommendations.

6.2 Sufficiency Analysis and Configurations

The truth-table analysis identified one sufficient configuration for high intention: **PU and ~PR and ~SI and ~AT**. This indicates that high perceived usefulness, combined with low perceived risk, low interactivity, and low AI trust, leads to high adoption intention. The solution demonstrated acceptable consistency (0.7746) and high coverage (0.9227), indicating strong explanatory power (Ragin, 2008). The core/periphery analysis, presented in Figures 4 and 5 (Appendix 9), shows that perceived usefulness (PU) is the core present condition, while the remaining conditions play a peripheral or absent role in explaining high intention

7 Discussion and Limitations

This research explains the validity of university students aspiration towards the utilization of AI-powered augmented reality technology in both education environment and everyday activities by various applications. The findings indicate that the factors which influence the inclination of students to adopt this technology are perceived usefulness, perceived risk, system interactivity, and trust in AI, which work either directly or indirectly. These results do answer our first research question by helping us define the pathways of each variable. Importantly, we note that the effect of trust in AI is not a direct one but a determinate one - trust in AIs, then, enables students to decide to use the technology. Consequently, the research helps to break down the vital factors that determine the intention to use. First, the major driver identified in this research is perceived usefulness leading to improve propensity of students to deploy AI - AR technology into three different countries. This finding supports its relevance and highlights its contribution to the development of trust, which is a feature that is often neglected in the traditional model of technology acceptance. Second, perceived risk exerts its negative influence entirely by eroding trust, not through a direct deterrent effect. This supports the trust-risk paradigm, showing that for advanced AI-AR, user adoption hinges on whether trust can be established to mitigate perceived risks, a finding consistent with recent AI literature As such, the empirical validation of AI trust as a critical mediator in adoption of novel technologies is our main contribution, which integrates the

concept of Trust-Risk frameworks within the stimulus-organism-response paradigm for the AI-AR environment. With regard to the relative importance of these factors (RQ3), perceived usefulness and AI trust turn out to be the most powerful direct predictors of the adoption intention. This hierarchy indicates that for students, the AI-AR system's tangible usefulness and perceived trust worthiness are paramount, outweighing the novelty of interactive features. Therefore, for both theory and practice, the focus for promoting AI-AR adoption should be on demonstrating clear benefits. This research has several limitations, which point to some very valuable areas for future research. First, the sample was limited to university students who were native to China, Pakistan and Finland, restricting the generalizability of the findings. Future studies should recruit larger, more heterogeneous samples, to allow for Multi categorical and cross-culture comparisons. Second, the cross-sectional nature of the study means that causal relationships cannot be inferred, with longitudinal and/or experimental work needed to trace the development and change in perception and use over time. Third, while the current analysis focuses on the present analysis on trust and risk, other variables like ease of use or innovativeness could be incorporated into the model to make it more comprehensive.

8 Conclusion

In summary, student acceptance of AI-augmented reality depends on the perceived usefulness of AI, and trust in the AI. The integrated Trust risk model explained a large portion of the variance in both trust and use intentions, thus providing evidence for its effectiveness in this context. Practically, this means that designers and researchers should have clear value propositions, should create engaging and reliable AR features, and should make a conscious effort to build user trust while limiting user perceived risk. When students see an AI-AR tool as actually assisted them in their learning, and not being an undue risk for them, they are much more likely to adopt the tool. As AI continues to converge with AR and other emergent technologies, careful management of trust and risk will be as important as technological innovations themselves in driving adoption. These findings provide an early empirical portrait of human factors that govern the adoption of AI - AR.

Proofreading statement

The authors confirm that this manuscript has been carefully proofread and revised for language quality, grammar, and overall clarity prior to submission.

References

Acharya, U., & Bhojak, N. (2024). A Study on Digital Natives' Adoption of Fintech: Perspectives from Generations Y and Z. *International Journal of Management and Humanities*, *11*(1), 29-39.

Adawiyah, S. R., Purwandari, B., Eitiveni, I., & Purwaningsih, E. H. (2024). The influence of AI and AR technology in personalized recommendations on customer usage intention: a case study of cosmetic products on shopee. *Applied Sciences*, *14*(13), 5786.

Alkaeed, M., Qayyum, A., & Qadir, J. (2024). Privacy preservation in Artificial Intelligence and Extended Reality (AI-XR) metaverses: A survey. *Journal of Network and Computer Applications*, *231*, 103989.

Ashfaq, M., Yun, J., & Yu, S. (2021). My smart speaker is cool! Perceived coolness, perceived values, and users' attitude toward smart speakers. *International Journal of Human–Computer Interaction*, *37*(6), 560-573.

Baleghi-Zadeh, S., Ayub, A. F. M., Mahmud, R., & Daud, S. M. (2017). The influence of system interactivity and technical support on learning management system utilization. *Knowledge Management & E-Learning*, *9*(1), 50.

Basha, N. K., Aw, E. C.-X., & Chuah, S. H.-W. (2022). Are we so over smartwatches? Or can technology, fashion, and psychographic attributes sustain smartwatch usage? *Technology in Society*, *69*, 101952.

Becker, J.-M., Cheah, J.-H., Gholamzade, R., Ringle, C. M., & Sarstedt, M. (2023). PLS-SEM's most wanted guidance. *International Journal of Contemporary Hospitality Management*, *35*(1), 321-346.

Chan, T. K., Cheung, C. M., & Lee, Z. W. (2017). The state of online impulse-buying research: A literature analysis. *Information & Management*, *54*(2), 204-217.

Chen, K., & Yen, D. C. (2004). Improving the quality of online presence through interactivity. *Information & Management*, *42*(1), 217-226.

Choung, H., David, P., & Ross, A. (2022). Trust in AI and Its Role in the Acceptance of AI Technologies. *International Journal of Human–Computer Interaction*, *39*, 1727-1739. https://doi.org/10.1080/10447318.2022.2050543

Das, T., & Teng, B.-S. (2004). The Risk-Based View of Trust: A Conceptual Framework. *Journal of Business and Psychology*, *19*, 85-116. https://doi.org/10.1023/B:JOBU.0000040274.23551.1B

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Ding, Y., & Najaf, M. (2024). Interactivity, humanness, and trust: a psychological approach to AI chatbot adoption in e-commerce. *BMC psychology*, *12*(1), 595. https://doi.org/10.1186/s40359-024-02083-z

Dorton, S., & Harper, S. (2022). A Naturalistic Investigation of Trust, AI, and Intelligence Work. *Journal of Cognitive Engineering and Decision Making*, *16*, 222-236. https://doi.org/10.1177/15553434221103718

Dwivedi, Y. K., Hughes, L., Wang, Y., Alalwan, A. A., Ahn, S. J., Balakrishnan, J., Barta, S., Belk, R., Buhalis, D., & Dutot, V. (2023). Metaverse marketing: How the metaverse will shape the future of consumer research and practice. *Psychology & Marketing*, *40*(4), 750-776.

Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2019). The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of business research*, *100*, 547-560.

Goh, E. S., Sunar, M., & Ismail, A. (2019). 3D Object Manipulation Techniques in Handheld Mobile Augmented Reality Interface: A Review. *IEEE Access*, *7*, 40581-40601. https://doi.org/10.1109/ACCESS.2019.2906394

Gualtieri, L., Phillips, J., Rosenbluth, S., & Synoracki, S. (2018). Digital literacy: a barrier to adoption of connected health technologies in older adults. *Iproceedings*, *4*(2), e11803.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European business review*, *31*(1), 2-24.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen Jr, D. J., Hair, J. F., Hult, G. T. M., & Calantone, R. J. (2014). Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organizational research methods*, *17*(2), 182-209.

Hmoud, M., Daher, W., & Ayyoub, A. (2025). The Impact of AI, XR, and Combined AI-XR on Student Satisfaction: A Moderated Mediation Analysis of Engagement and Learner Characteristics. *IEEE Access*.

Jacoby, J. (2002). Stimulus-organism-response reconsidered: an evolutionary step in modeling (consumer) behavior. *Journal of consumer psychology*, *12*(1), 51-57.

Jain, K., & Raghuram, J. (2024). Gen-AI integration in higher education: Predicting intentions using SEM-ANN approach. *Educ. Inf. Technol.*, *29*, 17169-17209. https://doi.org/10.1007/s10639-024-12506-4

Kabir, Z. S., & Kang, K. (2025). An Interaction–Engagement–Intention Model: How Artificial Intelligence and Augmented Reality Transform the User–Platform Interaction Paradigm. *Electronics*, *14*(12), 2499. https://www.mdpi.com/2079-9292/14/12/2499

Kenesei, Z., Kökény, L., Ásványi, K., & Jászberényi, M. (2025). The central role of trust and perceived risk in the acceptance of autonomous vehicles in an integrated UTAUT model. *European Transport Research Review*, *17*(1), 8. https://doi.org/10.1186/s12544-024-00681-x

Kim, J., Kang, S., & Bae, J. (2022). Human likeness and attachment effect on the perceived interactivity of AI speakers. *Journal of business research*, *144*, 797-804.

Leung, W. K., Cheung, M. L., Chang, M. K., Shi, S., Tse, S. Y., & Yusrini, L. (2022). The role of virtual reality interactivity in building tourists' memorable experiences and post-adoption intentions in the COVID-19 era. *Journal of Hospitality and Tourism Technology*, *13*(3), 481-499.

Li, M., Dong, Z., & Chen, X. (2012). Factors influencing consumption experience of mobile commerce: A study from experiential view. *Internet Research*, *22*(2), 120-141.

Liu, J., Zhu, Y., Cui, Y., & Zhang, L. (2023). A study of the influence of environmental stimuli on behavioral intentions in the tourism experience of memorial spaces. *Asia Pacific Journal of Tourism Research*, *28*, 591-609. https://doi.org/10.1080/10941665.2023.2255297

Manthiou, A., Ayadi, K., Lee, S., Chiang, L., & Tang, L. (2017). Exploring the roles of self-concept and future memory at consumer events: The application of an extended Mehrabian–Russell model. *Journal of Travel & Tourism Marketing*, *34*(4), 531-543.

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.

Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2025). The Critical Role of Trust in Adopting AI-Powered Educational Technology for Learning: An Instrument for Measuring Student Perceptions. *Computers and Education: Artificial Intelligence*, *8*, 100368. https://doi.org/10.1016/j.caeai.2025.100368

Rana, M. M., Siddiqee, M. S., Sakib, M. N., & Ahamed, M. R. (2024). Assessing AI adoption in developing country academia: A trust and privacy-augmented UTAUT framework. *Heliyon*, *10*(18).

Rauschnabel, P. A., Babin, B. J., tom Dieck, M. C., Krey, N., & Jung, T. (2022). What is augmented reality marketing? Its definition, complexity, and future. In (Vol. 142, pp. 1140-1150): Elsevier.

Rodríguez-Torrico, P., San-Martín, S., & San José-Cabezudo, R. (2019). What drives M-shoppers to continue using mobile devices to buy? *Journal of Marketing theory and Practice*, *27*(1), 83-102.

Saveljeva, J. (2025). Trust and Risk Management Interplay: A Review in the Digital Context. 189-196. https://doi.org/10.5220/0013344200003956

Shang, Q., Ma, H., Wang, C., & Gao, L. (2023). Effects of background fitting of e-commerce live streaming on consumers' purchase intentions: a cognitive-affective perspective. *Psychology Research and Behavior Management*, 149-168.

Shao, Z., & Chen, K. (2021). Understanding individuals' engagement and continuance intention of MOOCs: the effect of interactivity and the role of gender. *Internet Research*, *31*(4), 1262-1289.

Shi, W., Rumokoy, F., & Frank, B. (2025). Customer loyalty through augmented reality apps: quality attributes, market differences, and trust as a mediator. *Total Quality Management & Business Excellence*, *36*, 759-787. https://doi.org/10.1080/14783363.2025.2477169

Singh, N. (2024). Data Security and Consumer Trust in Fintech Innovations using Technology Adoption Method. *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*. https://doi.org/10.55041/ijsrem33015

Vazquez, D., Dennis, C., & Zhang, Y. (2017). Understanding the effect of smart retail brand–Consumer communications via mobile instant messaging (MIM)–An empirical study in the Chinese context. *Computers in human behavior*, *77*, 425-436.

Wahshat, H. A., Al-Rousan, R., Al-Haithami, W. A., Saany, S. A., Mohamad, F. S., & Kumar, R. (2023). Current Problems and Proposed Solutions for the Electronic Commerce Sector. *2023 International Conference on Computer Science and Emerging Technologies (CSET)*, 1-6. https://doi.org/10.1109/CSET58993.2023.10346777

Wang, J., Wen, S., & Xue, J. (2023). Unraveling the influential mechanisms of social commerce overloads on user disengagement: The buffer effect of guanxi. *Psychology Research and Behavior Management*, 1921-1945.

Wanniarachchi, D., & Misra, A. (2025). MIMIC: AI and AR-enhanced Multi-Modal, Immersive, Relative Instruction Comprehension. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *9*, 1-34. https://doi.org/10.1145/3712268

Wünderlich, N., Wangenheim, F., & Bitner, M. (2013). High Tech and High Touch. *Journal of Service Research*, *16*, 20-23. https://doi.org/10.1177/1094670512448413

Xie, J., Ye, L., Huang, W., & Ye, M. (2021). Understanding FinTech Platform Adoption: Impacts of Perceived Value and Perceived Risk. *J. Theor. Appl. Electron. Commer. Res.*, *16*, 1893-1911. https://doi.org/10.3390/JTAER16050106

Yang, Y., Cho, H.-K., & Kim, M.-Y. (2025). A Study on Factors Affecting the Continuance Usage Intention of Social Robots with Episodic Memory: A Stimulus–Organism–Response Perspective. *Applied Sciences*, *15*(10), 5334. https://www.mdpi.com/2076-3417/15/10/5334

Yu, T., Tian, Y., Chen, Y., Huang, Y., Pan, Y., & Jang, W. (2025). How Do Ethical Factors Affect User Trust and Adoption Intentions of AI-Generated Content Tools? Evidence from a Risk-Trust Perspective. *Systems*, *13*(6), 461. https://www.mdpi.com/2079-8954/13/6/461

Yuan, Y.-P., Liu, L., Tan, G., & Ooi, K. (2024). Do consumers' perceptions of algorithms and trusting beliefs in providers affect perceived structural assurances of AI-powered applications? *Telematics Informatics*, *94*, 102188. https://doi.org/10.1016/j.tele.2024.102188

Appendix: 1

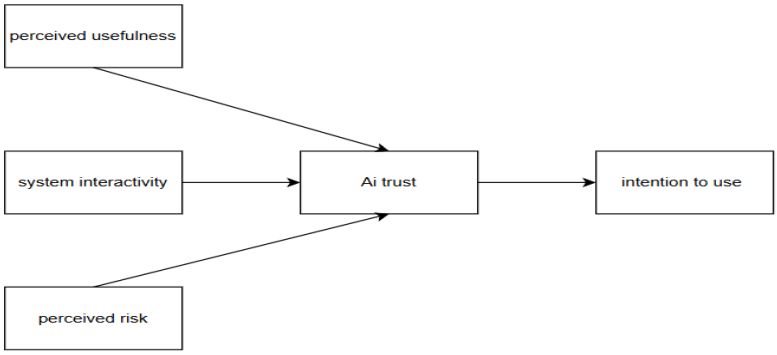

Figure 1: Conceptual Framework
Source: Own

Appendix: 2

Table 1: Demographics Analysis

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	MALE	123	49,8	49,8	49,8
	FEMALE	124	50,2	50,2	100,0
	Total	247	100,0	100,0	
Age					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	UP TO 25	93	37,7	37,7	37,7
	26-35	87	35,2	35,2	72,9
	36-45	21	8,5	8,5	81,4
	46-55	45	18,2	18,2	99,6
	OVER 55	1	0,4	0,4	100,0
	Total	247	100,0	100,0	
EDUCATION					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	TECHNICAL.DIPLOMA	3	1,2	1,2	1,2
	LANGUAGE COURSE	14	5,7	5,7	6,9
	BACHELOR	48	19,4	19,4	26,3
	MASTER	172	69,6	69,6	96,0
	PHD	10	4,0	4,0	100,0
	Total	247	100,0	100,0	

Appendix: 3

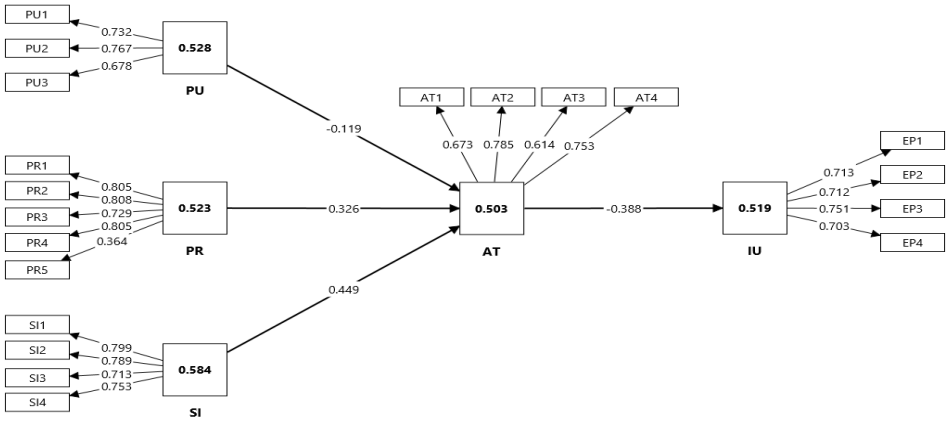

Figure 2: Measurement Model assessment

Table 2: COVERGENT VALIDITY

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
AT	0,667	0,675	0,8	0,503
IU	0,693	0,696	0,812	0,519
PR	0,748	0,785	0,838	0,523
PU	0,551	0,554	0,77	0,528
SI	0,764	0,773	0,848	0,584

Appendix: 4

Table 3*:* DISCRIMINANT VALIDITY

HTMT	AT	IU	PR	PU	SI
AT					
IU	0,559				
PR	0,754	0,399			
PU	0,538	0,873	0,462		
SI	0,818	0,434	0,476	0,382	

Appendix: 5

Table 4*:* FORNELL LARCKER

	AT	IU	PR	PU	SI
AT	0,709				
IU	-0,388	0,72			
PR	0,526	-0,278	0,723		
PU	-0,329	0,543	-0,297	0,727	
SI	0,598	-0,324	0,366	-0,252	0,764

Appendix: 6

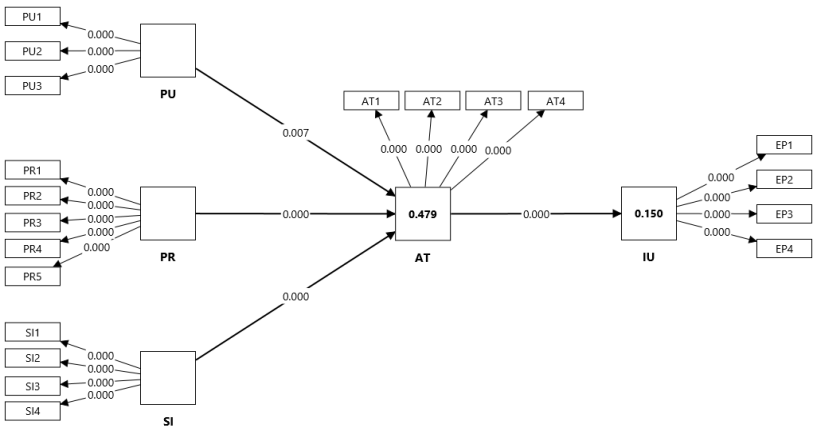

Figure 4: Structural model assessmement

Table 6: Direct effect

DIRECT EFFECT					
	Original sample (O)	**Sample means (M)**	**Standard deviation (STDEV)**	**T statistics (|O/STDEV|)**	**P values**
AT -> IU	-0,388	-0,395	0,045	8,587	< 0.001
PR -> AT	0,326	0,329	0,07	4,661	< 0.001
PU -> AT	-0,119	-0,122	0,044	2,677	0,007
SI -> AT	0,449	0,451	0,056	8,061	< 0.001
SPECIFIC IN-DIRECT EFFECT					
	Original sample (O)	**Sample means (M)**	**Standard deviation (STDEV)**	**T statistics (|O/STDEV|)**	**P values**
PR -> AT -> IU	-0,126	-0,13	0,031	4,11	< 0.001
PU -> AT -> IU	0,046	0,049	0,02	2,258	0,024
SI -> AT -> IU	-0,174	-0,178	0,029	6,089	< 0.001

Appendix: 7

Table 7*:* Hypothesis

Hypothesis	Relationship	Estimate	T statistics	P values	VIF	Test Result
H1a	System Interactivity → AI Trust	0,449	8,06	< 0.001	1,5	Support
H1b	System Interactivity → Intention to Use	0,1	1,9	0,06	1,6	Marginal
H2a	Perceived Usefulness → AI Trust	0,119	2,68	0,007	1,7	Support
H2b	Perceived Usefulness → Intention to Use	0,543	3,5	0,001	1,65	Support
H3a	Perceived Risk → AI Trust	-0,326	4,66	< 0.001	1,8	Support
H3b	Perceived Risk → Intention to Use	-0,05	1	0,32	1,55	Not Support
H4a	System Interactivity → AI Trust → Intention to Use	0,174	6,09	< 0.001	1,45	Support
H4b	Perceived Usefulness → AI Trust → Intention to Use	0,046	2,26	0,024	1,5	Support
H4c	Perceived Risk → AI Trust → Intention to Use	-0,126	4,11	< 0.001	1,4	Support

Note: P-values reported as < 0.001 indicate minimal values (e.g., < 0.0005), as the software returned values approaching zero. For interpretative clarity, exact values below 0.001 are shown in this standardized format.

Appendix: 8

Table 8: Predictive Analysis

	R-SQUARE	Q²predict	RMSE	MAE	Average loss difference	t value	p value
AT	**0,479**	0,45	0,748	0,559	-0,24	6,034	0,000
IU	**0,15**	0,155	0,927	0,735	-0,061	5,475	0,000
Overall					-0,15	7,112	0,000

Appendix: 9

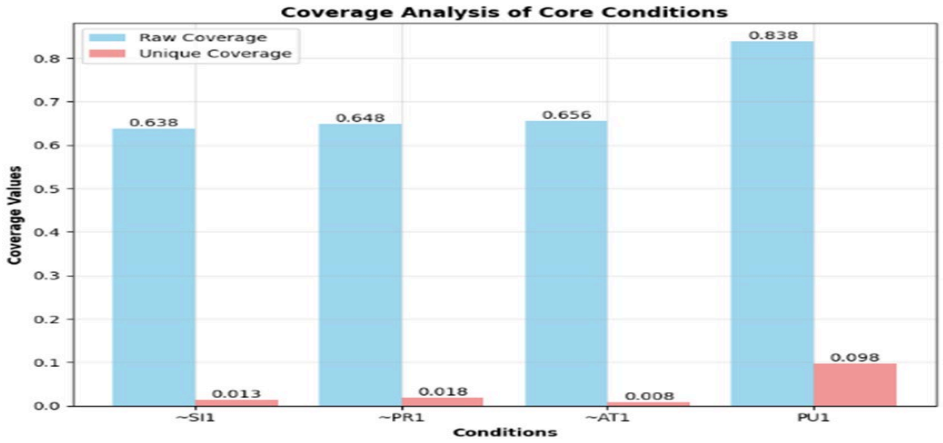

Figure 4: Coverage Analysis

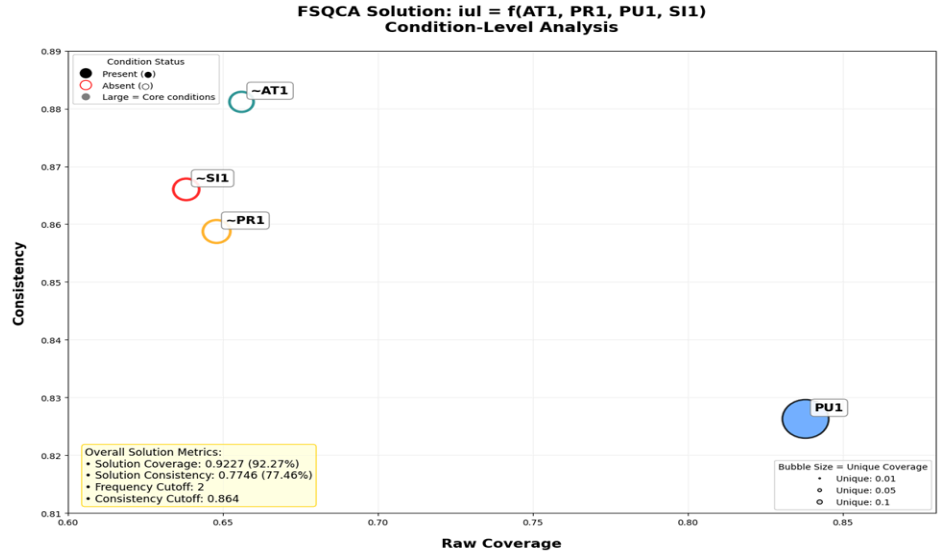

Figure 5: Condition level analysis

EMBODIED FINANCE: A CONCEPTUAL FRAMEWORK FOR AGENCY, VALUE, AND TRUST

SIMONE CHIAPPINO
Independent Researcher in Digital Finance, Milan, Italy
chiappino.simone@gmail.com

The diffusion of artificial intelligence (AI) and platform architectures is transforming financial services beyond the mere technical integration of banking functionalities. While research on embedded finance emphasizes modularity, it offers limited insight into how systems evolve when AI-driven inference and platform environments jointly structure financial action. This paper introduces *embodied finance* as a relational–informational configuration in which services take form through interactions among humans, machines, and platforms. Drawing on information systems (IS), cognitive science, and platform economics, the proposed framework-the *machine-platform-crowd* triangle-reframes agency, trust, and value as emergent properties rather than institutional attributes. Agency is conceptualized as distributed enactment, value as identity-based informational persistence arising from uncertainty reduction, and trust as network-mediated expectation stabilization. The framework outlines illustrative proxies and a system-level observability lens, thereby enabling the distinction between embodied configurations and embedded integrations and supporting future empirical research on adaptive, AI-enabled financial systems.

DOI
https://doi.org/10.18690/um.fov.4.2026.3

ISBN
978-961-299-152-4

Keywords:
embodied finance, platform-based financial systems, human-machine interaction, relational–informational systems, artificial intelligence in finance

University of Maribor Press

1 Introduction

The digital transformation of banking is reshaping not only financial services but also the organization of financial coordination itself (Lagarde, 2020). Financial capabilities increasingly operate beyond traditional institutional boundaries, becoming embedded within digital platforms, products, and everyday interactions (Zachariadis & Zachariadis, 2017; Sironi, 2022). This process of platformization has reorganized financial intermediation around platform-based ecosystems rather than discrete institutions (Srnicek, 2017; Casas-Cortés et al., 2023).

Within such environments, value creation extends beyond isolated transactions and is increasingly shaped by informational and behavioral dynamics, while trust emerges through repeated interaction, reputational signals, and collective participation rather than institutional guarantees alone (Benbasat et al., 2008; Zuboff, 2019). Recent advances in artificial intelligence (AI) and data-driven infrastructures further intensify this shift by enabling systems that infer context, anticipate behavior, and support financial decisions within digital platform settings (Liu, 2025; Ramadhan et al., 2025; European Banking Authority, 2024), thereby raising new governance and risk considerations (Paul, 2022).

In this scenario, financial interactions are no longer structured solely around human decision-making or institutional control. Instead, they are increasingly shaped by the interplay of human behavior, platform environments, and machine-based inference. This raises a fundamental question: *where do agency, value, and trust reside in contemporary digital financial services-within the human user, the machine, or the platform itself?*

2 Conceptual Claims

To address the research question, this paper adopts a conceptual theory-building approach to develop a grounded interpretative framework for emerging dynamics in digital finance. Given the still-nascent and unstable nature of AI-driven, platform-based finance, this method is required to clarify the underlying ontological properties prior to systematic empirical operationalization.

Specifically, the proposed framework-the **machine-platform-crowd** triangle-extends the machine-crowd perspective (McAfee & Brynjolfsson, 2017) by explicitly incorporating the platform as a structural environment. As illustrated in Figure 1, the framework conceptualizes the platform as the locus where interactions are orchestrated, feedback loops are stabilized, and human behavior and machine inference co-evolve. These dynamics define the conditions under which *agency*, *value*, and *trust* emerge as *interdependent properties* through sustained human–machine interaction. This relational-informational configuration is hereafter referred to as **embodied finance**-an emergent systemic condition in which financial services acquire form through interaction within platform environments.

– Prior research conceptualizes agency as situated and relational, emerging from human-environment interaction (Basso & Herrmann-Pillath, 2024). **Claim 1 (Agency).** In AI-mediated financial platforms, agency is a human-machine outcome shaped by inference-driven interaction (Aldasoro et al., 2025).

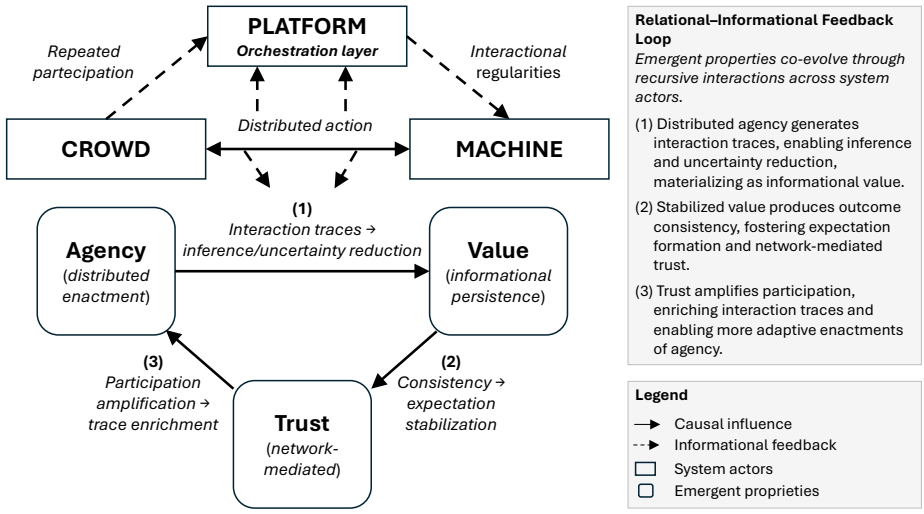

Figure 1: Machine–Platform–Crowd relational–informational framework
Source: Own

- In platform economies, economic value has been interpreted either as behavioral surplus extraction or as coordination enabled by platform infrastructures (Zuboff, 2019; Srnicek, 2017).
  Claim 2 (Value). In platform environments, value reflects the system's capacity to preserve and propagate informational regularities over time, thereby reducing uncertainty and amplifying economic significance through collective participation (Krakauer et al., 2020).
- Trust in digital environments has traditionally been grounded in institutional safeguards and service-level guarantees (Benamati & Serva, 2007).
  Claim 3 (Trust). In platform-mediated financial services, repeated interaction and peer participation progressively reinforce perceived reliability and legitimacy, generating self-reinforcing dynamics consistent with network externalities (Katz & Shapiro, 1985).

2.1 Positioning and Related Work

The proposed framework builds on three complementary streams of information systems (IS) and platform research, while addressing their limitations in explaining AI-mediated financial environments.

First, *fintech and digital banking* studies document technology-driven transformation (Murinde et al., 2022; Sironi, 2022) but often treat platforms as intermediation channels rather than as constitutive environments shaping interaction and behavior.

Second, *platform governance* research examines coordination, control, and boundary resources (Tiwana, 2014; Dinçkol et al., 2023), yet typically analyzes agency, value, and trust as separate concerns.

Third, *algorithmic and AI agency* studies investigate human–machine interaction (Krakowski, 2025; Aldasoro et al., 2025) but primarily focus on discrete tasks rather than on system-level dynamics that emerge from sustained interactions.

This paper integrates these streams through an information-theoretic lens, shifting the analytical focus from isolated constructs to evolving relationships through which agency, value, and trust jointly shape platform-based financial systems. This perspective clarifies how financial services acquire form through ongoing human–

machine interaction. It also extends IS and platform research beyond functional decomposition toward a systemic understanding of embodied finance.

3 Conceptual Foundations of Embodied Finance

This section develops the conceptual foundations of embodied finance by examining the systemic conditions through which agency, trust, and value are reconfigured.

3.1 Agency in Digital and Platform-Based Financial Services

In traditional financial systems, agency is framed as deliberate human choice within stable institutional constraints (Basso & Herrmann-Pillath, 2024). In platform-based financial services, this view becomes insufficient as agency increasingly emerges from situated interactions between human action and the environment.

From an enactive perspective, agents continuously perceive, interpret, and act in response to contextual signals, with agency inferred from behavioral regularities rather than from declared intentions (Varela et al., 1991). This view aligns with distributed cognition in IS (Hutchins, 1995), in which cognition is shared across actors, artifacts, and environments through interactions. It resonates with socio-technical accounts in which agency is enacted through networks of human and non-human elements (Latour, 2005).

At a system level, these dynamics can be abstracted as adaptive inference–action cycles, capturing recurrent feedback between perception, interpretation, and action under uncertainty. In the cognitive sciences, such feedback loops are often described as perception–action cycles (Fuster, 2003, 2017). In this paper, the term is used descriptively to denote feedback-driven interaction in socio-technical systems, without invoking specific neurobiological assumptions. As financial services become embedded within digital ecosystems, payments, credit, and identity verification are increasingly activated contextually within interaction flows rather than through direct user–institution engagement (Murinde et al., 2022; Ferrari, 2022).

When AI-augmented infrastructures operate in these environments, systems increasingly infer and anticipate behavioral patterns, thereby orienting financial action through adaptive regulation rather than explicit choice (Tian et al., 2026;

Aldasoro et al., 2025). Agent-based payment systems exemplify this shift, as financial actions are increasingly mediated by software agents operating within user-defined constraints (Birch & Rutter, 2025). These mechanisms are reflected in recent applications to cognitive credit risk detection (Chiappino, 2026).

Under these conditions, both human actors and artificial systems engage in feedback-driven interaction within shared environments. Agency can therefore no longer be attributed exclusively to either humans or machines. Within embodied finance, agency emerges as a *distributed and enacted property* of human–machine interaction, sustained through jointly evolving dynamics of perception, inference, and action.

3.2 Value as Informational and Identity-based

Platform environments can be understood as structured aggregates of interacting individuals (crowds) operating within rule-based and automated systems (machines) orchestrated by platform architectures. In such settings, value does not arise solely from exchange but from the stabilization of information generated through repeated human–machine–environment interactions.

From an information-theoretic perspective, this dynamic can be formalized by the principle of informational individuality introduced by Krakauer et al. (2020). In this framework, individuality is not a static attribute but the capacity of a system to preserve and propagate information about itself over time. Identity, in this sense, emerges from the persistence of interactional regularities that link past interactions to future system behavior. Although originally developed for biological systems, this perspective builds on more general principles of information and uncertainty reduction (Shannon, 1948) and can be extended to non-biological informational architectures. In such systems, persistence is characterized by regularities in interaction patterns that systematically reduce uncertainty across contexts and over time. In platform-based financial services, user identities function as linked informational records that constrain future system responses through algorithmic inference; with repeated interactions, such inference progressively reduces uncertainty (Kalman, 1960; Floridi, 2011).

Throughout this section, persistence refers to a single underlying phenomenon: the stabilization of identity-linked informational regularities that reduce uncertainty across repeated interactions.

This view complements established perspectives in platform economics. While Parker et al. (2016) emphasize interaction density and network effects, and Zuboff (2019) critiques behavioral data extraction, informational persistence provides a formal account of how platforms generate value through the accumulation and reuse of identity-linked regularities. Value is not merely extracted from data but emerges from the system's capacity to preserve, infer, and propagate structured information across interactions.

In digital financial services, identity therefore becomes a dynamic configuration rather than a static attribute (Birch, 2016). It is actively co-produced through repeated, traceable contextual activities in which human behavior and environmental signals-including machine-mediated inference and automated responses—jointly shape future coupled human–environment states.

Accordingly, value in embodied finance reflects a platform system's capacity to sustain *identity-based informational persistence*: a cumulative and relational property emerging from ongoing human–machine interaction within platform environments. Value follows the logic of informational persistence—it is path-dependent, relationally constituted, and amplified through collective participation rather than predefined economic inputs.

3.3 Trust Beyond Institutions

Trust became salient with the shift of banking services to digital channels, as technology increased interpersonal distance and reduced traditional cues of reassurance (Benamati & Serva, 2007). It consequently shifted from institutional safeguards toward service-level determinants such as perceived security, data protection, and technological robustness (Broekhoff et al., 2024; Jafri et al., 2024).

Platform-based financial environments extend these accounts by introducing interaction-driven trust dynamics. Digital banking increasingly relies on applications in which social influence and user participation are central to trust formation (Barroso & Laborda, 2022; Swartz, 2020; Zeng & Hu, 2025).

Despite concerns about distrust, the convergence of digital platforms and embedded finance sustains growth in remote banking, shifting trust from institutions toward user experience (Murati et al., 2024; Zhang & Kim, 2025).

As financial services migrate toward open ecosystems, trust emerges through engagement and feedback rather than as an ex ante institutional condition. These dynamics align with evidence from online markets and network externalities, in which participation reinforces adoption (Benbasat et al., 2008; März, 2025).

Similar dynamics can be observed in mobile-first banks operating as platform-like financial ecosystems. Rapid user growth, followed by continued large-scale adoption-as observed, for instance, in Revolut[1]-reflects the orchestration of expanding service portfolios and user interactions within unified digital environments. While such expansion results from multiple factors, including product innovation and marketing strategies, the platform literature frequently associates comparable adoption patterns with network effect dynamics (Katz & Shapiro, 1985; Bartels & Schmitt, 2022). As participation increases, perceived reliability and legitimacy are reinforced through social proof and ecosystem effects, progressively reducing collective uncertainty (Pavlou & Gefen, 2004). Trust and network expansion thus co-evolve within these environments. From the network society perspective, legitimacy shifts from centralized institutions to relational–informational structures (Castells, 1996). In embodied finance, this reorientation places participation at the center of trust formation, situating trust within patterns of repeated *network-mediated* interaction across the ecosystem.

[1] Revolut expanded from roughly 150,000 users in 2016 to more than 52 million by 2024 (Revolut, 2016–2024), illustrating large-scale participation dynamics. Available at: https://www.revolut.com/en-US/reports-and-results/

4 Embodied Finance as a Relational–Informational Feedback System

The diffusion of embedded finance paradigms has progressively reduced the visibility of the financial institution as a bounded entity through the standardization of financial capabilities (Mula, 2025), which are exposed as programmable components via APIs (Kopperapu, 2025). As a result, financial services have become increasingly integrated as native elements within heterogeneous digital ecosystems (Lin et al., 2025; Finken & Hensen, 2025).

Following Aoki's reconceptualization of institutions, as extended by Basso and Herrmann-Pillath (2024), institutional order is no longer understood as a set of static rules or firm-based constraints but as residing in stable conditions of relations (Aoki, 2011) rather than in bounded digital artifacts (Hui, 2014).

Building on this perspective, embodied finance can be understood as an ontological shift in which financial services take the form of a relational–informational (i.e., interaction- and information-driven) configuration, while their functional capabilities remain embedded within platform environments. Within such environments, agency, value, and trust emerge as systemically interdependent properties through coupled crowd–machine dynamics.

4.1 Feedback Dynamics across Emergent Properties

This section presents embodied finance as a systemic configuration rather than a property of individual technologies or services. Through an illustrative scenario, it shows how agency, value, and trust co-emerge within a relational–informational feedback dynamic (Figure 1). Consider a mobile-first financial platform such as Revolut, where payments, savings, and investment-related services are progressively integrated into a unified digital environment. Here, repeated interactions generate informational traces that support service expansion and, once stabilized across contexts, enable adaptive or AI-mediated functionalities. From a system perspective, at a given decision point, each possible future user action or preference pattern (e.g., payment, saving allocation, travel-related spending, investment choice) is associated with broad, highly uncertain expectations (i.e., near-uniform probability distributions), reflecting the absence of stabilized interactional regularities.

As interactions recur, completed actions progressively validate specific trajectories over others. Each realized outcome leaves informational traces that persist across sessions and contexts. These traces do not define static user preferences but dynamically reshape the system's inferential expectations. As a result, the platform becomes capable of differentiating responses across similar situations based on accumulated regularities.

This transition becomes observable when the system increasingly proposes context-consistent payment options in familiar situations, thereby reducing interactional friction. Agency, in this configuration, does not reside exclusively in either the user or the system but emerges as distributed enactment within adaptive interaction loops. Adaptive response latency and action-path diversity thus become visible markers of this process, indicating that financial actions are no longer purely rule-driven but contextually enacted.

As informational traces accumulate, inferential uncertainty progressively narrows, making future user action trajectories increasingly predictable. System outcomes become more consistent across comparable situations, and explanations of decisions-implicit or explicit-grow more coherent over time. This stabilization reflects the emergence of value as identity-based informational persistence: the system's capacity to preserve and reuse structured regularities linking past interactions to future responses. Explanation consistency and confidence-interval contraction, therefore, are observable manifestations of this informational consolidation.

Trust does not precede this process; rather, it stabilizes through repeated and verifiable interaction (Borgogno & Colangelo, 2020; Casolaro et al., 2025). As outcomes become predictable in familiar contexts and deviations trigger exception-based responses rather than blanket procedural enforcement, expectations consolidate. Participation becomes self-reinforcing, extending beyond individual users to the network level, where collective interaction patterns further enrich the informational environment. Verification stability and cross-platform coherence thus render trust observable as a network-mediated property grounded in interactional reliability rather than institutional assertion.

4.2 System Observability and Implications

The proxies summarized in Table 1 are designed to capture system-level dynamics through their joint variation over time, making changes in the underlying relational–informational configuration observable.

Table 1: Embedded finance capabilities and embodied relational–informational functions

Embedded Capability	Relational Function	Emergent Property (Embodied)	Informational Systemic Function	Indicative proxy (interpretive role)
Programmability Standardized financial capabilities (e.g., risk, KYC, compliance)	**Distributed action**	**Agency** (distributed enactment)	**Variation generation** (introduction of action diversity)	**Adaptive latency & action-path diversity** Deviation from predefined scripts under contextual signals, rendering agency observable as situated enactment.
Interoperability Cross-actor technical integration via APIs	**Repeated participation**	**Trust** (network-mediated)	**Expectation stabilization** (coordination and reliability)	**Verification stability & cross-platform coherence** Consistency of outcomes across services over time, expressing trust as stabilized expectation rather than institutional assumption.
Integrability Native embedding within platform journeys	**Interactional regularities**	**Value** (identity-based informational persistence)	**Uncertainty reduction** (informative inference)	**Explanation consistency & confidence-interval contraction** Persistence of identity-linked regularities enabling sustained uncertainty reduction under explanatory constraints.

In empirical settings, the proxies in Table 1 can be observed in platform data.

Agency proxies describe how distributed enactment becomes observable within interaction flows, including user actions and response times. *Action-path diversity* and *adaptive latency* capture the range of potential user interactions (e.g., pay now or later, allocate savings, or select an investment option) and the relative reaction time to a system trigger (e.g., from a suggested option to the user's action). Value proxies track the persistence and predictive relevance of identity-linked information over time. *Explanation consistency* tracks the repeated reuse of stable behavioral features across similar cases, while *confidence-interval contraction* indicates the increasing predictive relevance of such information as uncertainty decreases. Trust emerges when the platform generates predictable and coherent expectations under similar conditions. *Verification stability* reflects the consistency of outcomes across repeated interactions, while *cross-platform coherence* indicates whether similar users or situations produce consistent outcomes across integrated services over time. Together, these indicators signal reliable and sustainable platform behavior.

When adaptive latency decreases while action-path diversity increases, agency is not simply accelerating—it is reconfiguring. When explanation consistency and confidence intervals stabilize, value is not merely accumulating—it is persisting as structured informational identity. When verification outcomes become coherent across services and exceptions replace procedural enforcement, trust is not assumed—it is enacted and stabilized.

Observing any single proxy in isolation is insufficient. Changes acquire meaning only when interpreted relationally within the feedback system described in Section 4.1. For example, increased participation without corresponding reductions in uncertainty may signal marketing-driven adoption rather than embodied stabilization, while uncertainty reduction without sustained participation may indicate brittle inference rather than trust formation. Embodied configurations are therefore identifiable through coordinated shifts across agency-, value-, and trust-related manifestations rather than through absolute levels.

From an IS perspective, this joint observability provides a diagnostic lens rather than a predictive model, aligning with platform governance approaches that emphasize adaptive monitoring metrics (Dinçkol et al., 2023) and systemic alignment with

human intent[2]. It enables researchers and practitioners to distinguish between embedded integrations—where financial functions remain modular and procedural—and embodied configurations—where interaction, inference, and participation become structurally coupled.

At the governance level, this distinction shifts attention from static rule compliance toward traceability, contestability, and transparency of adaptive responses in such systems. The framework thus does not prescribe optimization targets but offers a means to observe whether a system is transitioning toward embodied dynamics, and where breakdowns in the feedback loop may undermine sustainability, legitimacy, or user alignment.

5 Conclusions and Future Directions

This paper has shown that the digital transformation of financial services extends beyond the functional integration of banking capabilities into non-financial platforms. Moving past embedded finance as a primarily technological paradigm, it introduces embodied finance as a relational–informational configuration in which financial services acquire form through ongoing interaction among humans, machines, and platforms. Building on the machine–platform–crowd framework, the analysis clarified how agency, trust, and value are reconfigured in digital environments not as static institutional attributes but as *distributed and enacted system-level properties*, stabilized through interaction, inference, and informational continuity. From a governance perspective, this view highlights how such relational–informational configurations introduce boundary conditions centered on traceability, explainability, oversight, and the protection of identity-based informational persistence rather than on static compliance alone.

At the same time, the framework is conceptual and entails important limitations. Future research is needed to (i) formalize embodied finance within a coherent theoretical and information-theoretic architecture, (ii) empirically validate the proposed emergent properties through system-level and longitudinal analyses based on the joint temporal co-variation of the proxies outlined in Table 1, and (iii) examine structural risks and limits of embodied financial systems, including

[2] These implications resonate with European regulatory frameworks such as the AI Act, PSD2, and eIDAS, which emphasize traceability, contextual oversight, and accountability in socio-technical financial systems.

scalability, loss of user control, opacity of inference, and concentration of informational power. Addressing these directions is essential to assess whether embodied financial configurations can remain sustainable, governable, and aligned with human intent as platform-based finance evolves.

References

Aldasoro, I., Gambacorta, L., Korinek, A., Shreeti, V., & Stein, M. (2025). Intelligent financial system: How AI is transforming finance. Journal of Financial Stability, 81, 101472.

Aoki, M. (2011). Institutions as cognitive media between strategic interactions and individual beliefs. Journal of Economic Behavior & Organization, 79(1), 20–34.

Barroso, M., & Laborda, J. (2022). Digital transformation and the emergence of the fintech sector: A systematic literature review. Digital Business, 2(2), 100028.

Bartels, N., & Schmitt, A. (2022). Developing network effects for digital platforms in two-sided markets: The NfX construction guide. Digital Business, 2(2), 100044.

Basso, F., & Herrmann-Pillath, C. (2024). Embodiment, political economy and human flourishing: An embodied cognition approach to economic life. Palgrave Macmillan.

Benamati, J., & Serva, M. A. (2007). Trust and distrust in online banking: Their role in developing countries. Information Technology for Development, 13(2), 161–175.

Benbasat, I., Gefen, D., & Pavlou, P. A. (2008). Trust in online environments. Journal of Management Information Systems, 24(4), 5–11.

Birch, D. G. (2016). Identity is the new money. London Publishing Partnership.

Birch, D., & Rutter, K. (2025). Digital banking in the artificial intelligence era: Strategies for serving nonhuman customers. *Journal of Digital Banking, 10*(1).

Borgogno, O., & Colangelo, G. (2020). The data sharing paradox: BigTechs in finance. European Competition Journal, 16(2–3), 492–511.

Broekhoff, M.-C., van der Cruijsen, C., & de Haan, J. (2024). Towards financial inclusion: Trust in banks' payment services among groups at risk. Economic Analysis and Policy, 82, 104–123.

Casas-Cortés, M., Cañedo-Rodríguez, M., & Diz, C. (2023). Platform capitalism. In Oxford Research Encyclopedia of Anthropology. Oxford University Press.

Casolaro, A. M., Rauber, G. N., & de Lima, U. S. (2025). Open banking: A systematic literature review. Journal of Banking Regulation, 26(3), 340–355.

Castells, M. (1996). The rise of the network society. Blackwell.

Chiappino, S. (2026). A proxy-information bottleneck strategy for cognitive credit risk detection. Engineering Letters, 34(2), 576–590.

Dinçkol, D., Ozcan, P., & Zachariadis, M. (2023). Regulatory standards and consequences for industry architecture: The case of UK open banking. Research Policy, 52(6), 104760.

European Banking Authority. (2024, November). Special topic – Artificial intelligence. Retrieved January 31, 2026, from https://www.eba.europa.eu/publications-and-media/publications/special-topic-artificial-intelligence

Ferrari, V. (2022). The platformisation of digital payments: The fabrication of consumer interest in the EU fintech agenda. Computer Law & Security Review, 45, 105687.

Finken, S., & Hensen, J. (2025). Embedded finance: An evaluation of current and future use cases. Journal of Digital Banking, 9(4), 294–316.

Floridi, L. (2011). The philosophy of information. Oxford University Press.

Fuster, J. M. (2003). Cortex and mind: Unifying cognition. Oxford University Press.

Fuster, J. M. (2017). Prefrontal executive functions predict and preadapt. In E. Goldberg (Ed.), Executive functions in health and disease (pp. 3–19). Academic Press.

Hui, Y. (2014). Form and relation: Materialism on an uncanny stage. Intellectica, 61(1), 105–121.

Hutchins, E. (1995). Cognition in the wild. MIT Press.
Jafri, J. A., Amin, S. M., Rahman, A. A., & Nor, S. M. (2024). The role of trust and security on fintech adoption in banking: A systematic literature review. Heliyon, 10(1), e22980.
Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82(1), 35–45.
Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. American Economic Review, 75(3), 424–440.
Kopperapu, R. (2025). Embedded finance: How non-financial companies are shaping the future of financial services. Journal of Science & Technology, 10(2), 23–38.
Krakauer, D., Bertschinger, N., Olbrich, E., Flack, J. C., & Ay, N. (2020). The information theory of individuality. Theory in Biosciences, 139(2), 209–223.
Krakowski, S. (2025). Human–AI agency in the age of generative AI. Information and Organization, 35(1), 100560.
Lagarde, C. (2020). Payments in a digital world. European Central Bank. Retrieved January 31, 2026, from https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200910~31e6ae9835.en.html.
Latour, B. (2005). Reassembling the social: An introduction to actor-network theory. Oxford University Press.
Lin, X., Zhang, S. S., & Zachariadis, M. (2025). Open data and API adoption of U.S. banks. Journal of Financial Intermediation, 63, 101162.
Liu, S. (2025). The impact of fintech on financial consumer behavior: Digital payment platforms as an example. Advances in Economics, Management and Political Sciences, 225, 90–98.
März, G. (2025). Shaping participation: How digital platform design influences user participation. In A. Pucihar et al. (Eds.), 38th Bled eConference Proceedings (pp. 281–304). University of Maribor Press.
McAfee, A., & Brynjolfsson, E. (2017). Machine, platform, crowd: Harnessing our digital future. W. W. Norton & Company.
Mula, K. (2025). The transformative impact of open banking on financial services. World Journal of Advanced Engineering Technology and Sciences, 15(1), 1959–1969.
Murati, A., Skau, O., & De Geer, A. (2024). Embedded finance: How banks and customer platforms are converging. McKinsey & Company.
Murinde, V., Rizopoulos, E., & Zachariadis, M. (2022). The impact of the fintech revolution on the future of banking. International Review of Financial Analysis, 81, 102103.
Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). Platform revolution. W. W. Norton & Company.
Paul, M. (2022). Artificial intelligence in financial services: New risks and the need for more regulation? In S. Voeneky et al. (Eds.), The Cambridge handbook of responsible artificial intelligence (pp. 359–376). Cambridge University Press.
Pavlou, P. A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. Information Systems Research, 15(1), 37–59.
Ramadhan, A., Vidianto, M. A., Muharam, H., & Mawardi, W. (2025). Fintech and financial inclusion: Evidence from emerging markets. Research Horizon, 599–612.
Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423.
Sironi, P. (2022). Banks and fintech on platform economies. Wiley.
Srnicek, N. (2017). Platform capitalism. Polity Press.
Swartz, L. (2020). New money: How payment became social media. Yale University Press.
Tian, Y., Zhu, D., Nazar, R., Ali, S., & Mirza, A. (2026). From algorithms to access: The role of artificial intelligence in financial inclusion. Technology in Society, 84, 103098.
Varela, F. J., Rosch, E., & Thompson, E. (1991). The embodied mind: Cognitive science and human experience. MIT Press.

Zachariadis, M., & Zachariadis, P. (2017). The API economy and digital transformation in financial services. SWIFT Institute Working Paper.
Zeng, F., & Hu, B. (2025). The impact of social trust on the development of digital finance. International Journal of Financial Studies, 13(4), 232.
Zhang, Q., & Kim, H. (2025). From social to financial: Understanding trust in extended payment services on social networking platforms. Behavioral Sciences, 15(5), 659.
Zuboff, S. (2019). The age of surveillance capitalism. PublicAffairs.

About the author

Simone Chiappino is an independent researcher in digital finance and a senior executive in digital transformation within the financial services industry.

His work focuses on platform-based financial ecosystems, cognitive finance, and AI-enabled financial systems, with particular attention to how inference, interaction, and information shape decision-making and value creation.

He holds a Ph.D. in Science and Technology for Information and Knowledge from the University of Genoa (2015), an MBA from Bologna Business School, and degrees in Telecommunications Engineering. He has published in international journals and conferences on cognitive systems, information theory, and intelligent architectures.

Summary

This paper introduces embodied finance as a novel conceptual framework for understanding how financial services evolve in AI-enabled, platform-based environments. Moving beyond embedded finance as a model of technical integration, it argues that financial services increasingly take form through interactions among humans, machines, and platforms. The proposed machine–platform–crowd framework reconceptualizes agency, value, and trust as emergent system-level properties: agency as distributed enactment, value as identity-based informational persistence, and trust as network-mediated expectation stabilization. The paper also develops an observability perspective through illustrative proxies, showing how embodied configurations can be distinguished from modular embedded integrations. This conceptualization provides a basis for future empirical research on how agency, value, and trust evolve in AI-enabled financial ecosystems.

GREEN BEHAVIOR AND WELL-BEING: A BIBLIOMETRIC AND NETWORK ANALYSIS OF RESEARCH TRENDS

HAMED ABBASZADEH AVILAGH, PARINAZ PAZHUHANRAD, KRISTINA ERIKSSON-BACKA

Åbo Akademi University, Faculty of Human and Social Sciences, Turku, Finland
hamed.abbaszadehavilagh@abo.fi, parinaz.pazhuhanrad@abo.fi, kristina.eriksson-backa@abo.fi

Green behavior and well-being is an interdisciplinary research domain linking environmental sustainability with mental health, happiness, and overall quality of life. Over the past three decades, it has gained substantial scholarly attention and emerged as a significant contemporary field. To examine its intellectual structure and publication trends, this study applies a bibliometric analysis of Web of Science–indexed articles. Using keyword co-occurrence analysis, six major thematic clusters were identified: nature-based public health, sustainable lifestyles, sustainable workplaces, climate perception and behavior, nature-related well-being, and subjective well-being. These clusters demonstrate that the field integrates insights from public health, environmental psychology, organizational studies, climate psychology, and happiness research, highlighting its strongly interdisciplinary nature. The results indicate that green behavior and well-being is evolving into a coherent and structured domain of inquiry. Moreover, the findings emphasize the reciprocal relationship between environmental quality and human flourishing at individual, organizational, and societal levels.

DOI
https://doi.org/10.18690/um.fov.4.2026.4

ISBN
978-961-299-152-4

Keywords:
green behavior, well-being, bibliometric analysis, sustainability, network analysis

University of Maribor Press

1 Introduction

Global climate change is necessitating a fundamental shift in human behavior and consumption patterns toward environmental sustainability (Krumm, 2024). Previously, pro-environmental behavior was often viewed as a sacrifice, but contemporary research presents it as a positive "co-benefit," meaning that environmental actions can simultaneously benefit both the planet and individual well-being (Zawadzki et al., 2020).Empirical evidence suggests that the intention to engage in green behaviors such as preferring environmentally friendly products despite higher prices and purchasing green products is positively associated with life satisfaction (Xiao & Li, 2011).To analyze this complex relationship, it is crucial to clearly define the constructs of a green lifestyle and well-being. Pro-environmental behavior refers to actions intended to reduce negative impacts on the environment; however, its psychological payoff often depends less on the objective action itself and more on an individual's internal evaluation or green self-concept (Binder & Blankenberg, 2017a). Well-being is multidimensional and is typically divided into hedonic and eudaimonic components. The hedonic dimension, known as subjective well-being (SWB), includes cognitive evaluations, that is, "life satisfaction" and affective responses or "happiness or sense of belonging" (Welsch et al., 2021). In contrast, the eudaimonic tradition focuses on self-actualization and leading a meaningful life, measured through outcomes such as mental vitality and a sense of belonging (Guillen-Royo, 2019). Psychologically, engaging in non-material habits and activities that require "undivided attention", often found in green Activities, helps shift the stressed and reactive "red mind" into a more relaxed "green mind," thereby directly linking sustainable choices to improved neurological health (Pretty et al., 2017).

Binder et al. (2025) showed, based on data from a municipality in Germany, that the available evidence is not sufficient to confirm a strong and definitive positive relationship between pro-environmental behavior and life satisfaction. Their findings suggest that life satisfaction may not be driven solely by green behaviors and that other factors are likely to play a role in this relationship. Critical studies point to several important nuances. First, the positive link between well-being and pro-environmental behavior appears mainly in low-impact actions, such as small changes in consumption. There is no evidence that this relationship holds for behaviors that demand major effort or come with significant personal costs

(Krumm, 2024). Second, timing matters. When people expect very high levels of future well-being, those expectations can actually discourage them from engaging in pro-environmental behavior right now (Kaida & Kaida, 2016). Context also plays a major role. Participating in collective environmental activities, including activism, can sometimes be associated with lower life satisfaction. Yet this negative effect may weaken or even turn positive when individuals feel strongly connected to nature (Ibáñez-Rueda et al., 2020). Even so, recent quasi-experimental studies provide causal evidence that small efforts toward pro-environmental behavior can immediately lift positive emotions. For example, simply finding a new use for an old item can boost hedonic well-being (Prinzing, 2024). Longitudinal findings reinforce this insight. Engaging in pro-environmental behaviors on a daily basis is causally linked to higher eudemonic well-being, especially a greater sense of meaning in life and stronger feelings of social connectedness (Caldaroni et al., 2025).

Despite these insights, existing studies show substantial variation. They differ in context, define green behaviors in various ways-such as individual versus collective actions-and use different tools to measure well-being. As a result, the evidence is fragmented. A key issue is that many studies rely solely on cross-sectional or observational designs. Such approaches make it impossible to determine whether pro-environmental behaviors truly cause increases in well-being (Liu et al., 2022).

Among the 17 Sustainable Development Goals (SDGs) [see official UN link[1]] adopted by the member states of the United Nations in 2015, the relationship between SDG 3 (Good health and well-being) and 12 (Responsible consumption and production), is the main focus of the study by Siva Rama Krishna et al. (2024), which looks at how green behavior affects sustainable health outcomes. Despite this contribution, there is still a clear research gap regarding the relationship between green behavior and well-being because no study has directly performed a quantitative mapping or network-based analysis in this area. The increased focus on sustainability in recent years has resulted in a substantial and varied body of research in various contexts. But this quick growth has also caused a significant amount of fragmentation in the literature.

[1] https://sdgs.un.org/goals

Because of this, it is now more crucial than ever to conduct a thorough bibliometric analysis in the area of green behavior and wellbeing. By using quantitative scientific mapping and network analysis, such an analysis can make the structure of current research more understandable. The objective is to show how the relationship between green behavior and well-being has changed over time while also identifying new research trends, theoretical underpinnings, and important thematic clusters. Additionally, this method can draw attention to the fields and writers who have had the biggest impact on this multidisciplinary conversation. The research questions of this study are:

What are the main publication trends, influential journals, and leading contributors in the field connecting green behavior and well-being? Which countries, and collaborations have played key roles in shaping this research area? What are the dominant research themes, theoretical perspectives, and conceptual clusters linking green behavior and well-being? How has the intellectual and thematic structure of this field evolved over time, and what emerging topics indicate future research directions?

2 Methodology

Bibliometric analysis has become highly popular due to the accessibility of online research databases and increased research output (Zhang et al., 2024). This quantitative approach helps researchers analyze the output of academia by identifying publication trends, mapping influential authors or institutions, and examining citation patterns (Liu et al., 2019). Following Donthu et al. (2021), this study uses performance analysis, science mapping, and network analysis to assess and visualize the current state of research on green behavior and well-being. We applied these bibliometric techniques to analyze research outputs concerning green behavior within the field of well-being. To achieve this goal, data analysis was performed in the Google Colab environment using Python (version 3.11). In this regard, the Pandas, NumPy, Matplotlib, and Seaborn libraries were used. Also, VOSviewer (version 1.6.20) software was used for clustering and visualization of scientific networks and portraying thematic clusters and influential countries.

Data collection

Web of Science, established in 1964, is one of the earliest and most authoritative citation databases. Given the interdisciplinary nature of this study, it was selected as the primary data source due to its selective indexing policy (Singh et al., 2021) and the depth and quality of its citation data (Aghaei Chadegani et al., 2013). Moreover, Web of Science is widely employed in bibliometric research because of its rigorous indexing standards, comprehensive citation metadata, and cross-disciplinary consistency, making it a reliable foundation for quantitative scientific mapping. Keywords were extracted through a review of the sustainability and well-being research area. We refined this initial list through iterative testing in the Web of Science database to ensure that the final search terms were both relevant to the topic and reflective of the broader academic discourse. The first component included terms related to green and sustainability concepts and their synonyms. The second focused on behavior and its related constructs. The third encompassed well-being and associated terms to capture the majority of relevant studies. The final component excluded terms unrelated to the study's objective, such as animal, fish, plant, and similar keywords. Applying the following search string yielded 11,696 records: ("green" OR "eco-friendly" OR "sustainable" OR "pro-environmental" OR "environmentally conscious" OR "eco conscious" OR "environmentally responsible" OR "sustainability-oriented") AND ("behavio?r" OR "attitude*" OR "intention*" OR "belief*" OR "perception*" OR "practice*" OR "lifestyle*" OR "decision*" OR "choice*") AND ("wellbeing" OR "well-being" OR "subjective wellbeing" OR "subjective well-being" OR "mental health" OR "quality of life" OR "happiness" OR "life satisfaction") NOT (plant* OR animal* OR species OR fish OR marine OR soil OR agriculture OR forestry OR "energy efficiency" OR "energy consumption" OR "waste management" OR "air pollution" OR "carbon emission*" OR "emission reduction" OR "building design" OR "urban planning" OR "renewable energy").

Aiming to focus on human behavior and psychological well-being, this research does not delve into areas such as urban planning or building design. Narrowing the scope of the research in this way ensures that our findings remain aligned with the main goal of the research, which is to understand sustainability at the individual level. Subsequently, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) procedure was employed to refine the dataset and identify the

final sample of publications. Several filters were applied. First, only peer-reviewed articles published in English were retained. Given the interdisciplinary scope of green behavior (Li et al., 2019) and well-being (Fabian, 2021; Sigfusdottir et al., 2017), relevant research domains included psychology, sociology, business and economics, public, environmental, and occupational health, and social sciences. In the final stage, documents were screened based on keywords and abstracts, followed by full-text assessment to confirm relevance. Screening of titles, abstracts, and full text of articles was performed by the authors to ensure consistent application of inclusion and exclusion criteria. The selection process was guided by PRISMA protocols and involved systematic steps, including identification, removal of duplicates, and rigorous screening of eligible articles. Figure 1 provides a graphical detail of how the final sample was obtained. This process resulted in a final dataset of 3,210 articles.

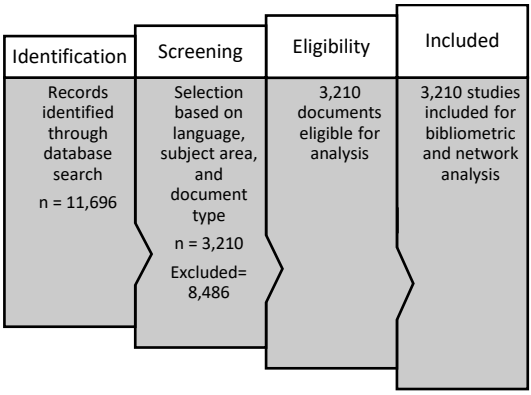

Figure 1: PRISMA technique for selecting documents
Source: Authors' elaboration based on PRISMA procedure

3 Results and Finding

Table 1 summarizes the bibliometric characteristics of the 3,210 documents identified. The dataset spans the period 1986–2025, encompassing the full range of publications retrieved from the database based on the applied search query. At the time of data collection (November 13, 2025), the most cited article had received 1,703 citations, with an average of 22.48 citations per document. The International Journal of Environmental Research and Public Health is the most influential outlet

in terms of publication volume, while Heesup Han is the most active author, with 25 publications.

Table 1: overall overview of the dataset

Metric	Value
Total Publications	3,210
Year Range	1986-2025
Total Unique Authors	12,070
Total Journals	963
Document Average Age	4.03
Total Keywords	9,710
Average Authors per Paper	4.07
Average Citations per Paper	22.48
Total Citations	72,168
Most Cited Paper (Citations)	1,703
Most Productive Author	Han, Heesup
Most Productive Author Papers	25
Most Common Journal	*International Journal of Environmental Research and Public Health*

Source: Authors' calculations based on Web of Science dataset

3.1 Research output trend

Figure 2 shows the trend of publications over the years. This trend started in 1986 with one publication, and until 1993, there were no other publications. However, since the 1990s, a slow but steady research stream continued until the mid-2010s. By that time, yearly publication output never reached 50 per year, but from 2015 to the present, the research in this field has grown substantially, and in 2025, a record number of 685 articles was reached.

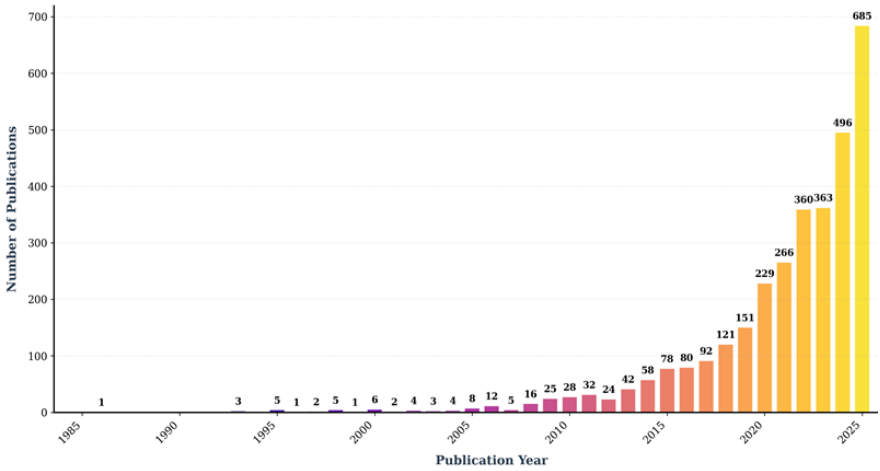

Figure 2: research trend of green behavior and well-being
Source: Authors' calculations

3.2 Top Article Analysis

The table 2 shows the top-cited articles in the field of green behavior and well-being. Lyubomirsky et al. (2005), with the article titled "Pursuing happiness: The architecture of sustainable change" and 1,703 citations, is the most cited work, receiving almost twice as many citations as the second article. According table 2, the top ten most cited articles follow two main research streams. One stream is grounded in environmental psychology and public health, while the other focuses on green business and sustainable management. In these studies, nature connectedness, which refers to how individuals perceive their relationship with the natural environment, is directly examined in some of the top-cited articles such as Capaldi et al. (2014) and Martin et al. (2020). This concept is also closely related to other themes found in the table, including green space and neighborhood greenness, which are commonly used to explore how contact with nature can enhance happiness, well-being, and pro-environmental behavior. Despite our interdisciplinary Approach in selecting research areas, Table 3 shows that most of the top-cited articles are published in the fields of psychology, public health, and sustainability.

Table 2: Top-citation green behavior and well-being

Title	Authors	Year	Citations	Journal	Journal IF*
Pursuing happiness: The architecture of sustainable change	Lyubomirsky, S; Sheldon, KM; Schkade, D	2005	1,703	Review of General Psychology	3.2
Mainstreaming Green Product Innovation: Why and How Companies Integrate Environmental Sustainability	Dangelico, R & Pujari, D	2010	859	Journal of Business Ethics	6.7
Green space as a buffer between stressful life events and health	van den Berg, AE; Maas, J; Verheij, RA; Groenewegen, PP	2010	804	Social Science & Medicine	5
Morbidity is related to a green living environment	Maas, J; Verheij, RA; de Vries, S; Spreeuwenberg, P; Schellevis, FG; Groenewegen, PP	2009	754	Journal of Epidemiology and Community Health	3.7
The relationship between nature connectedness and happiness: A meta-analysis	Capaldi, CA; Dopko, RL; Zelenski, JM	2014	748	Frontiers in Psychology	2.9
Associations of neighborhood greenness with physical and mental health: do walking, social coherence and local social interaction explain the relationships?	Sugiyama, T; Leslie, E; Giles-Corti, B; Owen, N	2008	616	Journal of Epidemiology and Community Health	3.7
Understanding how consumers view green hotels: how a hotel's green image can influence behavioural intentions	Lee, JS; Hsu, LT; Han, H; Kim, Y	2010	597	Journal of Sustainable Tourism	7.8
Are psychological and ecological well-being compatible? The role of values, mindfulness, and lifestyle	Brown, KW; Kasser, T	2005	590	Social Indicators Research	2.8
Nature contact, nature connectedness and associations with health, wellbeing and pro-environmental behaviours	Martin, L; White, MP; Hunt, A; Richardson, M; Pahl, S; Burt, J	2020	552	Journal of Environmental Psychology	7
Building Sustainable Organizations: The Human Factor	Pfeffer, J	2010	543	Academy of Management Perspectives	5.7

Source: Authors' calculations

***Table note:** Journal IF refers to the journal's Impact Factor, which indicates the average number of citations received per paper published in that journal during the previous two years.

3.3 Top Journal Analysis

Table 3: top 10 most influential journal of green behavior and well-being

Journal Name	Number of Publications	Percentage (%)	Total Citations	Avg Citations	Journal impact factor	publisher	SJR* 2024
International Journal of Environmental Research and Public Health	188	5.86	4522	24.05	0.919	MDPI	Q2
Frontiers in Psychology	111	3.46	3205	28.87	0.872	Frontiers Media SA	Q2
Frontiers in Public Health	83	2.59	661	7.96	1.027	Frontiers Media SA	Q1
BMC Public Health	82	2.55	1622	19.78	1.359	BioMed Central Ltd	Q1
Ecological Economics	58	1.81	3078	53.07	2.088	Elsevier	Q1
Journal of Environmental Psychology	44	1.37	2132	48.45	2.208	Academic Press	Q1
Journal of Sustainable Tourism	43	1.34	2499	58.12	2.733	Taylor and Francis Ltd.	Q1
Social Indicators Research	41	1.28	1568	38.24	0.933	Springer Netherlands	Q1
Health & Place	32	1.00	1838	57.44	1.456	Elsevier	Q1
Health Promotion International	24	0.75	380	15.83	0.830	Oxford University Press	Q1

Source: Authors' calculations
Table note: SJR* 2024 indicates the SCImago Journal Rank score for each journal based on Scopus data

Table 3 displays the top ten journals contributing most to the field of green behavior and well-being. The top position on this list goes to the International Journal of Environmental Research and Public Health (IJERPH), with 188 publications and an average of 24.05 citations. This citation rate is lower than that of six other journals with fewer Publications. This discrepancy can be seen in the next three ranks: Frontiers in Psychology, Frontiers in Public Health, and BMC Public Health. Since these three journals are fully open access, their articles benefit from wider visibility and accessibility, which can naturally lead to higher citation rates compared to journals that have not always followed the same access model. This can also be interpreted in relation to the varying degrees of specialization among journals. IJERPH publishes a broad range of topics and has the highest productivity, whereas

journals such as Ecological Economics and Health & Place publish fewer but more targeted studies within their specific domains. This narrower scope can result in higher citations per publication because the articles are more focused and directly relevant to specialized audiences. The result shows that public health journals like IJERPH, Frontiers in Public Health, BMC Public Health, Health Promotion International, and Health & Place account for 409 of the research output, which means 13% of all research output in this field. Among the publishers, Elsevier and Taylor and Francis, although having fewer contributions, have publications that are more influential in terms of average citations.

3.4 International Collaboration Network

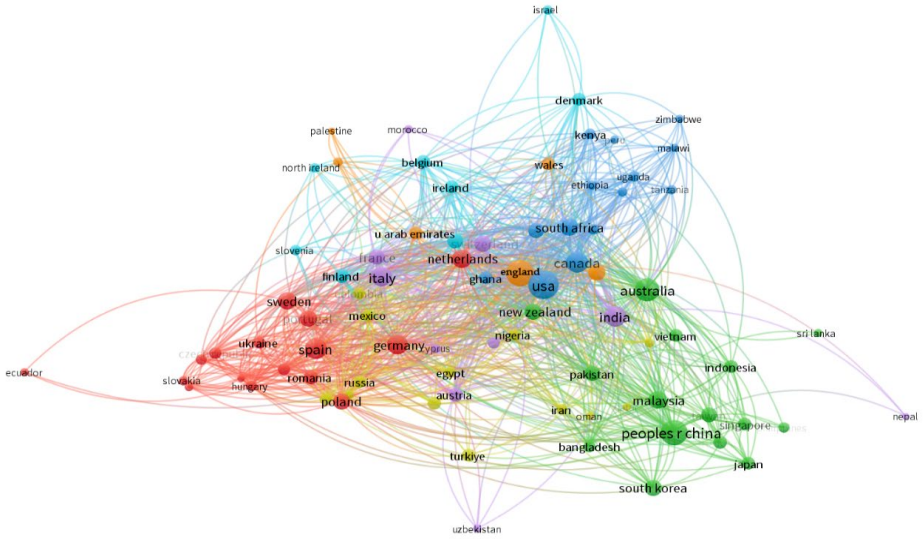

Figure 3: Global collaboration network of green behavior and well-being
Source: Authors' visualization using VOSviewer

Figure 3 presents a bibliometric map of international collaboration among countries in the field of green behavior and well-being. The VOSviewer program was used to generate this figure, with a threshold set at five articles, in order to focus on countries with a meaningful level of research output and to avoid visual clutter in the collaboration network. Out of 147 countries with publications in this field, only 82 met this criterion. Regarding documents, the USA leads with 618 publications,

followed by England (439), China (365), and Australia (283). For total citations, the USA, with 18,378 citations, remains the top country, followed by England (12,485), Australia (8,774), and Canada (7,511). As is clear, the USA is the leading power in the field and has the greatest influence on shaping the direction of the research. Strong co-authorship links of these countries with others show they serve as global research hubs, enabling regional knowledge sharing and facilitating collaborative research efforts.

3.5 Keyword Co-Occurrence and Thematic Clusters

The result of the co-occurrence analysis of author keywords using VOSviewer is shown in Figure 4. A minimum occurrence threshold of 15 was applied, meaning that only keywords appearing at least 15 times in the dataset were included. This value was chosen to ensure meaningful representation of frequently used terms and to prevent visual clutter on the map, in line with common practice in bibliometric mapping. Before using VOSviewer, a structured data-cleaning process was conducted, which reduced the initial 9,710 keywords to 9,628. This process included merging singular and plural forms, harmonizing British and American spellings, expanding abbreviations, and removing duplicates or irrelevant terms. Such steps are recommended in bibliometric methodology, as variations in keyword formats can distort co-occurrence structures and reduce the reliability of the analysis. (Donthu et al., 2021) As shown in Figure 4 and Table 4, six clusters with different colors and thematic labels were generated. The first cluster is the red cluster labelled Nature-based Public Health, comprising 20 items. It is the largest cluster with a biomedical and psychological focus, indicated by keywords such as *anxiety, depression, obesity, stress, and mental health. Keywords like green spaces, built environment, and nature point to the role of urban planning interventions in supporting community mental health.*

The second cluster is labelled Sustainable Lifestyles with a green color, due to the presence of keywords such as *lifestyle, values, adolescents, community, and education.* These keywords highlight the sociological aspects of sustainability. The presence of keywords like *sustainable tourism and tourism impacts* shows the depth of the sustainable lifestyle and indicates that this aspect is not limited to household-based issues.

The blue cluster, labelled Sustainable Workplaces, contains keywords related to the occupational aspects of sustainability, such as *burnout, job satisfaction, resilience, work*

engagement, and sustainable HRM. Clearly, this research area emphasizes how organizational and managerial sustainability can impact employee well-being and resilience. The presence of the COVID-19 keyword in this cluster can be an indication of the effect of the crisis on workplace sustainability.

The yellow cluster labelled Climate Perception & Behavior includes keywords like *climate change, eco-anxiety, emotion, materialism, and pro-environmental behavior.* This suggests that environmental changes can trigger specific perceptions and green behaviors. Overall, this cluster indicates the potential for expanding climate psychology research.

Purple cluster with seven items, including *positive psychology, nature connectedness, green exercise, and decision-making,* is labelled Nature-Connected Well-being. This cluster demonstrates that well-being is closely tied to engaging with the natural environment. Connecting with nature can enhance emotional health and psychological resilience.

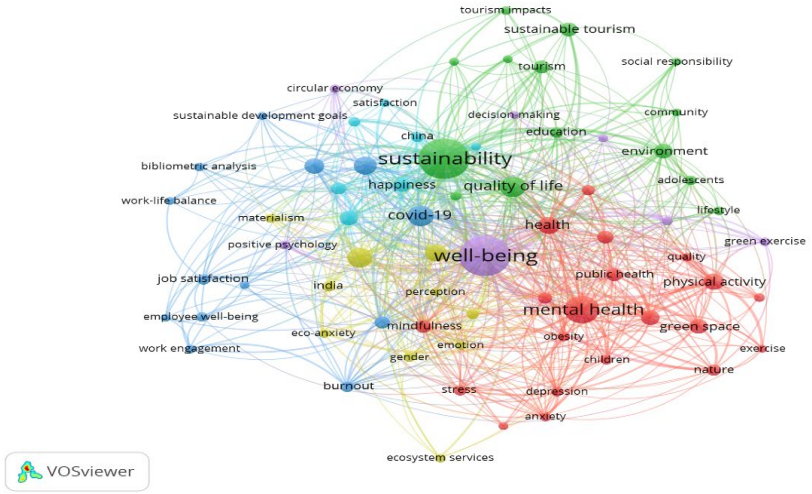

Figure 4: Keyword co-occurrence network of green behavior and well-being
Source: Authors' visualization using VOSviewe

Lastly, a light blue cluster with seven items is labelled Subjective Well-being. This cluster covers personal experiences of well-being, such as *happiness and life satisfaction.*

Additionally, the inclusion of the *machine learning* keyword indicates a data-driven approach in green behavior and well-being research. The presence of the keyword *China* in this cluster highlights China's geographical focus in this area.

Table 4: Thematic clusters derived from the co-occurrence network

Cluster	color	Labels	Items
Cluster 1 (20 items)	Red	Nature-based Public Health	built environment; anxiety; children; depression; exercise; green space; health; health promotion; implementation; mental health; mindfulness; nature; obesity; older adults; physical activity; public health; qualitative research; quality; stress; urban green spaces
Cluster 2 (14 items)	Green	Sustainable Lifestyles	adolescents; community; education; environment; environmental sustainability; innovation; lifestyle; quality of life; social responsibility; sustainability; sustainable tourism; tourism; tourism impacts; values
Cluster 3 (13 items)	Blue	Sustainable Workplaces	bibliometric analysis; burnout; corporate social responsibility; covid-19; employee well-being; job satisfaction; leadership; resilience; sustainable development goals; sustainable HRM; work engagement; work-life balance
Cluster 4 (10 items)	Yellow	Climate Perception & Behavior	Climate change; eco anxiety; ecosystem service; emotion; gender; India; materialism; perception; self-efficiency ; pro environmental behavior;
Cluster 5 (7 items)	Purple	Nature-connected Well-being	positive psychology; well-being; nature connectedness; green exercise; equity; decision making; circular economy
Cluster 6 (7 items)	Light blue	Subjective Well-being	China; happiness; life satisfaction; machine learning; place attachment; satisfaction; sustainable consumption

Source: Authors' calculations

3.6 Thematic Evolution Heatmap

Figure 5 presents the evaluation of the top 20 keywords from 1986 to the present. The keyword heatmap shows a consistent but low presence of green behavior and well-being–related keywords until 2014. However, after 2014–2015, an explosive growth occurred across almost all of the top twenty keywords. Sustainability, well-being, and mental health became the dominant themes in the field, and after 2019, a noticeable increase in their frequency appears. The post-2020 period can be considered the golden era of this field, as most keywords reached their highest values, indicating greater researcher interest, more diversified research output, and further consolidation.

The keyword "sustainable development goals" emerged after 2015, reflecting its entry into the literature on green behavior and well-being immediately following the introduction of the United Nations' 17 Sustainable Development Goals in March

2015 (Hák et al., 2016) Interestingly, during the same period as COVID-19–related keywords peaked around 2021, keywords such as happiness, mindfulness, and quality of life also appeared. This indicates a shift toward stress-reduction and sustainability-linked well-being research, which became a major focus during and after the pandemic, a period that in many countries involved lockdowns and prolonged stay-at-home conditions.

Figure 5: Top 20 keywords evolution heatmap
Source: Authors' visualization using Python

4 Conclusion

This research aimed to map the green behavior and well-being research output to date. The results highlighted driving themes of the field, current trends, influential journals, authors, and papers. According to the findings, it can be concluded that the year 2015 marks a clear turning point in this field, as it coincides with the introduction of the UN Sustainable Development Goals, which triggered a substantial rise in research output. Since then, keywords such as sustainability and well-being have appeared in nearly every publication, reflecting a broad integration of these themes into the core of scholarly work. The prominence of these concepts surged even further following the COVID-19 pandemic, suggesting a growing recognition that time spent in nature and eco-friendly lifestyle choices can serve as meaningful pathways to enhanced mental well-being and improved overall quality of life.

Research is mainly categorized into six main thematic clusters, namely: Nature-based Public Health, Sustainable Lifestyles, Sustainable Workplaces, Climate Perception & Behavior, Nature-connected Well-being, and Subjective Well-being. Despite the global interest in green behavior and the well-being area, the USA, England (as separate from the UK), China, and Australia dominate the global collaboration networks.The findings of this paper are novel and, in some cases, align with previous studies. In most of the studies with a bibliometric approach with a sustainability focus, the mid-2010s marks a turning point with a surge in research. This trend is reflected in studies such as Si et al. (2019), Font-Barnet & Nel-lo Andreu (2021), Hák et al. (2016), and Bai et al. (2025). Six different clusters also depict a thematic picture of the green behavior and sustainability field. The idea of the first cluster is supported by studies by Lackey et al. (2019) and Bratman et al. (2012), which emphasize nature's uplifting role and the promotion of mental health. The cluster of sustainable lifestyles is also discussed in Brülde's (2014) book, where he explains that sustainable lifestyles can lead to well-being and a happy life in the long term. A sustainable workplace is also another thematic cluster that was discussed in Cinar & Bilodeau's (2022) study. According to their research, a sustainable workplace positively impacts mental health and well-being, which are essential components for achieving sustainable development and are addressed in SDG 3 (Cinar & Bilodeau, 2022). The SDG 3 promotes good health and well-being, covering physical and mental wellbeing of individual and society (Mexico et al., 2018). The fourth cluster,

i.e., climate perception and behavior, was discussed in a systematic literature review covering the years 2012 to 2022, and it shows that climate change perception correlates with lower resilience and well-being (Gianfredi et al., 2024). The fifth cluster includes keywords indicating that nature connection enhances well-being. The findings of Capaldi et al. (2015) showed that disconnection from nature can lead to declines in both hedonic and eudaimonic well-being. The last cluster from the co-occurrence analysis was labeled Subjective Well-being. To measure subjective well-being, Steverink et al. (2001) focused on aging and individual resources, while Diener (1998) discussed that personality is an important factor in predicting long-term well-being. Additionally, more recent studies, such as the research by Jaidka et al. (2020), used machine learning approaches to measure subjective well-being, which was also mentioned in the sixth cluster.

These six clusters show how green behavior and well-being span different research disciplines. The results showed that this field is becoming an independent research area and should not be considered as an extension to sustainability or environmental psychology. The thematic breadth connects different research fields such as public health, pro-environmental behavior, organizational science, climate psychology, and subjective well-being, and understanding this field needs an interdisciplinary perspective. The similarity of the foundation of green behavior and well-being can be seen as the shift in the paradigm where environmental quality and human flourishing are increasingly studied as interdependent phenomena at the individual, organizational, and social levels. In the end, the evidence points to something important: living more sustainably is not just about protecting the planet; it also supports our own sense of fulfillment. By pulling together these diverse threads into one clear picture, this analysis gives future researchers a strong starting point and offers practical ideas for policies that can protect the environment while looking after people's mental health in a world facing serious challenges.

This study presents comprehensive bibliometric analyses of green behavior and well-being research, covering over thirty years of research. It highlights key authors, articles, journals, and countries, and employs co-occurrence analysis to identify research trends in the field. These clusters demonstrate the popularity of each theme, and future researchers can use these themes as guidelines to position their research within the existing body of knowledge. This study, in addition to expanding current bibliometric research, offers a comprehensive approach to the field of green

behavior and well-being, emphasizing the interconnection of different disciplines and methodologies. Despite the novelty of this research, this study is also constrained by some limitations. This study exclusively relied on Web of Science data, which may limit the study to capturing all publications in the field of green behavior and well-being.

Even though this study is a novel approach to the field of well-being and green behavior, future studies can promote the findings by considering the following suggestion. Relevant publications should encompass other databases, such as Scopus, PubMed, PsycINFO, ERIC, or those published in non-indexed outlets. Additionally, the analysis is affected by well-known citation biases that favor English-language journals and well-funded institutions. This can strengthen the advantage of historically prominent countries like the UK and the USA, while failing to recognize researchers in underrepresented regions, who often face citation disadvantages. To resolve this, future research should take into account the altmetrics or normalization methods to minimize these biases. Also, combining this technique with qualitative methodologies will improve the validity of the approach.

References

Aghaei Chadegani, A., Salehi, H., Md Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). A comparison between two main academic literature collections: Web of Science and Scopus databases. Asian Social Science, 9(5), 18–26. https://doi.org/10.5539/ass.v9n5p18.

Bai, J., Wang, Y., & Chang, Q. (2025). Cultural ecosystem services of urban green space for human well-being: A literature review based on bibliometrics and knowledge mapping analysis. Transactions in Earth, Environment, and Sustainability, 2(3-4), 191-212. https://doi.org/10.1177/2754124X251331602 (Original work published 2024)

Binder, M., & Blankenberg, A.-K. (2017). Green lifestyles and subjective well-being: More about self-image than actual behavior? Journal of Economic Behavior & Organization, 137, 304–323. https://doi.org/10.1016/j.jebo.2017.03.009

Binder, M., Blankenberg, A.-K., & Nickel, J. (2025). Pro-environmental behavior and life satisfaction: How strong is our evidence? Ecological Economics, 237, 108684. https://doi.org/10.1016/j.ecolecon.2025.108684

Bratman, G. N., Hamilton, J. P., & Daily, G. C. (2012). The impacts of nature experience on human cognitive function and mental health. Annals of the New York Academy of Sciences, 1249(1), 118-136. https://doi.org/10.1111/j.1749-6632.2011.06400.x

Brown, K. W., & Kasser, T. (2005). Are psychological and ecological well-being compatible? The role of values, mindfulness, and lifestyle. Social Indicators Research, 74(2), 349–368. https://doi.org/10.1007/s11205-004-8207-8 IDEAS/RePEc

Brülde, B. (2014). Well-being, happiness, and sustainability. In Well-being in contemporary society (pp. 157-176). Cham: Springer International Publishing.

Caldaroni, S., Gerbino, M., Schmiedek, F., Neubauer, A. B., Manfredi, L., Gregori, F., Pastorelli, C., Corbelli, G., & Zuffianò, A. (2025). The positive effect of pro-environmental behavior on eudaimonic well-being in young adults: A daily diary study using the within-person encouragement design. Journal of Personality. https://doi.org/10.1111/jopy.13021

Capaldi, C. A., Dopko, R. L., & Zelenski, J. M. (2014). The relationship between nature connectedness and happiness: A meta-analysis. Frontiers in Psychology, 5, 976. https://doi.org/10.3389/fpsyg.2014.00976 Frontiers

Capaldi, C. A., Passmore, H.-A., Nisbet, E. K., Zelenski, J. M., & Dopko, R. L. (2015). Flourishing in nature: A review of the benefits of connecting with nature and its application as a wellbeing intervention. International Journal of Wellbeing, 5(4), 1–16. https://doi.org/10.5502/ijw.v5i4.449

Chadegani, A. A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ebrahim, N. A. (2013). A comparison between two main academic literature collections: Web of Science and Scopus databases. arXiv preprint arXiv:1305.0377

Cinar, A. B., & Bilodeau, S. (2022). Sustainable workplace Mental well Being for sustainable SMEs: How? Sustainability, 14(9), 5290. https://doi.org/10.3390/su14095290

Dangelico, R. M., & Pujari, D. (2010). Mainstreaming green product innovation: Why and how companies integrate environmental sustainability. Journal of Business Ethics, 95(3), 471–486. https://doi.org/10.1007/s10551-010-0434-0 SpringerLink

Diener, E. (1998). Subjective well-being and personality. In Advanced personality (pp. 311-334). Boston, MA: Springer US.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Fabian, M. (2021). Improving interdisciplinary research in well-being—A review with further comments of Michael Bishop's The Good Life: Unifying the Philosophy and Psychology of Well-Being. Journal of Happiness Studies, 22(6), 2829–2844. https://doi.org/10.1007/s10902-020-00333-6

Font-Barnet, A., & Nel-lo Andreu, M. (2021). Research on tourism, well-being, and nature: a bibliometric analysis. Anatolia, 34(2), 163–175. https://doi.org/10.1080/13032917.2021.2002699

Gianfredi, V., Mazziotta, F., Clerici, G., Astorri, E., Oliani, F., Cappellina, M., Catalini, A., Dell'Osso, B. M., Pregliasco, F. E., Castaldi, S., & Benatti, B. (2024). Climate Change Perception and Mental health. results from a Systematic Review of the Literature. European Journal of Investigation in Health, Psychology and Education, 14(1), 215–229. https://doi.org/10.3390/ejihpe14010014

Guillen-Royo, M. (2019). Sustainable consumption and wellbeing: Does on-line shopping matter? Journal of Cleaner Production, 229, 1112–1124. https://doi.org/10.1016/j.jclepro.2019.05.061

Hák, T., Janoušková, S., & Moldan, B. (2016). Sustainable development goals: A need for relevant indicators. Ecological Indicators, 60, 565–573. https://doi.org/10.1016/j.ecolind.2015.08.003

Ibáñez-Rueda, N., Guillén-Royo, M., & Guardiola, J. (2020). Pro-environmental behavior, connectedness to nature, and wellbeing dimensions among Granada students. Sustainability, 12(21), 9171. https://doi.org/10.3390/su12219171

Jaidka, K., Giorgi, S., Schwartz, H. A., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2020). Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. Proceedings of the National Academy of Sciences, 117(19), 10165–10171. https://doi.org/10.1073/pnas.1906364117

Kaida, N., & Kaida, K. (2016). Pro-environmental behavior correlates with present and future subjective well-being. Environment, Development and Sustainability, 18(1), 111–127. https://doi.org/10.1007/s10668-015-9629-y

Krumm, L. (2024). The relationship between pro-environmental behavior, subjective well-being, and environmental impact: a meta-analysis. Environmental Research Letters, 19(9), 094056. https://doi.org/10.1088/1748-9326/ad6888

Lackey, N. Q., Tysor, D. A., McNay, G. D., Joyner, L., Baker, K. H., & Hodge, C. (2019). Mental health benefits of nature-based recreation: a systematic review. Annals of Leisure Research, 24(3), 379–393. https://doi.org/10.1080/11745398.2019.1655459

Lee, J. S., Hsu, L.-T., Han, H., & Kim, Y. (2010). Understanding how consumers view green hotels: How a hotel's green image can influence behavioural intentions. Journal of Sustainable Tourism, 18(7), 901–914. https://doi.org/10.1080/09669581003777747 Taylor & Francis Online

Liu, W., Wang, J., Li, C., Chen, B., & Sun, Y. (2019). Using bibliometric analysis to understand the recent progress in agroecosystem services research. Ecological Economics, 156, 293–305.

Liu, Y., Cleary, A., Fielding, K. S., Murray, Z., & Roiko, A. (2022). Nature connection, pro-environmental behaviours and wellbeing: Understanding the mediating role of nature contact. Landscape and Urban Planning, 228, 104550. https://doi.org/10.1016/j.landurbplan.2022.104550

Lyubomirsky, S., Sheldon, K. M., & Schkade, D. (2005). Pursuing happiness: The architecture of sustainable change. Review of General Psychology, 9(2), 111–131. https://doi.org/10.1037/1089-2680.9.2.111 SAGE Journals+2ResearchGate+2

Maas, J., Verheij, R. A., de Vries, S., Spreeuwenberg, P., Schellevis, F. G., & Groenewegen, P. P. (2009). Morbidity is related to a green living environment. Journal of Epidemiology and Community Health, 63(12), 967–973. https://doi.org/10.1136/jech.2008.079038 PubMed

Martin, L., White, M. P., Hunt, A., Richardson, M., Pahl, S., & Burt, J. (2020). Nature contact, nature connectedness and associations with health, wellbeing and pro-environmental behaviours. Journal of Environmental Psychology, 68, 101389. https://doi.org/10.1016/j.jenvp.2020.101389 Pearl

Pfeffer, J. (2010). Building sustainable organizations: The human factor. Academy of Management Perspectives, 24(1), 34–45. https://doi.org/10.5465/amp.24.1.34

Pretty, J., Rogerson, M., & Barton, J. (2017). Green mind theory: How brain-body-behaviour links into natural and social environments for healthy habits. International Journal of Environmental Research and Public Health, 14(7), 706. https://doi.org/10.3390/ijerph14070706

Prinzing, M. M. (2024). Proenvironmental behavior increases subjective well-being: Evidence from an experience-sampling study and a randomized experiment. Psychological Science, 35(9), 951–961. https://doi.org/10.1177/09567976241251766

Si, H., Shi, J.-g., Tang, D., Wen, S., Miao, W., & Duan, K. (2019). Application of the theory of planned behavior in environmental science: A Comprehensive Bibliometric Analysis. International Journal of Environmental Research and Public Health, 16(15), 2788. https://doi.org/10.3390/ijerph16152788

Sigfusdottir, I. D., Kristjansson, A. L., Thorlindsson, T., & Allegrante, J. P. (2017). Stress and adolescent well-being: the need for an interdisciplinary framework. Health promotion international, 32(6), 1081-1090.

Singh, V.K., Singh, P., Karmakar, M. et al. The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. Scientometrics 126, 5113–5142 (2021). https://doi.org/10.1007/s11192-021-03948-5

Siva Rama Krishna, J., Kumar, S., & Kirmani, M. D. (2024). Green behaviour engagement towards the achievement of sustainable health: A systematic review and bibliometric analysis. Social Marketing Quarterly, 30(4), 183–222. https://doi.org/10.1177/15245004241289389

Steverink, N., Westerhof, G. J., Bode, C., & Dittmann-Kohli, F. (2001). The personal experience of aging, individual resources, and subjective well-being. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 56(6), P364-P373.

Sugiyama, T., Leslie, E., Giles-Corti, B., & Owen, N. (2008). Associations of neighbourhood greenness with physical and mental health: Do walking, social coherence, and local social

interaction explain the relationships? Journal of Epidemiology and Community Health, 62(5), e9. https://doi.org/10.1136/jech.2007.064287 PubMed

van den Berg, A. E., Maas, J., Verheij, R. A., & Groenewegen, P. P. (2010). Green space as a buffer between stressful life events and health. Social Science & Medicine, 70(8), 1203–1210. https://doi.org/10.1016/j.socscimed.2010.01.002 nivel.nl+1

Welsch, H., Binder, M., & Blankenberg, A.-K. (2021). Green behavior, green self-image, and subjective well-being: Separating affective and cognitive relationships. Ecological Economics, 179, 106854. https://doi.org/10.1016/j.ecolecon.2020.106854

Xiao, J. J., & Li, H. (2011). Sustainable consumption and life satisfaction. Social Indicators Research, 104(2), 323–329. https://doi.org/10.1007/s11205-010-9746-9

Zawadzki, S. J., Steg, L., & Bouman, T. (2020). Meta-analytic evidence for a robust and positive association between individuals' pro-environmental behaviors and their subjective wellbeing. Environmental Research Letters, 15(12), 123007. https://doi.org/10.1088/1748-9326/abc4ae

Zhang, X., Panatik, S. A., & Zhang, N. (2024). Employee green behavior: bibliometric-content analysis. Heliyon, 10(10), e31045. https://doi.org/10.1016/j.heliyon.2024.e31045

Declaration of AI Use

The authors used AI-assisted tools for language editing and improving the clarity and readability of the manuscript. The intellectual content, research design, analysis, interpretation, and conclusions are the authors' own.

Designing a Layered Learning Infrastructure with a Digital Twin Component

Satu Aksovaara, Hannu Haapala,
Minna Silvennoinen
Jamk University of Applied Sciences, Jyväskylä, Finland
satu.aksovaara@jamk.fi, hannu.haapala@jamk.fi, minna.silvennoinen@jamk.fi

The increasing availability and complexity of data across professional domains require decision-makers to integrate diverse information in uncertain situations, highlighting the role of professional education in developing data-informed decision-making competence. While digital twins are widely used to model and optimise complex systems, their potential as pedagogically structured learning environments remains underexplored. This study presents the design of a layered learning infrastructure for data-informed decision-making, in which a digital twin functions as a component within a simulation-based learning environment. The study proposes a reconceptualisation of digital twins as components of learning processes. Drawing on multi-professional co-creation process in a smart agriculture context, the study suggests that pedagogical orchestration makes learners' reasoning visible by supporting multi-source data integration, trade-off evaluation, and engagement with uncertainty. The paper contributes a design-oriented framework and argues that the educational value of digital twins depends on pedagogical orchestration that renders decision-making processes explicit and open to reflection.

DOI
https://doi.org/10.18690/um.fov.4.2026.5

ISBN
978-961-299-152-4

Keywords:
data-informed decision-making,
digital twin,
simulation-based learning,
professional education,
smart agriculture

University of Maribor Press

1 Introduction

Across data-intensive professional sectors such as smart farming, decision-making increasingly requires the integration of heterogeneous data sources under conditions of uncertainty (Wolfert et al., 2017; Zhang et al., 2016). Smart farming refers to the use of digital technologies, data analytics, and automation to enhance agricultural productivity and sustainability (Wolfert et al., 2017; Klerkx et al., 2019). In this context, professionals are expected not only to access data, but also to interpret it, weigh alternatives, justify decisions, and reflect on consequences (Klerkx et al., 2019; Eastwood et al., 2019).

While the terms are often used interchangeably in the literature, some scholars distinguish data-driven decision-making as emphasising systematic data processing and analytic infrastructures, whereas data-informed decision-making highlights professional judgement, contextual interpretation, and ethical considerations (e.g., Brazauskienė, 2025). As decision-making in increasingly complex data-rich environments requires specific forms of professional competence, evidence suggests that data literacy training initiatives, for example in education, prioritise data analysis and interpretation, while the explicit development of decision-making competence remains less systematically addressed (e.g. Sandoval-Rios et al., 2025). Developing these competences has become a central challenge spanning individuals, organisations, and the education systems responsible for preparing and upskilling professionals (e.g. Sandoval-Rios et al., 2025; Coners et al., 2025; Brazauskienė, 2025).

Accordingly, this study adopts a data-informed decision competence (DIDC) perspective framing decision-making as a human, reflective competence rather than a purely technological process. Through intensified digitalization, AI-driven analytics become increasingly embedded in professional systems, research has raised concerns about over-reliance on AI models (Lai et al., 2021). This development highlights the need to preserve and strengthen human decision-making competencies (Comez et al., 2025).

Across sectors such as healthcare, engineering, energy, and smart agriculture, decision-making increasingly relies on multiple data sources. In smart agriculture, for example, decisions related to fertilisation strategies, harvesting timing, or

investment planning increasingly rely on multiple data sources, including production, economic, environmental, and contextual indicators (Wolfert et al., 2017). To support decision-making in such complex data-rich contexts, digital twins (DT) are increasingly used to model, monitor, and simulate complex systems, supporting analysis, optimisation, and strategic decision-making (Pylianidis et al., 2021; Jans-Senegacnik et al., 2023). However, modelling capability alone does not automatically translate into decision competence. Although digital twins are increasingly discussed in educational contexts (Bachmann et al., 2024), their conceptualisation as pedagogically structured learning infrastructures remains under-theorised. This paper conceptualises DTs as pedagogically embedded modelling environments. It presents the design and exploratory evaluation of layered learning infrastructure in a smart agriculture context.

Smart agriculture provides a prototypical example of a data-intensive sustainability domain characterised by multi-source data integration and systemic uncertainty (Wolfert et al., 2017). Decisions in this context must balance economic performance, environmental impact, and system resilience, underscoring the need for structured opportunities to practise decision-making without immediate real-world economic consequences.

The study makes three contributions

First, it reframes DTs not merely as immersive or adaptive learning technologies, but as components of pedagogically orchestrated learning infrastructures aimed at developing decision competence. **Second,** it proposes a layered learning infrastructure that integrates a DT modelling environment with structured decision–consequence–reflection cycles. **Third,** it offers exploratory empirical insights into how these cycles can scaffold decision practice in sustainability-oriented, data-intensive contexts.

Accordingly, the study is guided by the following research question:

RQ How can a digital twin component be pedagogically orchestrated within a simulation-based learning environment to structure and scaffold data-informed decision competence through iterative decision–consequence–reflection cycles?

By shifting the focus from technological optimisation to pedagogical orchestration, this study extends the discourse on DTs as learning environments that intentionally develop human competence in complex data-rich systems.

2 Theoretical Background

2.1 Data-informed decision competence in data-rich environments

Data-informed decision-making extends beyond access to data or the ability to interpret isolated indicators. While data-driven approaches emphasise systematic data utilisation and predictive analytics in complex agricultural systems, for example (Agbonika et al., 2025), data-informed decision-making integrates professional judgment, digital evidence, and ethical considerations (Brazauskienė et al., 2025). Professional judgment incorporates experiential insight, domain-specific expertise, and theoretical understanding (Schön, 1983), often emerging through distributed interaction rather than individual cognition. In complex socio-technical systems, decision-makers must compare alternatives, evaluate data reliability, and interpret trade-offs across temporal and systemic scales.

In this study, DIDC is conceptualised as a professional competence that enables decision-makers to (1) integrate heterogeneous data and knowledge sources, (2) critically evaluate the reliability and relevance of available data, (3) reason transparently under conditions of uncertainty, and (4) articulate and justify decisions through explicit reflection on potential consequences.

While maturity models such as Cech et al. (2018) conceptualise data capability primarily at the organisational level, we extend the developmental logic to the level of professional decision competence. In smart agriculture, data-driven approaches are often framed in technological and optimisation terms (e.g., McBratney et al., 2005; Sáiz-Rubio et al., 2020), with limited attention to how decision-making competence can be systematically developed in data-intensive environments. Consequently, there remains a conceptual gap between data-driven agriculture infrastructures and pedagogically structured competence development. Although data-driven infrastructures are increasingly sophisticated, explicit pedagogical models for human decision competence development within such environments seem under-theorised.

Such competence is particularly relevant in sustainability-oriented and data-rich decision contexts, such as smart farming, where economic, environmental, and systemic considerations interact and often conflict (Klerkx, Jakku & Labarthe 2019). In these contexts, integrated data does not eliminate tensions between alternative priorities; rather, decision-makers must interpret information, compare projected consequences, and justify trade-offs across temporal and systemic scales. Because DIDC involves interpretative and distributed reasoning, its development requires learning environments that make reasoning processes visible and open to critical examination.

2.2 Simulation pedagogy and decision–consequence–reflection cycles

Simply presenting integrated datasets or modelling system behaviour does not automatically support competence-oriented learning. Learning environments must provide structured opportunities for articulating decisions, comparing alternatives, and reflecting on consequences.

Simulation-based learning (SBL) has been widely used in professional education to create structured environments where learners can experiment with decisions and examine outcomes in a controlled setting (Lateef, 2010). Meta-analysis across 145 empirical studies demonstrates a large positive overall effect on the development of complex skills in higher education, particularly when SBL is combined with appropriate scaffolding (Chernkova et al., 2020). From an experiential and reflective learning perspective (Kolb, 1984; Schön, 1983), SBL may provide a structured framework within which processes associated with DIDC can be enacted, examined, and discussed. Evidence indicates that SBL is particularly effective when supported by appropriate scaffolding and instructional structuring (Chernikova et al., 2025).

Through structured engagement with integrated data and comparison of alternative interpretations, learners can practise multi-source reasoning, trade-off evaluation, and uncertainty awareness. Structured debriefing and guided reflection provide key mechanisms for examining assumptions and reconsidering decisions (e.g., Dieckmann, 2009; Cheng et al., 2014). Iterative decision–consequence–reflection cycles are particularly relevant because they make reasoning processes visible and open to examination, thereby supporting explicit justification of data-informed choices.

From an experiential learning perspective (Kolb, 1984), modelling environments, such as DTs, can provide a structured environment for concrete experimentation. However, the possibility to run DT-enabled simulations does not in itself constitute simulation pedagogy. By enabling learners to test decisions, observe simulated consequences, and revise their assumptions, DT environments can function as platforms for iterative transitions between experience, reflection, conceptualisation, and renewed action, positioning digital twins not only as technological tools but as pedagogically structured learning environments.

2.3 Digital twins as pedagogically orchestrated modelling environments

DTs are commonly conceptualised as dynamic digital representations of physical systems that integrate heterogeneous data sources to enable monitoring, simulation, and predictive analysis (Tao & Zhang, 2017). In agricultural contexts, digital twin approaches combine production, environmental, machinery, and economic data into system-level models that support scenario-based exploration of alternative decisions (Verdouw et al., 2021; Pylianidis, C., Osinga, S., & Athanasiadis, 2021; Peladarinos et al., 2023). More broadly, digital twin technologies are associated with data-intensive system monitoring, predictive analysis, and scenario-based modelling in contexts such as autonomous farming and smart agricultural technologies (Papadopoulos et al., 2025).

Recent research has explored the integration of DTs with virtual learning environments to enhance experiential and simulation-based learning, particularly in manufacturing education (e.g., Karanam & Hartman, 2025). Such work demonstrates the potential of DT-based immersive and adaptive learning environments. Nevertheless, comparatively less attention has been given to how DT environments can be pedagogically structured to support the development of human decision competence. From an educational perspective, the central question is not how accurately a DT represents a system, but how it can be embedded within pedagogical orchestration that structures decision-making processes and reflection. For the purposes of this study, DT is thus understood as a data-integrating modelling environment that enables scenario-based simulation. This perspective positions DTs as pedagogically orchestrated environments that can support the development of data-informed decision competence.

3 Methodological Approach

This study adopts a design-oriented exploratory approach focusing on the development and evaluation of a layered learning infrastructure. The design and evaluation were conducted across iterative cycles of design, implementation, reflection, and refinement. It examines how DT-based modelling environments can be intentionally integrated with structured pedagogical processes to scaffold DIDC within simulation-based learning environment.

The development followed a multi-professional co-creation process involving domain experts, vocational educators, pedagogical specialists, technical developers, and potential end-users. Workshops were organised to identify stakeholder expectations, perceived benefits, needs, and key decision-making challenges across different user groups (Aksovaara, et al. 2025). During this process, for example, learner personas were collaboratively developed to articulate anticipated user needs and the practical value the environment should provide. The workshops also included facilitated discussions to inform explicit competence objectives tailored to the learner personas, emphasising multi-source data integration, transparent justification of decisions, evaluation of uncertainty, and reflective examination of projected consequences across economic, environmental, and systemic dimensions.

In parallel with the workshops, a pedagogical orchestration layer was implemented within a learning management system (LMS) as part of the intervention design to structure iterative decision–consequence–reflection cycles. Core design outputs included (1) competence objectives for DIDC, (2) scenario-based decision tasks and related prompts, and (3) a DT simulation interface for testing solutions and exploring consequences. Based on these stakeholder-driven objectives, relevant data sources were identified and integrated into the DT modelling environment. Scenario-based decision tasks were constructed to represent realistic sub-decisions and their projected consequences.

Pilot session setup and participants

The first pilot was implemented as a six-hour group-based training session in a dual-platform learning environment consisting of a LMS and the DT simulation interface. The pilot was implemented in a hybrid format, combining an online webinar with

on-site group-based decision-making activities on campus. Participants were farmers (N = 5), recruited through the project network and participating on a voluntary basis. Their ages ranged from 30 to 65 years; two had over 20 years of experience in agriculture, while the others had shorter professional experience. Participants demonstrated moderate but uneven experience with digital tools: most used basic farm management and financial software (4/5), whereas fewer reported using advanced data-driven technologies such as sensors, automation, or telemetry (1–2/5).

The session comprised (1) an introduction to the DT simulation interface, which provided visualisations of key variables such as yield, cost, and environmental indicators, enabling comparison of alternative decision scenarios, (2) two to three decision rounds where participants selected and justified sub-decisions (e.g., fertilisation strategy and timing) using multiple data sources, (3) simulation of consequences in the DT, and (4) structured debriefing and reflection. The pedagogical orchestration layer in the LMS provided structured written prompts for decision justification and post-round reflection and served as a repository for participant artefacts.

Evaluation of the pilot session

Evaluation data included written participant reflections, collaboratively produced decision justifications, and survey responses. The analysis focused on examining how participants justified decisions, engaged with data, and reflected on consequences. The material was analysed using an iterative, theory-informed interpretative approach focusing on how participants articulated decisions, justified trade-offs, and engaged with uncertainty across the decision cycles. The analysis focused on recurring patterns, tensions, and illustrative examples in how participants used data, justified decisions, and engaged with uncertainty. The findings informed refinements to the pedagogical orchestration and supported examination of how decision-making processes became visible in practice. Insights were synthesised into relationships between pedagogical design, learner activity, and decision-process visibility, forming the basis of the conceptual model presented in the following section. Given the exploratory, design-oriented pilot, the small sample size was deemed appropriate for generating initial process-level insights.

4 Results

4.1. Layered learning infrastructure

The infrastructure developed in this study comprises two analytically distinct yet functionally interconnected layers (Figure 1): (1) a DT modelling layer and (2) a pedagogical orchestration layer.

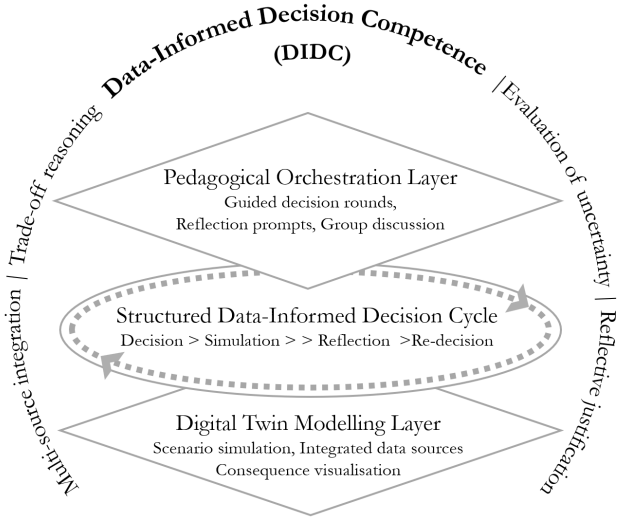

Figure 1: A layered learning infrastructure for data-informed decision-making with a digital twin component

The DT modelling layer functions as a virtual representation of the target system. It integrates production-related, economic, environmental, weather, and contextual data into a unified and temporally structured dataset. This enables scenario-based simulation of alternative decisions and examination of their projected consequences. By visualising outcomes across multiple criteria and time horizons, the DT renders systemic interdependencies and trade-offs analytically observable. The modelling environment itself does not prescribe how data should be interpreted or how decisions should be justified. Rather than a static analytics interface, the DT functions as an epistemic experimentation environment that externalises consequences and makes system dynamics visible

The pedagogical orchestration layer structures how the modelling environment is used within SBL. Implemented within a learning management system, it organises learning into iterative decision–consequence–reflection cycles. Within these guided rounds, learners articulate and justify sub-decisions based on multi-source data, implement and test alternatives within the DT environment, and engage in structured reflection. Reflection prompts and facilitated discussion make reasoning processes explicit, highlighting assumptions, uncertainty, data reliability, and trade-offs. A concluding reflection phase supports synthesis across decision rounds and connects simulated outcomes to broader sustainability-oriented system dynamics. Overall, the architecture creates structured opportunities for collaborative reasoning and examination of trade-offs in data-intensive contexts.

The interaction between the modelling and pedagogical layers forms an iterative learning cycle. Decisions are informed by integrated data, consequences are explored through the DT simulation interface, and reasoning is critically examined before proceeding to subsequent rounds. In relation to the RQ, the layered learning infrastructure illustrates how a DT component can be pedagogically orchestrated within a simulation-based learning environment, where the orchestration layer structures iterative decision–consequence–reflection processes.

4.2. Exploratory observations from the first pilot test

During the first six-hour pilot training session implemented as group-based learning, participants reported positively on the possibility to test alternative decisions and observe simulated consequences. The qualitative responses suggest engagement in processes associated with the components of DIDC.

Participants described how DT modelling outputs made system dynamics visible (e.g., "*It was interesting to see what kinds of results can be predicted through calculations – everything affects everything*"; "*The modelling made the relationships clearly visible*"). Decision-making drew on multiple heterogeneous data sources, including weather, soil sensor data, and historical data. Reflections such as *"We gained insights into variations in field growing conditions based on soil sensor data"* illustrate engagement in data-based reasoning.

The DT was explicitly described as a learning resource: *"I learned to use simulation comparison – it was a useful tool*", "*The simulation supported my own thinking*", and "*Through the simulations, we were able to compare alternatives.*" Participants also reflected critically on modelling outputs and data reliability (e.g., "*Historical data is not a law*," "*The computer can also be wrong*"), suggesting awareness of uncertainty and model limitations.

Trade-off reasoning between yield, costs, and environmental conditions was frequently articulated: "A *surprisingly high price is needed to reach break-even in grain farming,*" "*Even if everything is done well, market prices can still make the outcome unprofitable*," and "*Fertilisation does not always increase yield.*" Observations such as "*Inputs affect the final outcome*" and "*An optimum must be sought between yield and costs*" indicated comparative evaluation of alternatives. Reflective comments "*One must dare to make unconventional decisions,*" "*It is worth thinking more broadly when selecting seed varieties*" further pointed toward explicit articulation and justification of decisions. Peer interaction played a central role in the learning process. Participants emphasised the importance of discussion and shared reasoning "*We made decisions together as a team,*" "*The team's support was important*", indicating that decision-making was enacted collaboratively despite the strong presence of technological tools.

These observations provide empirical indications supporting the proposed design principles rather than generalisable findings. In relation to the research question, the findings suggest that pedagogically orchestrated decision–consequence–reflection cycles, with the digital twin functioning as a modelling and experimentation resource, can create conditions for practising processes associated with data-informed decision competence. In particular, the layered learning infrastructure appears to make multi-source reasoning, trade-off evaluation, uncertainty awareness, and explicit justification of decisions more visible within a simulation-based learning process. In the present pilot, these processes were enacted collaboratively, which may have further supported the articulation and externalisation of reasoning.

5 Discussion and conclusions

This study conceptualised DT as a component of pedagogically structured learning environments that make the consequences of decisions visible over time and enable learners to explore alternative options. The proposed layered learning infrastructure integrates a DT simulation interface with pedagogically orchestrated decision–

consequence–reflection cycles. This integration enables learners to test decisions, observe simulated consequences, and reflect on their reasoning in a structured manner, making decision-making processes explicit and open to examination. A key contribution of this study is the distinction between the modelling layer and the orchestration layer, which clarifies how competence-oriented learning depends on pedagogical structuring in addition to DT simulation interface capability. The exploratory examination of participant materials indicates that the pedagogically orchestrated decision–consequence–reflection cycles created conditions in which learners articulated and examined their reasoning in relation to multi-source data and projected system outcomes. Initial observations suggest that sustainability-oriented contexts, such as smart farming, amplify the need for practicing DIDC through structured articulation of assumptions, comparison of alternatives, and guided reflection. When decisions involve trade-offs across economic, environmental, and temporal dimensions, integrated data does not automatically resolve tensions. Instead, learners must interpret, prioritise, and justify their reasoning. The layered infrastructure appears to make these processes more visible and open to examination.

This exploratory study suggests that digital transformation requires advanced modelling technologies alongside pedagogical solutions. Although developed in a smart agriculture context, the learning infrastructure may be transferable to other data-intensive domains to support data-informed decision competence. Coupling DTs with scalable pedagogical orchestration may support structured engagement with decision processes across diverse domains characterised by uncertainty.

Despite these contributions, this study has limitations. The exploratory pilot involved a small sample (N=5), and the findings are based on interpretative analysis rather than formal qualitative coding, and should therefore be understood as indicative rather than generalisable. Future research should examine how pedagogically orchestrated DT environments, potentially augmented with AI, support the development of data-informed decision-making competence in larger and more diverse contexts.

End notes

Use of Generative AI in the Writing Process. Generative AI was used solely for language editing. All research design, data analysis, modelling decisions, and interpretation represent the authors' independent intellectual contributions. The authors take full responsibility for the content of this manuscript.

References

Agbonika, D. A., Abah, E. O., Fidelis, E. S., Hannah, E. U., & Haruna, S. O. (2025). A systematic review of management science integration in farm decision tools: Advancing theory for agricultural problem-solving. International Journal of Multidisciplinary Research and Growth Evaluation, 6(4), 656-664.

Aksovaara, S., Haapala, H., & Silvennoinen, M. 2025. Exploring Reflective Data-Driven Decision-Making Through Simulation Pedagogy and Digital Environments. In Ravesteijn, P., & Jung, J. (2025, June 18). Research for Impact in Digital Innovation: Proceedings of the 3rd RIDI Workshop in conjunction with 38th Bled eConference Bled (SI), June 8, 2025. Reesarch for Impact in Digital Innovation (RIDI), Bled, Slovenia. https://doi.org/10.5281/zenodo.15691349

Bachmann, J. E. C., Silveira, I. F., & Martins, V. F. (2024). Digital twins for education: A literature review. Simpósio Brasileiro de Informática na Educação (SBIE), 722-736.

Brazauskienė, E. (2025). Educational Decision-Making in Digital Education: A Conceptual Review of Data-Driven, Data-Based, and Data-Informed Approaches. Revista de Pedagogie Digitala, 4(1), 63-72. https://doi.org/10.61071/RPD.2562.

Brown, T. (2009). Change by design: How design thinking creates new alternatives for business and society. Harper Business.

Cech, T. G., Spaulding, T. J., & Cazier, J. A. (2018). Data competence maturity: developing data-driven decision making. Journal of Research in Innovative Teaching & Learning, 11(2), 139-158. https://doi.org/10.1108/JRIT-03-2018-0007.

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D. A. (2014). Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. Medical Education, 48(7), 657-666. https://doi.org/10.1111/medu.12432.

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-based learning in higher education: A meta-analysis. Review of educational research, 90(4), 499-541.

Chernikova, O., Sommerhoff, D., Stadler, M., Holzberger, D., Nickl, M., Seidel, T., Kasneci, E., Küchemann, S., Kuhn, J., Fischer, F., & Heitzmann, N. (2025). Personalization through adaptivity or adaptability? A meta-analysis on simulation-based learning in higher education. *Educational Research Review, 46*, 100662. https://doi.org/10.1016/j.edurev.2024.100662.

Comez, C., Cho, S. M., Ke, S., Huang, C. M., & Unberath, M. (2025). Human-AI collaboration is not very collaborative yet: A taxonomy of interaction patterns in AI-assisted decision making from a systematic review. Frontiers in Computer Science, 6, 1521066. https://doi.org/10.3389/fcomp.2024.1521066.

Coble, K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. Applied Economic Perspectives and Policy, 40(1), 79–96. https://doi.org/10.1093/aepp/ppx056

Coners, A., Matthies, B., Vollenberg, C., & Koch, J. (2025). Data skills for everyone!(?)–An approach to assessing the integration of data literacy and data science competencies in higher education. Journal of Statistics and Data Science Education, 33(1), 90-115. https://doi.org/10.1080/26939169.2024.2334408.

Dieckmann, P., Molin Friis, S., Lippert, A., & Østergaard, D. (2009). The art and science of debriefing in simulation: Ideal and practice. Medical teacher, 31(7), e287-e294.

Eastwood, C., Ayre, M., Nettle, R., & Dela Rue, B. (2019). Making sense in the cloud: Farm advisory services in a smart farming future. NJAS – Wageningen Journal of Life Sciences, 90–91, 100298.

Fischer, F., Bauer, E., Seidel, T., Schmidmaier, R., Radkowitsch, A., Neuhaus, B. J., ... & Fischer, M. R. (2022). Representational scaffolding in digital simulations–learning professional practices in higher education. Information and Learning Sciences, 123(11-12), 645-665.

Jans-Senegacnik, M., Kooistra, L., & Athanasiadis, I. N. (2023). Digital twins in smart farming: State of the art and future directions. Agricultural Systems, 203, 103534.

Karanam, S. A. K., & Hartman, N. W. (2025). A systematic review of digital twin (DT) and virtual learning environments (VLE) for smart manufacturing education. Manufacturing letters, 44, 1597-1608.

Klerkx, L., Jakku, E., & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0. NJAS – Wageningen Journal of Life Sciences, 90–91, 100315.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.

Lai, V., Chen, C., Liao, Q. V., Smith-Renner, A., & Tan, C. (2021). Towards a science of human-ai decision making: a survey of empirical studies. arXiv preprint arXiv:2112.11471.

Lateef, F. (2010). Simulation-based learning: Just like the real thing. *Journal of Emergencies, Trauma and Shock, 3*(4), 348. https://doi.org/10.4103/0974-2700.70743.

McBratney, A., Whelan, B., Ancev, T., & Bouma, J. (2005). Future directions of precision agriculture. Precision Agriculture, 6(1), 7–23. https://doi.org/10.1007/s11119-005-0681-8.

Nowak, B. (2021). Precision agriculture: Where do we stand? A review of the adoption of precision agriculture technologies on field crop farms in developed countries. Agricultural *Research, 10*, 515–522. https://doi.org/10.1007/s40003-021-00539-x.

Papadopoulos, G., Papantonatou, M., Uyar, H., Nychas, K., Psiroukis, V., Kasimati, A., Nieuwenhuizen, A., Van Evert, F., & Fountas, S. (2025). Stakeholders' perspective on smart farming robotic solutions. Smart Agricultural Technology, 11, 100916. https://doi.org/10.1016/j.atech.2025.100916.

Peladarinos, N., Piromalis, D., Cheimaras, V., Tserepas, E., Munteanu, R. A., & Papageorgas, P. (2023). Enhancing smart agriculture by implementing digital twins: A comprehensive review. Sensors, 23(16), 7128. https://doi.org/10.3390/s23167128.

Pylianidis, C., Osinga, S., & Athanasiadis, I. N. (2021). Introducing digital twins to agriculture. Computers and Electronics in Agriculture, 184, 105942.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Sandoval-Ríos, F., Gajardo-Poblete, C., & López-Núñez, J. A. (2025, March). Role of data literacy training for decision-making in teaching practice: A systematic review. In Frontiers in Education (Vol. 10, p. 1485821). Frontiers Media SA. https://doi.org/10.3389/feduc.2025.1485821.

Saiz-Rubio, V., & Rovira-Más, F. (2020). From smart farming towards agriculture 5.0: A review on crop data management. Agronomy, 10(2), 207.

Silvennoinen, M. & Aksovaara, S. (2025). Simulaatiopedagogiikka ammattiosaamista tukemassa. Jamk Arena Pro. https://urn.fi/urn:nbn:fi:jamk-issn-2984-0783-179.

Verdouw, C., Tekinerdogan, B., Beulens, A., & Wolfert, S. (2021). Digital twins in smart farming. Agricultural Systems, 189, 103046. https://doi.org/10.1016/j.agsy.2020.103046.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming – A review. Agricultural Systems, 153, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023.

Zhang, Y., Wang, L., & Wang, Y. (2016). Decision-making for precision agriculture using data analytics: A review. Precision Agriculture, 17(3), 349–372.

Pedagogical Data Production as the Generative Foundation of Meaningful Learning Analytics

Satu Aksovaara, Minna Silvennoinen
Jamk University of Applied Sciences, Jyväskylä, Finland
satu.aksovaara@jamk.fi, minna.silvennoinen@jamk.fi

Learner-centred and distributed higher education increases the need for information that makes learning visible beyond system-generated traces. While learning analytics has advanced methods for analysing learning processes, less attention has been paid to how learning design generates data on which such analytics depend. Drawing on a multi-year, design-oriented and practice-based research process, this process paper introduces a model that conceptualises pedagogical data production as a configuration of interconnected learning design dimensions. It specifies how the learning-related data, defined by its interpretive function, is produced through the alignment of learning design elements, which structure how such data is generated, becomes visible, and is interpreted, making learners' engagement, reasoning, and competence development visible. The proposed model positions learning analytics as a practice emerging from pedagogically structured learning processes and defines learning design as the generative condition of pedagogically meaningful learning-related data, through which learner agency and pedagogical action can be supported.

DOI
https://doi.org/10.18690/um.fov.4.2026.6

ISBN
978-961-299-152-4

Keywords:
learning analytics, pedagogical data production, learning-related data, learning design, learner agency

University of Maribor Press

1 Introduction

In learner-centred distributed higher education (HE), making learning processes visible beyond system-generated traces has become increasingly important. Distributed HE refers to learning across multiple settings (e.g., online, onsite, hybrid, and work-related), where learning involves group work and personalised pathways, and both learners and teachers require timely, context-sensitive information as it unfolds.

While learning analytics (LA) enables large-scale data collection, the pedagogical production of learning-related data remains insufficiently conceptualised within learning design (LD). In this paper, LD is understood as the pedagogical structuring of learning activities, tasks, and assessment. LA has been defined as a discipline for analysing educational data to support learner progress and pedagogical decision-making (Siemens, 2013; Gašević et al., 2015). In practice, LA often relies on indicators such as submissions, log-ins, and time-on-task. These traces enable monitoring at scale but provide only a partial representation of learning processes. This is particularly evident when considering cognitive, motivational, and self-regulatory processes, which are central to learning (Panadero, 2017; Zimmerman, 2002; Buckingham Shum et al., 2019). As a result, what becomes visible is shaped by what is easily measurable, and important dimensions of learning remain less visible (Motz et al., 2023). This indicates a key limitation in current LA approaches. The generation of pedagogically meaningful learning-related data within LD remains insufficiently conceptualised in prior research (Mangaroska & Giannakos, 2019).

The limitations of trace-based approaches become particularly evident in HE contexts characterised by collaborative, project-based, and individually paced learning. While system traces may indicate activity, such as assignment submission, they primarily reflect group-level outcomes and provide limited insight into individual contributions, progression, or understanding. A central challenge is that learning-related trace data does not adequately capture how learning unfolds in social and contextual processes, nor does it represent learners' experiences and sense-making, which require pedagogically designed opportunities for articulation and reflection.

Structural and pedagogical transformations in HE increase the demand for timely, pedagogically meaningful information about learning processes (Ifenthaler & Yau, 2020; Teräs et al., 2020). Growing learner diversity, larger groups, and limited resources reshape how learning takes place. At the same time, in HE, learning-related data is often fragmented across platforms and interactions. The key issue is not the amount of data, but whether it is pedagogically meaningful, aligned with learning goals, and structured to inform action. This points to a central principle: LA must be grounded in LD, where learning activities structure what becomes visible as data.

In this study, meaningful LA refers to the intentional generation of learning-related data that makes learners' understanding, reasoning, and competence development visible in ways that inform pedagogical action and support learner agency. To address this, the study introduces a LA–integrated model of pedagogical data production. The model is articulated through the following research questions:

> RQ1: What constitutes pedagogically meaningful learning-related data within distributed higher education contexts?
> RQ2: How can such data be intentionally produced through LA in ways that support learner agency and pedagogical action?

2 Conceptual Background

While recent scholarship has emphasised aligning LD with educational theory and evidence (Neelen & Kirschner, 2020), the pedagogical generation of meaningful learning-related data remains under-conceptualised. This section develops a conceptual foundation for pedagogical data production through learner agency, LD, and ethical conditions and addresses the current gap in understanding how this data is intentionally generated.

2.1 Pedagogically meaningful learning-related data and learner agency

LA research in higher education has frequently focused on instructional decision-making and institutional monitoring, particularly performance prediction and administrative uses (Leitner et al., 2017). In these approaches, learners often appear as data subjects rather than active participants in interpreting learning-related data (Buckingham Shum et al., 2019; West et al., 2019). While such applications support

large-scale coordination, they do not clarify what constitutes pedagogically meaningful learning-related data within learning processes at the course level.

From a learner-centred perspective, data is defined not by its technical availability but by its relevance to learners' learning processes. Such data becomes pedagogically meaningful for learners when it supports their understanding, regulation, and informed decision-making. Access to learning-related data enables learners to plan, monitor, and adjust their learning, thereby supporting their agency. In this sense, agency refers to learners' capacity to make informed decisions about their learning.

For teachers, pedagogically meaningful learning-related data enhances pedagogical awareness and enables adaptive guidance and timely support within ongoing learning processes (Banihashem et al., 2022, Teräs et al., 2020). In collaborative and distributed contexts, this applies to both individual and shared learning processes. Such data supports teachers in recognising learners' emerging understandings and areas of difficulty, thereby informing pedagogical action while sustaining learner agency.

2.2 LD as a mechanism for meaningful data production

While substantial research has focused on analytic techniques and visualisations, fewer studies have theorised the pedagogical generation of learning-related data within everyday teaching practice. A substantial proportion of studies do not directly measure learning outcomes or implement interventions in learning environments (Motz et al., 2023), and evidence of sustained pedagogical impact remains limited (Viberg et al., 2018). These observations have prompted calls to connect LA more explicitly with learning theory and learners' regulatory processes (Ifenthaler & Yau, 2020; Wong et al., 2019). Access to dashboards or tracking metrics alone does not ensure learner agency, particularly when representations focus primarily on behavioural traces detached from pedagogical context (Bodily & Verbert, 2017).

If pedagogically meaningful learning-related data is defined by its relevance to learning processes, the remaining question is how such data is produced and made visible in practice. Through LD, teachers actually influence which aspects of learning processes may become articulated, documented, and rendered interpretable as learning-related information (Mangaroska & Giannakos, 2019). This perspective

differs from approaches that treat LD primarily as a context for interpreting analytic outputs; instead, LD is understood here as the condition under which learning-related data is generated. What becomes visible as data is therefore shaped by how learning activities are configured and how opportunities for articulation, reflection, and data use are embedded within the learning process.

In competence-based HE, LD commonly follows an alignment logic connecting learning outcomes, assessment, and activities (Biggs, 1996). Beyond this alignment perspective, LD can be understood as a configuration of interrelated pedagogical elements forming a structured and coherent system (van den Akker, 2004) which shapes both learning processes and what becomes visible. From this perspective, LA is not external to LD but embedded within it, as part of the pedagogical system that shapes the production of learning-related data.

While prior research has explored connections between LD and LA, the focus has largely been on analysing learning processes or supporting instructional and design-related decision-making (e.g., Gašević et al., 2015; Er et al., 2019; Lim et al., 2023; Sajja et al., 2025). The perspective proposed here shifts attention to how pedagogical design shapes the data that is subsequently analysed. The central question thus moves from analysing available traces to configuring the pedagogical conditions under which meaningful learning-related data can emerge.

2.3 Responsible and Sustainable Pedagogical Data Production

Pedagogical data production is not value-neutral. LA ethics goes beyond legal compliance to require transparency, intentionality, and proportionality in data practices (Drachsler & Greller, 2016). If the generation of data is not grounded in pedagogically designed activities, or if its use is not transparent, attention may shift towards forms of monitoring that do not directly support learning and may limit learners' opportunities to act as active agents in interpreting and using such information. At the course level, teachers' pedagogical design decisions should determine what data is produced, for whom and how it becomes visible, and how it informs learning. In practice, however, this is not yet consistently recognised or implemented as part of the design phase.

Sustainable pedagogical data production requires that data collection, management, and storage are conducted in an ethical, transparent, and institutionally aligned manner (Drachsler & Greller, 2016). This entails alignment with institutional systems, clear data management practices, and well-defined responsibilities for data handling. At the course level, LA should therefore be grounded in consistent data practices. However, when separate tools outside the Learning Management System (LMS) are used to address emerging information needs, data may become fragmented, increasing management complexity and shifting responsibility for data use—including access rights and security—to individual teachers (Prinsloo & Slade, 2017; Drachsler & Greller, 2016). This fragmentation may weaken the connection between learning activities and the data, for example when learning-related data is distributed across simulation environments or learning tools, making it difficult to use as part of a coherent learning process.

Together, these considerations highlight that pedagogically meaningful data derives its value, for both learners and teachers, from how it is generated through LD and the ethical, transparent, and institutionally aligned conditions under which it is produced, managed, and interpreted. These dimensions also align with prior research on LA and feedback (e.g., Banihashem et al., 2022), which highlights similar forms of learning-related data. Figure 1 visualises these shared forms of learning-related data and their distinct interpretive functions in supporting learner agency and pedagogical action.

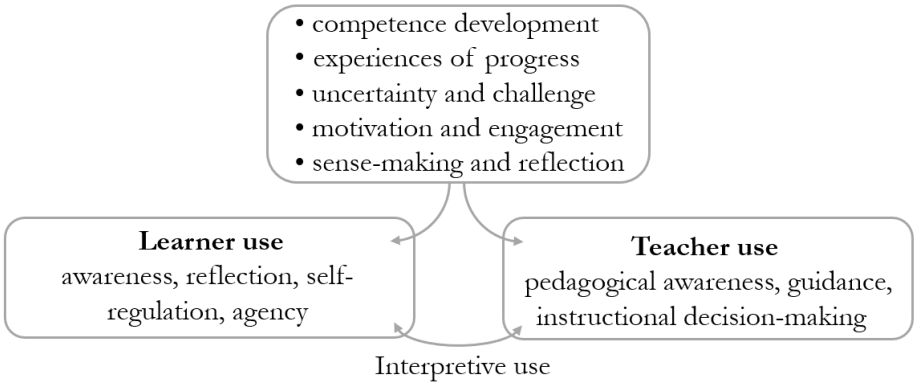

Figure 1: Shared forms of pedagogically meaningful learning-related data and their differentiated interpretation by learners and teachers

As illustrated in Figure 1, these forms of data are shared between learners and teachers but differ in how they are interpreted and used within the learning process. In response to RQ1, pedagogically meaningful learning-related data in distributed HE contexts consists of intentionally configured and pedagogically aligned information defined by its interpretive function rather than its technical origin. Such data makes learners' progress, challenges, motivation, and sense-making processes visible in ways that support learner agency and teacher decision-making under transparent and sustainable conditions.

3 Methodological Approach

This study draws on design-oriented and practice-based research (McKenney & Reeves, 2018; Sandoval, 2014), in which pedagogical interventions are iteratively developed, enacted, and analysed to identify transferable design principles. It presents a process-informed conceptual model developed through a multi-year research and development process (2018–2025) which also builds on five years of doctoral research. The aim is to abstract a conceptual model from sustained design-based work, rather than to report findings from a single empirical intervention.

The empirical development centred on two structurally distinct courses (2 and 10 ECTS), involving approximately 1,500 students annually and resulting in several thousand student participations across diverse cohorts and institutional contexts. The courses were selected to represent the organisational and pedagogical diversity. Together, they capture key variations typical of distributed HE, including differences in pedagogical scope, delivery modes (face-to-face, hybrid, and online), and learning approaches such as project-based, collaborative, small-group, work-integrated, and self-directed study.

Importantly, this variation was inherent to the courses themselves, as they were delivered to diverse learner cohorts and implemented by approximately 50 teachers annually. Participants included both national and international students from HE with selected extensions to other educational settings such as vocational upper secondary education, resulting in both structural and demographic diversity. This enabled examination of whether the emerging design mechanisms remained stable across differing pedagogical and institutional conditions.

Course development proceeded through multiple iterative cycles conducted annually, involving pedagogical co-design, implementation, reflection, and revision of learning tasks and assessment structures. The aim was to configure learning activities to intentionally generate learner-produced, competence-oriented data embedded within the LD.

The generation of learning-related data was pedagogically structured through embedded reflection and self-assessment processes positioned at key points in the learning process. These included diagnostic self-assessment at the beginning of the course to identify learners' initial state, ongoing reflective checkpoints to support monitoring and sense-making, and summative self-assessment at the end to support evaluation of competence development and learning experience. Reflection tasks included prompts related to learners' emotions, perceived progress, activity, development of competence in relation to learning objectives, experiences of teamwork and collaboration, and feedback on teaching and learning materials, capturing both performance-related and experiential aspects of the learning process.

Pedagogical design, course enactment, and research activities were closely intertwined with the configuration of LA within the institutional LMS, which functioned both as the pedagogical environment and as a documentary infrastructure for the development process. Documentation included archived course structures, design artefacts, analytics configurations, learner-produced materials, research notes, and documented teacher–researcher discussions.

This work was conducted in close collaboration between teachers and researchers, with pedagogical decisions and analytic interpretations jointly examined and refined in regularly organised meetings. Students contributed by producing learning-related data through pedagogically designed reflection points and delayed course feedback.

Across iterative cycles, course implementations were compared using a qualitative approach to identify relationships between LD decisions, learner articulation and feedback, data visibility, and pedagogical action. Teacher–researcher meeting discussions were documented to ensure traceability. In one course iteration, the focus was on understanding team functioning and how collaboration unfolded, using student reflections on roles, challenges, and decision-making as the primary data source.

These data informed subsequent design decisions and revealed recurring patterns in how pedagogical design shaped the visibility of learning-related data. Over successive iterations, these patterns were progressively abstracted through systematic comparison into relationships linking design decisions to the production of interpretable evidence, forming the basis of the conceptual model presented in the following section.

4 LD–Integrated Data Production

The model conceptualises pedagogical data production as a **configuration of interconnected LD dimensions** as outlined in Section 2. Figure 2 illustrates how these interconnected dimensions enable the intentional production of pedagogically meaningful data to support learner agency and pedagogical action, particularly in distributed HE contexts.

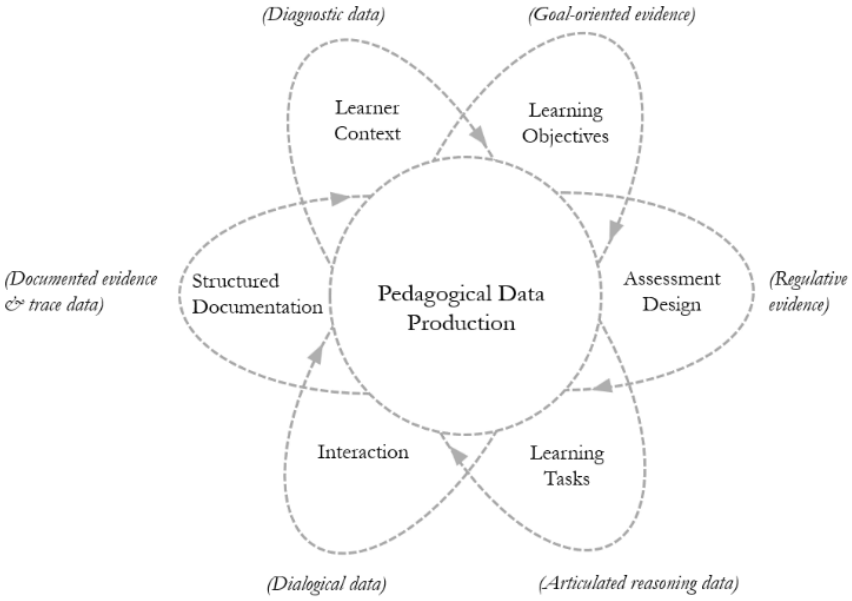

Figure 2: Pedagogical data production model where each dimension structures and contributes to the generation of learning-related data within ongoing learning process

Learner context shapes what becomes visible about learners' starting points, experiences, and constraints. Through structured onboarding tasks and reflection activities embedded within the course design, this configuration generates *diagnostic data* related to prior knowledge, expectations, and emerging uncertainties.

Learning objectives define what counts as learning and thereby shape the nature of the evidence to be produced as *goal-oriented evidence*. Within the course design, they guide the configuration of assessment events and learning tasks that prompt learners to articulate reasoning, demonstrate applied understanding, justify decisions, and reflect in relation to stated objectives.

Assessment design structures when and how *regulative learning-related evidence* is generated by integrating diagnostic, formative, and summative elements into an ongoing assessment process. These assessment moments across the learning process organise the production of data, such as self- and peer-feedback, and evidence of competence development. Assessment design thus organises the timing, form, and use of learning-related data.

Learning task configuration determines what learners are required to articulate, shaping the generation of *articulated reasoning data*. Tasks that involve comparison, explanation, justification, or reflection prompt learners to make their thinking explicit. Such tasks extend beyond behavioural indicators and produce traces of reasoning through which learners produce data about their thinking and experiences. Positioned at specific moments in the learning process, learning tasks also function as points at which learners are activated to produce data.

Interaction structures determine how learners engage in collaborative phases of the learning processes, shaping the generation of *dialogical data*. Discussions, peer and teacher interaction, and guided interaction prompt learners to articulate perspectives and negotiate meaning, producing traces of participation, collaboration, and shared understanding. Through interaction structures learners produce data about social and dialogical dimensions of learning.

Documentation structures, supported by the LMS, provide the infrastructure for recording and making learning-related data available for use within the course. An LMS enables the documentation of pedagogically structured artefacts. This may

include data originating from different learning contexts (e.g. external tools), which is brought into the LMS through learning activities or tasks. Centralising data within institutional LMS environments supports consistent, transparent, and secure handling of it. While the LMS organises this *documented evidence and trace data*, it does not determine its pedagogical meaning.

The model conceptualises LD as a data-generating mechanism in which pedagogical elements such as objectives, tasks, assessment, interaction, and documentation structures function as conditions for the generation, visibility, and interpretability of learning-related data within ongoing learning processes. Together, these elements form a coherent pedagogical configuration in which data is intentionally produced, made visible, and can be used as part of the learning process. When data production is embedded in ongoing learning processes, learners become active producers of such data, while both learners and teachers act as its interpreters and users. In this way, LA can support learner-centred understanding and foster learner agency (e.g., Hooshyar et al., 2023), while also providing a basis for data-informed pedagogical decision-making.

5 Implications for LA as pedagogical practice

Experiential and competence-related dimensions of learning become analytically visible through pedagogical structures that invite learners to articulate, justify, document, and reflect on their activity. This repositions learners not only as recipients of analytics but as active contributors to the production of interpretable evidence within ongoing learning processes (see also e.g. Hooshyar et al., 2023; Viberg et al. 2018).

Rather than treating LA as a technological capacity emerging from trace infrastructures (e.g., Gašević et al., 2015; Drachsler & Greller, 2016), the model conceptualises analytics as a property of pedagogical structure. What becomes analytically visible is shaped by design decisions guided by anticipated informational needs, rather than by system capability alone. Pedagogical dimensions generate distinct forms of meaningful learning-related data-diagnostic, goal-oriented, regulative, dialogical, and reasoning-based-thereby addressing RQ1, while specifying how such data can be intentionally produced through learning design (RQ2). In distributed HE contexts, the key challenge is how learning-related data is

pedagogically produced and made meaningful across learning processes that span multiple environments. This shifts the focus of learning analytics from system-generated traces to pedagogically structured learning processes, enabling the model to extend beyond digitally mediated environments to diverse learning contexts. Repositioning LA as a pedagogical property does not remove institutional constraints, but clarifies that meaningful analytics depends on pedagogical conditions under which learning becomes visible, interpretable, and actionable.

In this perspective, LA primarily supports learning during the process by informing self-regulation and pedagogical guidance, rather than retrospective monitoring. By defining learning-related data structurally, the model extends beyond system-specific implementations and can operate across diverse learning contexts. This reconceptualisation aligns LA with a data-informed perspective, where pedagogically configured evidence complements professional judgment and positions teachers as designers of the conditions under which learning-related data can emerge.

As AI-supported and data-informed education expands, its value depends on the availability of pedagogically meaningful learning-related data (e.g., Yan et al., 2024; Sajja et al., 2025). This study emphasises that this value ultimately depends on how such data is pedagogically produced and made meaningful within learning processes.

6 Conclusion

This study demonstrates that learning-related data becomes interpretable through pedagogical configuration. LA is best understood as a pedagogically embedded practice that emerges from intentionally designed learning processes producing interpretable evidence. By conceptualising LD as the generative condition of learning-related data, the study shows how such data can be intentionally produced within ongoing learning processes to support learner agency and pedagogical action. In this perspective, learners are not only recipients of analytics but active contributors to the production of learning-related evidence.

End notes

Use of Generative AI in the Writing Process. Generative AI was used solely for language editing. All research and interpretation are the authors' own, and they take full responsibility for the manuscript.

References

Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. A. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review, 37*, 100489. https://doi.org/10.1016/j.edurev.2022.100489

Biggs, J. (1996). Enhancing teaching through constructive alignment. Higher Education, 32(3), 347–364. https://doi.org/10.1007/BF00138871

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-Centred Learning Analytics. *Journal of Learning Analytics*, *6*(2), 1–9. https://doi.org/10.18608/jla.2019.62.1

Drachsler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue: A checklist for trusted learning analytics. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 89–98). ACM. https://doi.org/10.1145/2883851.2883893

Er, E., Gómez-Sánchez, E., Dimitriadis, Y., Bote-Lorenzo, M. L., Asensio-Pérez, J. I., & Alvarez-Alvarez, S. (2019). Aligning learning design and learning analytics through instructor involvement: A MOOC case study. *Interactive Learning Environments*, 27(5–6), 685–698. https://doi.org/10.1080/10494820.2019.1610455

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Hooshyar, D., Tammets, K., Ley, T., Aus, K., & Kollom, K. (2023). Learning analytics in supporting student agency: A systematic review. *Sustainability, 15*(18), 13662. https://doi.org/10.3390/su151813662

Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success: Reflections on current practice and future directions. Educational Technology Research and Development, 68(4), 1963–1988. https://doi.org/10.1007/s11423-020-09788-z

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education - A literature review. *Learning Analytics: Fundaments, Applications, and Trends* (pp. 1-23). Springer, Cham. https://doi.org/10.1007/978-3-319-52977-6_1

Lim, L. A., Atif, A., Heggart, K., & Sutton, N. (2023). In Search of Alignment between Learning Analytics and Learning Design: A Multiple Case Study in a Higher Education Institution. *Education Sciences*, 13(11), 1114. https://doi.org/10.3390/educsci13111114

Mangaroska, K., & Giannakos, M. (2019). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. IEEE Transactions on Learning Technologies, 12(4), 516–534. https://doi.org/10.1109/TLT.2018.2868673

McKenney, S., & Reeves, T. (2018). Conducting Educational Design Research (2nd ed.). Routledge. https://doi.org/10.4324/9781315105642

Motz, B. A., Bergner, Y., Brooks, C. A., Gladden, A., Gray, G., Lang, C., Li, W., Marmolejo-Ramos, F., & Quick, J. D. (2023). A LAK of Direction: Misalignment Between the Goals of Learning Analytics and its Research Scholarship. *Journal of Learning Analytics*, *10*(2), 1–13. https://doi.org/10.18608/jla.2023.7913

Neelen, M., & Kirschner, P. A. (2020). *Evidence-informed learning design: Creating training to improve performance*. Kogan Page.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, 422. https://doi.org/10.3389/fpsyg.2017.00422

Prinsloo, P., & Slade, S. (2017). *An elephant in the learning analytics room: The obligation to act.* In J. A. Larusson & B. White (Eds.), Learning analytics: From research to practice (pp. 269–282). Springer.

Sajja, R., Sermet, Y., Cwiertny, D., & Demir, I. (2025). Integrating AI and learning analytics for data-driven pedagogical decisions and personalized interventions in education. *Technology, Knowledge and Learning*, 1–31. https://doi.org/10.1007/s10758-025-09897-9

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*(1), 18–36. https://doi.org/10.1080/10508406.2013.778204

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Teräs, M., Silvola, A., Teräs, H., Hartikainen, S., Aksovaara, S., Hietaniemi, R., & Muukkonen, H. (2020). Learning analytics for students: Synthesis of two user needs studies in Finnish higher education. In *EdMedia + Innovate Learning* (pp. 455–463). Association for the Advancement of Computing in Education (AACE). https://urn.fi/URN:NBN:fi-fe2022021519072

van den Akker, J. (2004). Curriculum Perspectives: An Introduction. In *Curriculum Landscapes and Trends*. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-1205-7_1

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in Human Behavior, 89, 98–110. https://doi.org/10.1016/j.chb.2018.07.027

West, D., Luzeckyj, A., Toohey, D., Vanderlelie, J., & Searle, B. (2019). Do academics and university administrators really know better? The ethics of positioning student perspectives in learning analytics. *Australasian Journal of Educational Technology, 35*(6), 27–41. https://doi.org/10.14742/ajet.4653

Wong, J., Baars, M., de Koning, B. B., van der Zee, T., Davis, D., Khalil, M., Houben, G.-J., & Paas, F. (2019). Educational theories and learning analytics: From data to knowledge. In *Utilizing Learning Analytics to Support Study Success* (pp. 3–25). Springer, Cham. https://doi.org/10.1007/978-3-319-64792-0_1

Yan, L., Martinez-Maldonado, R., & Gašević, D. (2024). Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle. In *Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 101–111). ACM.https://doi.org/10.1145/3636555.3636856

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

DOI
https://doi.org/10.18690/um.fov.4.2026.7

ISBN
978-961-299-152-4

LEVERAGING TRANSACTIONAL BUSINESS DATA TO PREDICT EMPLOYEE WORKLOAD SATISFACTION IN OPERATIONS: AN EMPIRICAL STUDY – PART 1

VALMIR BEKIRI,[1] STEFAN STÖCKLER,[1]
KATRIN OETTMEIER[2]

[1] Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland
valmir.bekiri@ost.ch, stefan.stoeckler@ost.ch
[2] University of Liechtenstein, Vaduz, Liechtenstein
katrin.oettmeier@uni.li

We examine whether transactional Enterprise Resource Planning (ERP) data can predict employee-perceived workload satisfaction in a metal-processing facility. Drawing on 127 working days of daily employee surveys across four logistics gates, we link each day's responses to operational records from four SAP tables: (COOIS) production confirmations, (LT060) external transport orders, (LT061) internal transport orders, and (MB51) material movements. Two gates register mean stress scores five to six times higher than a third. Internal transport activity correlates negatively with perceived problems at two gates, suggesting a buffering effect; at others, material volume rather than transport frequency seems to drive stress. Mid-week workload is consistently elevated. Our results confirm that ERP metrics can explain a bounded but meaningful portion of subjective workload variance, and that workplace-/ gate-level modeling substantially outperforms generalized approaches. These findings offer actionable insights for operations managers seeking to proactively monitor workplace social sustainability through data already available in existing information systems.

Keywords:
social sustainability, business information systems, employee satisfaction, continuous improvement, operations management

University of Maribor Press

1 Introduction

In the corporate context, sustainability refers to an organization's impact on the natural and social environment as well as its financial performance (Albers Mohrman and E. Lawler III 2014). Social sustainability, which is frequently framed as corporate social responsibility, thereby encompasses the people side of sustainability, such as human and labour rights, diversity, employee engagement, and supply chain responsibility (Missimer and Mesquita 2022; Nijhof et al. 2003). An important indicator for operationalizing social sustainability at the workplace is employee satisfaction, which is commonly understood as "a measure of how happy workers are with their job and working environment". Satisfied employees tend to be more committed and productive, while dissatisfaction can lead to low performance and absenteeism (Sageer 2012).

Although frequent employee surveys are beneficial for organizations - the frequency of feedback was found to be positively linked to both, performance and satisfaction (Mertens et al. 2021; Yaneva 2018) - systematic feedback collection remains uncommon in practice. In the DACH[1] region, 48% of companies conduct employee opinion surveys only every two years, and 17% annually (Maximilian Siessegger 2025). This sporadic approach risks overlooking short-term changes in employee satisfaction, hindering proactive social sustainability management.

Research identifies multiple determinants of employee satisfaction, including wages, leadership, promotion opportunities, work safety, and operational job characteristics such as working hours and task type and quantity (Myskova 2011; Sageer 2012). For day-to-day monitoring, short-term business process data is particularly relevant, as it offers insights into workload and tasks performed.

This paper addresses this gap in two ways. First, it extends a previously proposed idea for a Social Sustainability Insights Tool (Oettmeier et al. 2025) with a full empirical validation conducted at the logistic operations department of a manufacturing company. Second, it contributes to the emerging literature on people analytics in operations management by demonstrating how standard ERP transaction data can be used to build predictive models of employee workload

[1] D…Deutschland/Germany, A…Österreich/Austria, CH…Schweiz (DACH normaly includs also Liechtenstein)

perception together with seven hypothesis described in chapter 7. The central research questions are:

RQ1: *To what extent can subjective employee workload ratings be explained by objective ERP-derived operational metrics at the workplace (warehouse gate) level?*

RQ2: *Which operational factors correlate most strongly with perceived workload, and which warehouse gate-specific patterns provide actionable insights for continuous improvement?*

2 Related Work

2.1 Employee Satisfaction and Operational Workload

The literature broadly distinguishes between distal job satisfaction determinants - such as pay, career advancement, and organizational culture - and more proximal, task-related factors that fluctuate on shorter time horizons (Hackman and Oldham 1976; Myskova 2011). Demand-Control-Support (DCS) models (Karasek 1979) and the Job Demands-Resources (JD-R) framework (Bakker and Demerouti 2007) suggest that high task demands unaccompanied by adequate resources or autonomy are primary drivers of occupational stress and reduced satisfaction. In manufacturing and logistics contexts, empirical studies have linked objective workload metrics - such as order volumes, handling frequencies, and shift lengths - to subjective wellbeing outcomes (Lin and Mon 2025; Inegbedion et al. 2020).

Temporal patterns in workload have received comparatively less empirical attention. However, studies on production scheduling confirm that mid-week periods often accumulate higher order backlogs due to Monday start-up effects and end-of-week completion pressures (J. Buzacott and J. George Shanthikumar 1993). Such rhythmicity suggests that day-of-week may serve as a meaningful temporal control variable in workload models.

2.2 People Analytics and ERP-Driven Insights

People analytics - the collection and analysis of employee-related data to support evidence-based HR decision-making - has emerged as a significant research and practice domain (Opatha 2020). Recent work highlights the transformative potential

of integrating diverse data sources, including operational metrics, to enhance satisfaction and retention (Ajiga et al. 2024; Snigdha and Singh 2024). ERP systems have increasingly incorporated people analytics dashboards to analyse staff attrition, hiring, and labour costs, demonstrating correlations between operational data and workforce metrics (Megan O'Brien 2021).

Empirical evidence also supports correlations between ERP system usability and employee satisfaction: studies on ERP implementations have found that streamlined workflows reduce stress and enhance satisfaction, while usability issues increase cognitive load and frustration (Bradley and Lee 2007; Dachawar et al. 2025). Overall, this study establishes a theoretical rationale for using ERP transaction data as a proxy for operational demands, yet no study has linked production order volumes, transport quantities, and material movements directly to daily satisfaction survey data at the production gate level.

3 Methodology

3.1 Research Design

This study is positioned within the Design Science Research (DSR) methodology of Peffers et al. (2007). The overall research program targets the development and validation of a Social Sustainability Insights Tool for ERP-based workplace monitoring. The problem identification, objectives definition, and initial design and development phases - including the conceptual architecture and requirements elicited from HR managers and data protection experts - were reported in prior work (Oettmeier et al. 2025). The present paper addresses the subsequent demonstration and evaluation phases: it demonstrates the feasibility of instantiating the tool's core mechanism - daily aggregation of ERP transactions and their alignment with workload survey data - in a real manufacturing setting and evaluates the explanatory and predictive value of this mechanism using quantitative analyses. The findings are fed into the design cycle..

3.2 Research Site and Survey Instrument

The study was conducted at a manufacturing facility that operates four distinct production warehouse logistics gates (referred to as Gates 3, 4, 5, and 6). Each gate represents a dedicated operational area with specific workflows involving inbound and outbound logistics as well as material handling. Employees at each gate complete as a group (not individually) a brief anonymized daily workload survey as part of the existing shopfloor management routine. The survey captures three dimensions on a 0 - 4 ordinal scale, where 0 denotes no workload and 4 denotes very high workload:

Routine Tasks (a): Perceived burden from standard, planned activities.

Irregular Tasks (u): Perceived burden from unplanned, ad-hoc activities.

Problems (p): Perceived frequency/severity of operational problems encountered.

3.3 ERP Data Sources and Feature Engineering

Operational data was extracted from three SAP transactions covering the same facility. Table 1 provides an overview of all data sources employed.

Table 1: Overview of data sources used in the analysis

Source	Records	Content	Role in Analysis
workload_ratings.csv	127	Daily employee workload ratings per gate (a, u, p)	Target variables
COOIS.csv	29,986	Production order confirmations by gate & date	Feature source (all gates)
LT22_061.csv	39,084	Internal warehouse transport orders	Feature source (Gates 5, 6)
LT22_060.csv	6,501	External transport orders (inbound/outbound)	Feature source (Gate 4)
MB51.csv	4,945	Material movement documents (goods receipts/issues)	Feature source (Gates 3, 4)
workplace.csv	129	Workstation-to-gate mapping table	Mapping reference
bom-assignment.csv	2,545	Bill-of-materials component assignments	Complexity reference

Source: Own

A key design decision was the adoption of gate-specific feature engineering. Given the structural heterogeneity of ERP data availability across gates - for instance, internal transport records (LT061) are exclusively associated with Gates 5 and 6, while external logistics records (LT060) are almost entirely attributable to Gate 4 - pooling all gates in a single feature space would introduce systematic missingness and mask gate-level dynamics. Table 2 summarizes the feature availability profile per gate.

Table 2: ERP data availability profile by production gate

Gate	Production Orders (COOIS)	Internal Transport (LT061)	External Transport (LT060)	Material Movements (MB51)
Gate 3	998 records	None	None	109 records
Gate 4	3,196 records	None	6,286 records	4,836 records
Gate 5	8,502 records	6,719 records	None	None
Gate 6	5,377 records	13,141 records	None	None

Source: Own

For each gate, daily-level features were aggregated from the raw transaction records: order count, transport count, total quantity (units or kilograms), and confirmed production quantity. Day-of-week was encoded as an ordinal integer (0 = Monday through 4 = Friday) and included as a temporal feature in all gate models. After matching ERP features to available survey days and removing rows where all operational features were zero, the final clean dataset comprised 127 survey days across all four gates.

3.4 Analytical Strategy

The analysis proceeded in three stages. First, descriptive statistics were computed for each satisfaction dimension by gate and by day-of-week to characterize the structural distribution of workload. Second, Pearson correlation coefficients were calculated between each operational feature and each target dimension (a, u, p) to identify the most informative predictors. Feature engineering extensions including lag variables and rolling means (3-day and 5-day windows) were systematically tested but yielded no statistically significant improvements in correlation coefficients.

4 Results

4.1 Descriptive Statistics by Gate

Table 3 presents the descriptive statistics for all three workload dimensions. Results are ordered by descending overall workload intensity defined as the arithmetic mean of the three satisfaction dimensions.

Table 3: Descriptive statistics of workload dimensions by gate (scale 0 = very low workload - 4 = very high workload; SD = standard deviation; N = 127 days)

Gate	Routine (a) μ ± SD	Irregular (u) μ ± SD	Problems (p) μ ± SD	Overall μ	N
Gate 3	1.65 ± 1.21	1.66 ± 1.18	1.56 ± 1.21	**1.62 (high-stress)**	127
Gate 4	1.61 ± 1.27	1.47 ± 1.26	1.47 ± 1.27	**1.52 (high-stress)**	127
Gate 6	1.02 ± 1.19	0.72 ± 1.28	0.46 ± 1.04	0.73	127
Gate 5	0.43 ± 0.83	0.21 ± 0.60	0.20 ± 0.62	**0.28 (low-stress)**	127

Source: Own

Gates 3 and 4 exhibit markedly elevated workload across all three dimensions, with mean overall scores of 1.62 and 1.52 respectively. In contrast, Gate 5 has a mean overall score of 0.28 - approximately one-sixth of Gate 3's level - despite hosting the highest production order volume of all gates (8,502 COOIS records; 31.1 confirmed orders per day on average). This dissociation between production volume and perceived workload points to fundamental differences in task characteristics and logistics support structures across gates. Gate 6 occupies an intermediate position with a mean overall score of 0.73, reflecting moderate workload with substantially higher variance in routine task ratings.

4.2 Day-of-Week Temporal Patterns

Table 4 presents mean overall workload by day-of-week for each gate. A consistent temporal pattern emerges across all four gates. The workload intensifies mid-week, peaking between Tuesday and Wednesday, and subsides toward Friday.

Table 4: Mean overall workload (arithmetic mean of a, u, p) by day-of-week per gate. Arrows indicate peak (↑) and trough (↓) days

Gate	Monday	Tuesday	Wednesday	Thursday	Friday
Gate 3	1.68	1.71	**1.75** ↑	1.46	1.41 ↓
Gate 4	1.58	**1.76** ↑	1.65	1.41	1.23 ↓
Gate 5	0.22	0.30	**0.49** ↑	0.26	0.12 ↓
Gate 6	0.60	**0.88** ↑	0.78	0.86	0.47 ↓

Source: Own

The Wednesday peak at Gates 3 and 5 and the Tuesday peak at Gates 4 and 6 are consistent with theoretical expectations from production scheduling literature. Monday's lower utilization (system start-up, deferred confirmations) is followed by mid-week accumulation of open orders and deliveries. The universal Friday decline may reflect reduced new-order initiation, deliberate managerial workload reduction ahead of the weekend, or the anticipation effect described in job satisfaction research. Fridays often yield more positive self-assessments regardless of objective workload (Kahneman et al. 2004).

4.3 Correlation Analysis by Gate

Table 5: Selected Pearson correlation coefficients between ERP features and workload dimensions per gate[2].

Gate	Feature	r (a)	r (u)	r (p)
Gate 3	Daily production order count	**+0.348***	**+0.390***	**+0.438***
	Confirmed production quantity	+0.299	+0.329	**+0.386***
	Material movement count	+0.072	+0.046	+0.048
Gate 4	Total material weight (kg)	**+0.314***	**+0.318***	**+0.313***
	External delivery volume	+0.177	+0.196	+0.157
	External transport count	+0.023	-0.034	-0.065
Gate 5	Internal transport volume (units from)	-0.082	-0.175	**-0.356***
	Internal transport count	-0.062	-0.176	**-0.339***
	Production order count	+0.120	-0.073	-0.028
Gate 6	Internal transport count	+0.161	-0.151	**-0.220***
	Internal transport volume	**+0.201***	-0.099	-0.189

Source: Own

[2] * indicates meaningful effect. a = routine tasks; u = irregular tasks; p = problems

Table 5 presents Pearson correlation coefficients between the most informative ERP features and each workload dimension per gate. Features with near-zero variance or less than 30 observations were excluded.

Several noteworthy patterns emerge from the correlation analysis. At Gate 3, feature "daily production order count" consistently shows the highest correlations with all three workload dimensions, peaking at r = +0.438 for problem perception - suggesting that higher daily order throughput is associated with more frequently reported operational disruptions. "Material movement count" adds negligible explanatory power, likely due to the limited number of observations (43 days with non-zero material movement count at Gate 3).

At Gate 4, feature "total material weight" emerges as the dominant predictor across all three dimensions (r = +0.31 to +0.32), while the raw count of external transports shows near-zero correlations. This asymmetry reveals an important volume-versus-frequency dynamic. It is the physical mass of goods processed rather than the number of transport events that appear to drive employee burden, plausibly because large shipments require extended handling, coordination, and physical effort.

The most striking finding concerns Gates 5 and 6: internal transport count/activity correlates negatively with problem perception (r = -0.356 and r = -0.220 resp.). Contrary to the intuitive expectation that more logistics activity means more work, higher internal transport throughput at these gates is associated with fewer reported problems. This suggests that timely and sufficient material provisioning may coincide with fewer reported work interruptions and wait states - a logistics buffer effect that we formalize as a testable hypothesis in Section 5. At Gate 6, the "internal transportation count" also shows a positive correlation with routine task ratings (r = +0.201), indicating that internal logistics work is classified as routine rather than problematic by these employees.

5 Discussion: Derived Hypotheses and Managerial Implications

5.1 Hypotheses Derived from Empirical Findings

The empirical results give rise to seven hypotheses that warrant confirmation in future studies with larger samples and ideally a control group design.

H1 (Logistics Buffer Hypothesis): Well-functioning internal logistics processes (LT061) reduce perceived workload, particularly problem perception. Evidence: Consistent negative correlations at Gates 5 ($r = -0.356$) and 6 ($r = -0.220$) suggest that timely material provisioning prevents work interruptions and ad-hoc problem-solving episodes.

H2 (Volume-over-Frequency Hypothesis): It is the physical volume of materials processed - not the frequency of transport events – that appear to drive employee workload. Evidence: At Gate 4, material weight (mb51_menge, $r \approx +0.31$) substantially predicts transport count (lt060_transporte, $r \approx 0.02$).

H3 (Mid-Week Stress Hypothesis): Mid-week (Tuesday to Wednesday) represents the peak workload period due to cumulative order and delivery pressure. Evidence: Confirmed across all four gates in Table 4, consistent with production scheduling theory (Buzacott & Shanthikumar 1993).

H4 (Task Character Hypothesis): Internal logistics tasks are perceived as substantially less burdensome than external logistics and material handling tasks, irrespective of volume. Evidence: Gate 5, with the highest production order volume, reports the lowest overall workload; Gates 3 and 4, which manage external material flows, report five to six times higher stress.

H5 (Task Heterogeneity Hypothesis): Gates with multi-source ERP profiles (i.e., heterogeneous task types) report higher workload than gates with focused operational profiles. Evidence: Gate 4 is the only gate active in three ERP data sources and reports the second-highest overall stress, while Gates 5 and 6 show cleaner operational focus and lower stress.

H6 (Data Coverage Hypothesis): ERP sources beyond those currently captured may correlate more strongly with perceived workload at Gate 3 than those used here. Evidence: Gate 3 has the highest mean workload (1.62) but the sparsest ERP coverage (998 COOIS records; no LT061/LT060 data), and observed correlations reach only $r = +0.44$.

H7 (Under-Utilization Hypothesis): Insufficient logistics activity - resulting in idle time and waiting - elevates perceived workload through boredom or frustration. Evidence: The negative LT061 correlation at Gates 5 and 6 may partially reflect the psychological burden of unproductive waiting when deliveries are delayed, complementing the logistics buffer mechanism in H1 or being moved to other Gates or position in company temporarily which could also lead to frustration.

5.2 Managerial Implications

The findings carry several concrete implications for operations managers seeking to leverage existing ERP data for social sustainability management. First, the identification of Gates 3 and 4 as structurally high-stress environments warrants targeted intervention. For Gate 4, the strong relationship between material volume and workload suggests that smoothing the incoming delivery schedule - for example, by negotiating more evenly distributed delivery windows with suppliers - could potentially reduce workload peaks. For Gate 3, the lack of ERP coverage suggests that the initial priority should be to improve data capture before continuing.

Second, the logistics buffer effect at Gates 5 and 6 (H1) suggests a cross-gate best-practice transfer opportunity. The internal transport scheduling practices that keep these gates well-supplied could be studied and partially adopted at Gates 3 and 4. Third, the consistent mid-week stress peak (H3) suggests that targeted workload relief measures - such as additional staffing, prioritized maintenance, or reduced ad-hoc task assignments - should be scheduled for Tuesday and Wednesday rather than distributed uniformly across the week.

Fourth, this study validates the feasibility of the Social Sustainability Insights Tool idea proposed in (Oettmeier et al. 2025). Even with existing ERP data and without any new IT infrastructure investments, it is possible to build gate-specific monitoring dashboards that flag workload-intensive operational configurations.

6 Limitations and Future Research

Several limitations of the present study must be acknowledged. First, the observation window of 127 working days (April to November 2024) covers only two-thirds of the annual calendar, excluding winter months and thus potentially seasonal variation

in workload and satisfaction. A minimum of 24 months of continuous data collection is recommended before the predictive models are used for managerial decision-making.

Second, the survey instrument relies on self-reported ratings, which are susceptible to self-presentation bias, response habituation, and anchoring effects (Podsakoff et al. 2003). The convergence of individual employees' ratings into a single daily gate value conceals inter-individual variation that may be practically important. Future studies should consider supplementing subjective ratings with objective workload indicators such as absenteeism rates, task completion times, or quality error rates.

Third, data protection regulations - specifically the Swiss DSG (Bundesversammlung der Schweizerischen Eidgenossenschaft 2022) - constrain the use of personal employee attributes (age, experience, shift assignment, health data) that are theoretically important covariates in workload models. Future research should explore privacy-preserving aggregation techniques, such as differential privacy or federated learning, to enable the inclusion of such variables within legal constraints.

7 Conclusion

This paper provides the first empirical validation of the claim that transactional ERP data can serve as a meaningful, even if only partial, predictor of daily employee workload satisfaction in manufacturing operations - proving the answer to RQ1. The finding that internal transport activity (LT061) may function as a protective factor - reducing rather than increasing perceived problems at internal logistics gates - challenges the intuitive assumption that more operational activity always means more employee burden. The volume-over-frequency dynamic at the external materials gate and the consistent mid-week stress peak across all gates provide concrete, actionable signals that operations managers can integrate into shift planning and continuous improvement cycles. Therefore, RQ2 has also been answered.

By confirming both, the feasibility and the bounded nature of ERP-based workload prediction, this study advances the design science research program for workplace social sustainability tools and contributes to the intersection of people analytics, operations management, and information systems. The empirical evidence supports

a workplace-specific, context-sensitive approach to deploying such tools rather than generic, facility-wide models. It underscores the importance of first improving data quality and ERP-to-gate assignment completeness at high-workload areas before attempting predictive monitoring.

References

Ajiga, David Iyanuoluwa; Hamza, Oladimeji; Eweje, Adeoluwa; Kokogho, Eseoghene; Odio, Princess Eloho (2024): Assessing the Role of HR Analytics in Transforming Employee Retention and Satisfaction Strategies. In IJSSER 3 (1), pp. 87–94. DOI: 10.54660/IJSSER.2024.3.1.87-94.

Albers Mohrman, Susan; E. Lawler III, Edward (2014): Designing organizations for sustainable effectiveness. In Journal of Organizational Effectiveness: People and Performance 1 (1), pp. 14–34. DOI: 10.1108/JOEPP-01-2014-0007.

Bakker, Arnold B.; Demerouti, Evangelia (2007): The Job Demands-Resources model: state of the art. In Journal of Managerial Psychology 22 (3), pp. 309–328. DOI: 10.1108/02683940710733115.

Bradley, Joseph; Lee, C. Christopher (2007): ERP Training and User Satisfaction. In International Journal of Enterprise Information Systems 3 (4), pp. 33–50. DOI: 10.4018/jeis.2007100103.

Bundesversammlung der Schweizerischen Eidgenossenschaft (2022): Bundesgesetz über den Datenschutz. Available online at https://www.fedlex.admin.ch/eli/oc/2022/491/de.

Conlon, Sumali; Simmons, Lakisha; Liu, Feng (2021): Predicting Tech Employee Job Satisfaction Using Machine Learning Techniques Sumali J. Conlon1 Lakisha L. Simmons2 Feng Liu3. In IJMIT 16, pp. 97–113. DOI: 10.24297/ijmit.v16i.9072.

Dachawar, Madhavi; Ainapure, Bharati; Singh, Bhupender; Kulkarni, Nitiraj (2025): Understanding and Exploring the Relationship Between ERP Usage and Cognitive Workload: A Gender and Age Perspective 13, pp. 417–430.

Espinosa-Pinos, Carlos Alberto; Acosta-Pérez, Paúl Bladimir; Valarezo-Calero, Camila Alessandra (2024): Applying Classification Techniques in Machine Learning to Predict Job Satisfaction of University Professors: A Sociodemographic and Occupational Perspective. In Computers 13 (12), p. 344. DOI: 10.3390/computers13120344.

Hackman, J.Richard; Oldham, Greg R. (1976): Motivation through the design of work: test of a theory. In Organizational Behavior and Human Performance 16 (2), pp. 250–279. DOI: 10.1016/0030-5073(76)90016-7.

Inegbedion, Henry; Inegbedion, Emmanuel; Peter, Adeshola; Harry, Lydia (2020): Perception of workload balance and employee job satisfaction in work organisations. In Heliyon 6 (1), e03160. DOI: 10.1016/j.heliyon.2020.e03160.

J. Buzacott; J. George Shanthikumar (1993): Stochastic models of manufacturing systems (13).

Kahneman, Daniel; Krueger, Alan B.; Schkade, David A.; Schwarz, Norbert; Stone, Arthur A. (2004): A survey method for characterizing daily life experience: the day reconstruction method. In Science (New York, N.Y.) 306 (5702), pp. 1776–1780. DOI: 10.1126/science.1103572.

Karasek, Robert A. (1979): Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. In Administrative Science Quarterly 24 (2), p. 285. DOI: 10.2307/2392498.

Lin, Han Htut; Mon, Thu Ya (2025): Review on the Implementation of Enterprise Resource Planning Software on Employee Job Satisfaction. In J. Multidiscip. Res. SMET 1 (1), pp. 82–89. DOI: 10.64297/jmrsmet.v1i1.8.

Maximilian Siessegger (2025): Zahlen und Fakten: So werden Mitarbeitende befragt. Available online at https://www.hrtoday.ch/de/article/zahlen-und-fakten-so-werden-mitarbeitende-befragt, checked on 2/21/2026.

Megan O'Brien (2021): Optimizing HR Costs With People Analytics, and the CFO's Role. Available online at https://www.netsuite.com/portal/resource/articles/human-resources/people-analytics.shtml.

Mertens, Shana; Schollaert, Eveline; Anseel, Frederik (2021): How much feedback do employees need? A field study of absolute feedback frequency reports and performance. In Int J Selection Assessment 29 (3-4), pp. 326–335. DOI: 10.1111/ijsa.12352.

Missimer, Merlina; Mesquita, Patricia Lagun (2022): Social Sustainability in Business Organizations: A Research Agenda. In Sustainability 14 (5), p. 2608. DOI: 10.3390/su14052608.

Myskova, Renata (2011): A New Measure Of Employee Satisfaction. In Global Journal of Business Research (1), pp. 101–110. DOI: None.

Nijhof, André; Cludts, Stephan; Fisscher, Olaf; Laan, Albertus (2003): Measuring the Implementation of Codes of Conduct. An Assessment Method Based on a Process Approach of the Responsible Organisation. In Journal of Business Ethics 45 (1-2), pp. 65–78. DOI: 10.1023/A:1024172412561.

Oettmeier, Katrin; Bekiri, Valmir; Stöckler, Stefan (2025): Towards a Social Sustainability Insights Tool for Operations. In : Proceedings of the SAP Academic Community Conference 2025 (D-A-CH), p. 90.

Opatha, H.H.D.P.J. (2020): HR Analytics: A Literature Review and New Conceptual Model. In IJSRP 10 (06), pp. 130–141. DOI: 10.29322/IJSRP.10.06.2020.p10217.

Peffers, Ken; Tuunanen, Tuure; Rothenberger, Marcus A.; Chatterjee, Samir (2007): A Design Science Research Methodology for Information Systems Research. In Journal of Management Information Systems 24 (3), pp. 45–77. DOI: 10.2753/MIS0742-1222240302.

Podsakoff, Philip M.; MacKenzie, Scott B.; Lee, Jeong-Yeon; Podsakoff, Nathan P. (2003): Common method biases in behavioral research: a critical review of the literature and recommended remedies. In The Journal of applied psychology 88 (5), pp. 879–903. DOI: 10.1037/0021-9010.88.5.879.

Sageer, Alam (2012): Identification of Variables Affecting Employee Satisfaction and Their Impact on the Organization. In IOSRJBM 5 (1), pp. 32–39. DOI: 10.9790/487X-0513239.

Snigdha; Singh, Animesh (2024): Enhancing Job Satisfaction and Retention in IT Companies Through HR Analytics-Driven Employee Engagement. In : 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT). 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT). Faridabad, India, 28.11.2024 - 29.11.2024: IEEE, pp. 898–905.

Talebi, Hojat; Bardsiri, Amid Khatibi; Bardsiri, Vahid Khatibi (2025): Developing a hybrid machine learning model for employee turnover prediction: Integrating LightGBM and genetic algorithms. In Journal of Open Innovation: Technology, Market, and Complexity 11 (2), p. 100557. DOI: 10.1016/j.joitmc.2025.100557.

Yaneva, Maya (2018): Employee Satisfaction vs. Employee Engagement vs. Employee NPS. In European Journal of Economics and Business Studies 4 (1), pp. 221–227. DOI: 10.2478/ejes-2018-0024.

The AI Implementation Paradox in Danish SMEs: Why Positive Intentions Fail to Translate into Sustained Use

Kenneth Dosanjh
Business Academy Copenhagen, Copenhagen, Denmark
kssd@ek.dk

DOI
https://doi.org/10.18690/um.fov.4.2026.8

ISBN
978-961-299-152-4

A substantial gap persists between organizational intention to adopt artificial intelligence (AI) marketing tools and their sustained use. This study investigates why Danish small and medium-sized enterprises (SMEs) with clear implementation intentions nonetheless fail to implement or scale AI adoption. Drawing on a qualitative multiple case study of six Danish SMEs and integrating the Technology Acceptance Model (TAM), Resource-Based View (RBV), and organizational AI capability research, we identify five mechanisms that systematically undermine intention-to-use translation: expectation-reality mismatches, time paradoxes, organizational structure constraints, tool proliferation, and support vacuums. We introduce the "intention-expansion gap," in which successful initial adoption stabilizes at partial use despite persistent expansion intentions. We extend TAM by positioning implementation capability as a moderator of the intention-behavior relationship and reconceptualize perceived ease of use as dynamically evolving; we extend RBV by showing resource constraints operate as implementation rather than adoption barriers.

Keywords:
SMEs, intention-behavior gap, implementation barriers, AI capability, technology acceptance, Denmark

University of Maribor Press

1 Introduction

The proliferation of accessible artificial intelligence (AI) tools has substantially reduced the technical and financial barriers to AI adoption for organizations of all sizes. Marketing applications, encompassing content generation, customer targeting, and data-driven decision support, are now technically within reach of small and medium-sized enterprises (SMEs) that previously lacked the specialist staff and budgets associated with such capabilities (Brynjolfsson & McElheran, 2016). Yet widespread availability has not produced commensurate adoption. A substantial body of evidence documents a persistent gap between organizational intentions to adopt AI and the sustained use of AI in practice (Ghobakhloo et al., 2021; Dwivedi et al., 2021; Alsheibani et al., 2020).

This gap is particularly salient in Denmark, where digital infrastructure is advanced and technological literacy high relative to international comparisons (Statistics Denmark, 2024). Danish SMEs operate in a high-readiness institutional environment characterized by strong broadband penetration, broad adoption of digital administrative tools, and government-supported digitalization initiatives (Danish Business Authority, 2023). If the intention-behavior gap in AI adoption persists in such an environment, the barriers are likely organizational and processual rather than infrastructural, making Denmark a theoretically informative context for investigating these dynamics.

With approximately 415,000 SMEs comprising 99.7% of Danish businesses (Danish Business Authority, 2023), understanding the AI implementation gap carries significant economic implications. Substantial prior research has examined technology adoption in SMEs, including barriers and drivers of digitalization (Ghobakhloo et al., 2021; Ramdani et al., 2022). However, most work focuses on whether SMEs adopt new technologies rather than on how adoption intentions succeed or fail during implementation. This distinction is particularly consequential for AI marketing tools, where the literature remains sparse and concentrated on large enterprise contexts (Alsheibani et al., 2020; Dwivedi et al., 2021).

This paper addresses three interrelated research questions:

RQ1: What is the nature and extent of the intention-behavior gap regarding AI marketing tool adoption among Danish SMEs?

RQ2: Which mechanisms explain this gap: why do SMEs with positive intentions nonetheless fail to implement AI marketing tools?

RQ3: How can Danish SMEs successfully bridge the intention-behavior gap in AI adoption?

We make three principal theoretical contributions. First, extending TAM, we demonstrate that implementation capability moderates the intention-behavior relationship in resource-constrained contexts, and that perceived ease of use evolves dynamically during implementation. Second, extending RBV, we reveal that resource constraints manifest as implementation barriers rather than adoption barriers. Third, we introduce the "intention-expansion gap", a stable partial-adoption equilibrium driven by satisficing dynamics that prevents broader diffusion despite strong expansion intentions.

2 Theoretical background

2.1 AI Capability, Digital Maturity, and Resource Constraints in SMEs

From a Resource-Based View (RBV), competitive advantage depends on resources that are valuable, rare, inimitable, and organizationally embedded (Barney, 1991). Recent research distinguishes the possession of AI tools from the organizational capacity to extract value from them (Mikalef & Gupta, 2021). AI capability encompasses data resources, technical infrastructure, analytical skills, and the process integration competencies required to embed AI meaningfully into business operations. Bharadwaj (2000) demonstrated that digital capabilities constitute a firm-level resource that translates technology investments into competitive outcomes; without such capabilities, technology adoption produces limited value regardless of tool quality. Alsheibani et al. (2020) extended this perspective to AI readiness, identifying organizational readiness as a prerequisite for realizing AI value at the firm level.

Owner and manager characteristics strongly shape adoption decisions in SMEs, where strategic, financial, and operational authority are frequently concentrated in one person (Thong, 1999). External support networks become critical compensatory resources given limited in-house expertise (Levy et al., 2001). SMEs often operate with informal processes for evaluating and implementing new information systems (Thong et al., 1996), which enables agility but also increases the risk of ad hoc implementations that fail to deliver sustainable value.

2.2 Technology Acceptance Model and the Intention-Behavior Link

The Technology Acceptance Model (TAM), developed by Davis (1989) and extended by Venkatesh et al. (2003), posits that perceived usefulness and ease of use shape attitudes toward technology, influencing behavioral intention and actual usage. However, a critical limitation is the assumption that intention reliably predicts behavior (Sheeran, 2002). Meta-analyses reveal that intentions account for only 28% of variance in behavior (Sheeran, 2002), indicating that even strong intentions frequently fail to materialize into action. This gap is particularly pronounced when behaviors require sustained effort, face implementation obstacles, or compete with other priorities (Gollwitzer & Sheeran, 2006), all characteristics describing AI implementation in resource-constrained SMEs.

In the context of AI adoption in SMEs, TAM remains useful for understanding how owners and managers form positive intentions based on perceived usefulness and ease of use. However, the intention-behavior gap literature suggests that strong intentions alone are unlikely to guarantee sustained use when implementation is time-consuming, technically demanding, and embedded in competing day-to-day operational priorities.

2.3 Intention-Behavior Gaps in Organizational Technology Implementation

Research on implementation intentions (Gollwitzer, 1999) suggests gaps can be bridged through specific planning mechanisms: defining when, where, and how to act; identifying potential obstacles in advance; and creating concrete action plans. However, this research has primarily focused on individual behaviors rather than organizational technology implementation. The translation of individual intentions

to organizational action involves additional complexities including coordination costs, consensus requirements, and the need to align implementation with existing routines and workflows.

For AI technologies specifically, studies show that many firms underestimate the organizational work required to benefit from AI (Alsheibani et al., 2020; Dwivedi et al., 2021). Value creation depends on aligning AI initiatives with business goals and processes rather than simply installing tools (Mikalef & Gupta, 2021). Empirical studies focusing on SMEs identify financial constraints, lack of technical expertise, and limited organizational readiness as persistent barriers (Ghobakhloo et al., 2021).

2.4 Integrated Framework

To understand the AI implementation paradox in Danish SMEs, this study integrates insights from RBV, TAM, AI capability research, and research on intention-behavior gaps. RBV and AI capability research highlight how limited organizational resources and digital maturity make SMEs particularly vulnerable during implementation. TAM explains how SME owners form positive intentions based on perceived usefulness and ease of use. Research on intention-behavior gaps demonstrates that strong intentions frequently fail to produce sustained behavior when implementation is effortful, contextually constrained, or dependent on complex organizational processes.

Within this integrated framework, we focus on the implementation stage of AI adoption. Positive intentions form when AI tools are perceived as useful and accessible. These intentions then encounter the organizational realities of SME resource constraints, digital maturity limitations, and the concrete characteristics of AI implementation. The intention-behavior gap emerges when three conditions coincide: AI tools are presented as attractive and accessible; SME decision-makers express clear adoption intentions; and resource constraints, organizational readiness gaps, and AI-specific implementation complexity undermine sustained use.

3 Methodology

3.1 Research Design

This study adopts an interpretive epistemological stance and employs a qualitative multiple case study design (Yin, 2018; Eisenhardt, 1989). The case study methodology is appropriate for examining contemporary phenomena in which the boundaries between phenomenon and context are blurred. We build theory from multiple cases by iterating between within-case analysis and cross-case comparison (Eisenhardt, 1989).

3.2 Case Selection and Overview

Table 1: Case Overview

Case	Sector	Size	Digital Maturity	AI Tools Attempted	Primary Goal	Outcome
Case 1	Restaurant (B2C)	6 employees	Low-Medium	ChatGPT, Canva	Automate social content; reduce owner time	Failed (abandoned, 3-4 months)
Case 2	Social services (B2G)	14 employees	Medium	ChatGPT, Canva, HubSpot AI	Accelerate client-facing content production	Failed (paused Q1 2025)
Case 3	Restaurant (B2C)	8 employees	Medium-High	ChatGPT, Jasper, Canva AI, Hootsuite, Copy.ai	Identify optimal AI tool for all marketing	Failed (abandoned, 2 months)
Case 4	Consultant (B2G)	1 employee	High	ChatGPT Plus	Augment document writing; 30%+ time savings	Success (sustained, 12+ months)
Case 5	Digital agency (B2B)	10 employees	Very High	Multiple AI tools	Integrate AI across all service delivery	Success (integrated workflows)
Case 6	E-commerce (B2C)	3 employees	High	ChatGPT	Automate product descriptions and customer responses	Partial (limited, failed to expand)

Source: Authors' analysis

We employed purposeful, maximum variation sampling to select six Danish SMEs spanning the main outcome types of AI marketing tool implementation (Patton, 2015). Selection required that each organization: (1) qualified as an SME (fewer than 250 employees); (2) employed marketing activities as a central component of its business model; and (3) had considered, piloted, or implemented AI-based marketing tools within the preceding two years. It is analytically important to distinguish adoption failure (declining to proceed after evaluation), implementation failure (beginning use but discontinuing before achieving sustained integration), and scaling failure (achieving initial success but failing to expand). All six cases experienced some form of post-intention failure, they represent not non-adoption but adoption attempts that did not achieve intended scope. This distinction shapes how we interpret the mechanisms identified in findings.

Recruitment through professional networks introduces a selection bias toward more digitally aware organizations than the average SME. The sample should be interpreted as representing relatively favorable contexts for AI implementation; similar or stronger challenges likely characterize less digitally oriented SMEs.

Note. All cases are Danish SMEs that expressed interest in or attempted AI marketing tool adoption within 2023-2025. Digital maturity assessed based on prior digital tool adoption, process documentation, and staff technical proficiency reported in interviews.

3.3 Data Collection

We employed multiple methods to enable triangulation. Semi-structured interviews were conducted with SME owners and key employees during March through October 2024, ranging from 45 to 90 minutes (mean: 62 minutes), yielding approximately 7.5 hours of recorded data. All interviews were audio-recorded with informed consent and subsequently transcribed.

Interview protocols followed a semi-structured format covering six thematic areas: (1) company background and marketing context; (2) prior technology adoption history and digital maturity; (3) initial intentions and expectations regarding AI tools; (4) the implementation journey, including specific steps taken and barriers encountered; (5) outcomes and current AI use status; and (6) retrospective

assessment and recommendations. Protocols were developed iteratively from the theoretical framework and refined through pilot discussions with two SME owners outside the final sample.

Documentary materials collected during the same period included examples of marketing content produced before and after AI introduction, screenshots from AI tools, configuration notes, and time-tracking records.

Table 2: Data Collection by Case

Case	Interviews (duration)	Observation Sessions	Documents
Case 1	1 interview (58 min)	None	Marketing content examples, tool screenshots
Case 2	2 interviews (45 min, 62 min)	None	Internal AI guidelines, content samples
Case 3	1 interview (71 min)	None	Tool comparison notes, content examples
Case 4	2 interviews (52 min, 48 min)	2 sessions	Work samples, time-tracking records
Case 5	3 interviews (90 min, 65 min, 58 min)	3 sessions	Workflow documentation, tool audit records
Case 6	2 interviews (55 min, 70 min)	1 session	Product descriptions pre/post AI, expansion planning notes

Source: Authors' analysis

3.4 Data Analysis

We followed a systematic five-stage analytical process. Stage 1 produced detailed case narratives capturing context, adoption journey, barriers, and critical turning points, shared with participants for member checking. Stage 2 involved open coding of all transcripts using NVivo, generating approximately 180 initial codes organized into five areas: decision-making processes, barriers encountered, implementation failures, success factors, and external support role.

To illustrate the coding approach: the code "time investment exceeds expectations" emerged from Case 1's owner's account that AI initially added one hour to each social media post, and from Case 2's director's observation that AI output required three to four times the anticipated editing effort. These codes, combined with "learning curve underestimated" and "operational urgency competing with implementation," formed the axial category "time constraints," which contributed to the cross-case mechanism "time paradox."

Stage 3 involved axial coding, organizing 180 initial codes into 12 larger categories. Stage 4 involved systematic cross-case analysis using comparison matrices examining expectation framing, implementation approach, time allocation, organizational structure, tool strategy, and support access. Stage 5 involved theoretical integration, interpreting empirical patterns through the integrated framework and developing propositions. Thematic saturation was reached after the fifth case, with Case 6 providing confirmatory evidence of the intention-expansion gap pattern.

3.5 Research Quality

We attended to multiple criteria for qualitative research quality (Lincoln & Guba, 1985). Credibility was supported through multiple data sources per case, member checking of case narratives, and active search for disconfirming evidence. Transferability was encouraged through thick description of organizational contexts and implementation processes. Dependability was supported by documenting sampling decisions, data collection procedures, and analytical steps in an audit trail. Confirmability was addressed by grounding interpretations in data excerpts and reflecting on researcher positionality.

4 Findings

4.1 Nature and Extent of the Intention-Behavior Gap (RQ1)

All six cases exhibited strong initial intentions toward AI adoption, with rated importance ranging from 7 to 10 out of 10, and all had taken concrete initial steps, including subscribing to tools, allocating time for experimentation, or aligning staff around adoption. Despite this, three cases (Cases 1-3) abandoned AI within two to six months, one case (Case 6) achieved partial implementation that did not expand

as intended, and even the two successful cases (Cases 4 and 5) reported implementation complexity substantially exceeding initial expectations.

The gap operates along two dimensions. A binary gap between intention and any sustained use characterizes Cases 1-3, where abandonment occurred despite strong initial intentions. A graduated gap between intention and intended scope characterizes Cases 4-6, where implementation was achieved but fell short of broader goals initially envisioned. This two-dimensional character refines the binary framing common in adoption literature, suggesting a spectrum of partial implementation outcomes that the conventional adopted/not-adopted distinction cannot capture.

4.2 Failed Implementations: The Expectation-Reality Collapse

All three failed cases (Cases 1-3) abandoned AI within two to six months despite initial enthusiasm and high importance ratings (6-7 out of 10), exhibiting progressive disillusionment as implementation reality contradicted expectations.

Case 1 conceptualized AI as a "digital employee" that would autonomously generate social media content. Initial trials produced outputs requiring substantial editing to match the restaurant's authentic voice and local cultural context. The owner reported: "I thought it would save time, but initially it took more. I had to both learn the tools and correct their output." Creating a single Instagram post using AI required approximately one hour more than manual creation. After three months, the owner abandoned AI entirely. Key barriers included time investment exceeding expectations, expectation mismatch between plug-and-play assumptions and learning curve reality, and absence of contextually appropriate guidance.

Case 2 encountered a values alignment problem apparent only through implementation. For a social services provider where authenticity and professional empathy constitute core organizational identity, AI-generated content carried a register that staff experienced as incompatible with their professional standards. The director reported: "We had to put it somewhat on pause after Q1 2025. The time it took to get outputs that matched our values was simply too much." Time requirements exceeded initial estimates by a factor of three to four. Organizational

structure concentrated all AI experimentation burden on the director, while field staff remained focused on service delivery.

Case 3, despite the owner's high digital fluency, illustrates how tool proliferation undermines implementation even where individual tools are technically manageable. Simultaneously testing ChatGPT, Jasper, Canva AI, Hootsuite, and Copy.ai, the owner reported: "I felt I was constantly switching between tools and never became really good at any of them. In the end, I couldn't justify continuing when I wasn't seeing clear results." Abandonment occurred after approximately two months.

Cross-case pattern: Failed implementations shared three characteristics: automation framing expecting AI to work autonomously rather than as augmentation requiring human expertise; ad hoc approach lacking structured implementation plans; and resource-time mismatches preventing sustained learning investment.

4.3 Successful Implementations: Realistic Expectations and Strategic Focus (RQ3, partial)

Case 4 achieved sustained success through four complementary factors: realistic augmentation framing, focused tool selection, deliberate time protection, and the structural advantage of solo operation. The consultant framed AI as a collaborative tool from the outset: "I knew from the start that it couldn't replace my expertise. I know the procurement processes, I know what evaluators look for. But I hoped AI could help with formulation and structuring." This framing transformed moments when AI required substantial expert input into confirmation of expectations rather than violations of them.

Strategic time allocation proved decisive: "I blocked 2 hours every Friday morning for experimenting and learning. I treated it like client work: non-negotiable time." This protected period enabled the sustained learning investment that failed cases could not maintain. As solo operator, she faced zero organizational barriers, all adoption decisions were concentrated in one technically capable individual with clear motivation and no coordination overhead. After approximately six weeks of dedicated learning, she achieved 30 to 40 percent time savings on document creation while maintaining output quality. Implementation has been sustained for over one year.

Case 5, the digital agency, succeeded through organizational capability compensating for the coordination complexity of a distributed ten-person team. The agency's technical sophistication meant team members framed AI adoption as an extension of existing digital practice. The founder reported: "In January 2023, I told the team: AI is not . We're going all-in. Everyone needs to start experimenting and learning." An internal Slack channel for knowledge sharing and the presence of team members with prior AI experimentation experience provided the contextual support that failed cases lacked.

The founder acknowledged, however, that success carried its own complexity: "I was looking at our tool subscriptions and realized we were spending over 1,500 euros per month on various AI tools. I tried to calculate ROI and couldn't clearly show which tools were worth it." One team member observed that they were "spending more time managing tools than saving with AI", a second-order challenge requiring ongoing organizational attention.

Cross-case pattern: Successful implementations shared augmentation framing, focused mastery strategy (one tool for one specific use case before expanding), and either structural simplicity (solo operation eliminating coordination costs) or organizational capability (established processes, technical sophistication, distributed implementation responsibility).

4.4 Partial Success: The Intention-Expansion Gap

Case 6 reveals a distinct failure mode: successful initial adoption that stabilizes without expanding despite persistent intentions. The e-commerce owner successfully adopted ChatGPT for product descriptions and customer service responses, achieving 40 to 50 percent time savings on these specific tasks over eight months of sustained use. The owner recognized clear expansion opportunities: "I know I should be using it for video scripts, newsletters, social media posts. I see others doing it successfully. I want to expand, but somehow it just doesn't happen."

4.5 Five Mechanisms Underlying the Intention-Behavior Gap (RQ2)

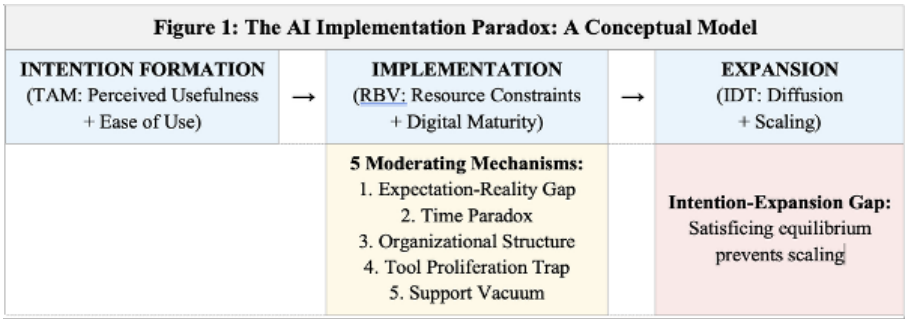

Figure 1: The AI Implementation Paradox: A Conceptual Model
Source: Authors' analysis

Synthesizing patterns across all six cases reveals five mechanisms that systematically undermine the translation of positive intentions into sustained AI implementation.

Mechanism 1: The Expectation-Reality Gap: All failed cases exhibited expectation-reality gaps across four dimensions. Capability expectations: failed cases framed AI as autonomous agent, successful cases as augmentation requiring human judgment. Ease-of-use expectations: failed cases anticipated plug-and-play simplicity; successful cases anticipated learning curves requiring dedicated effort. Time investment expectations: failed cases expected immediate savings; successful cases anticipated upfront investment before returns. Actual time requirements in failed cases ran three to four times higher than expected. Timeline expectations: failed cases expected value within days or weeks; successful cases planned for four to eight-week learning periods.

Mechanism 2: The Time Paradox: SMEs face a circular constraint where AI promises time savings but requires upfront time investment to realize those savings. For owners operating at full capacity with no organizational slack, no time is available to learn the tool that would eventually save time. Failed cases could not sustain learning investment long enough to realize eventual returns. Operational urgencies consistently won priority battles against implementation activities. Successful cases either possessed capacity slack enabling protected learning time or deliberately created such slack through strategic time blocking.

Mechanism 3: Organizational Structure as Implementation Moderator: Organizational structure proved more decisive than size in determining implementation success. The solo operator (Case 4) achieved the cleanest success by eliminating coordination overhead entirely. Small co-located teams (Cases 1-3) concentrated implementation burden on already-overextended owners. The distributed team (Case 5) encountered the highest coordination costs yet succeeded through strong leadership, clear role definitions, and distributed implementation responsibility across technically capable team members. Digital maturity moderated this relationship: Case 5's very high digital maturity provided the organizational infrastructure, documented processes, established communication channels, a culture of tool experimentation, that enabled coordination. Cases 1-3's lower maturity levels meant that AI implementation competed with, rather than built upon, existing digital workflows.

Mechanism 4: The Tool Proliferation Trap: Case 3's multi-tool exploration strategy created cognitive overload and decision paralysis. Attempting to simultaneously evaluate and learn five tools fragmented attention, preventing mastery of any single tool. This contrasts sharply with successful cases' focused mastery approach, selecting one tool, achieving competence in one specific use case, then potentially expanding. Even Case 5, which ultimately used multiple tools, built this breadth sequentially through organized team-based experimentation rather than simultaneous individual evaluation.

Mechanism 5: The Support Vacuum: All failed cases attempted implementation without access to contextually appropriate guidance. Professional consulting proved uniformly unaffordable. Peer learning networks provided validation but no concrete context-matched solutions. Vendor support addressed interface functionality but not business context integration. Successful cases accessed contextually relevant guidance: Case 4 found a LinkedIn group for consultants sharing practical AI prompts for procurement documents; Case 5 had internal team members with prior experimentation experience who could guide others.

5 Discussion

5.1 Theoretical Contributions

Extending the Technology Acceptance Model (RQ1, RQ2): Our findings extend TAM by demonstrating that in resource-constrained contexts, implementation capability moderates the relationship between behavioral intention and actual usage. Traditional TAM posits that perceived usefulness and perceived ease of use shape attitudes, which influence behavioral intention, leading to actual usage (Davis, 1989; Venkatesh et al., 2003). Cases 1 and 2 formed strong intentions (rating importance 7 out of 10) based on perceived usefulness and initial perceived ease of use, yet failed to achieve sustained implementation.

Our data reveal that implementation capability, comprising time availability, external support access, organizational readiness, and realistic expectation framing, moderates whether intentions translate into sustained behavior. When implementation capability is low, even strong intentions fail to produce sustained use. When implementation capability is high, moderate intentions can translate into successful implementation.

Proposition 1: In resource-constrained SME contexts, the relationship between behavioral intention and actual usage behavior is moderated by implementation capability, including time availability, external support access, organizational readiness, and realistic expectation framing.

Furthermore, our findings reveal that perceived ease of use evolves dynamically during implementation. TAM treats perceived ease of use as a relatively stable pre-adoption assessment (Davis, 1989). However, we observed perceptions shifting dramatically as users discovered complexity hidden behind simple interfaces. Tools appearing simple in demonstrations proved complex in practice, requiring prompt engineering skill, output evaluation capability, and workflow integration knowledge. Failed cases experienced this evolution as expectation violation; successful cases anticipated it.

Proposition 2: Perceived ease of use in AI adoption is not a stable pre-adoption perception but a dynamically evolving assessment that transforms during implementation as users discover the gap between interface simplicity and implementation sophistication.

Extending Resource-Based View (RQ2): Our findings extend RBV by demonstrating temporal dynamics of how resource constraints operate as implementation barriers rather than merely adoption barriers. Resource constraints often appear manageable at the intention stage but become prohibitive during implementation for three reasons: implementation costs exceed anticipated expenses as organizations discover hidden costs of learning and workflow adaptation; time demands compete with operational priorities while SMEs lack organizational slack to absorb both simultaneously; and knowledge gaps become apparent only during implementation when abstract understanding confronts practical reality.

Proposition 3: Resource constraints operate as implementation barriers rather than merely adoption barriers, preventing intention-to-behavior translation rather than intention formation itself, with this effect mediated by temporal dynamics where constraints appear manageable pre-adoption but become prohibitive during implementation.

Introducing the Intention-Expansion Gap (RQ1, RQ2): Case 6 reveals a previously undertheorized failure mode. Initial adoption succeeds with measurable value (40 to 50 percent time savings). Strong expansion intentions exist, repeatedly expressed and clearly articulated. Yet expansion persistently fails despite clear recognition of unrealized value, absence of technical barriers, and availability of resources.

This pattern challenges Innovation Diffusion Theory's assumption that successful initial adoption enables broader organizational diffusion (Rogers, 2003). We identify satisficing dynamics as the mechanism sustaining this equilibrium. Initial success creates a "good enough" state that reduces urgency for improvement. Time saved is absorbed into other activities.

5.2 Practical Contributions (RQ3)

Our findings suggest that the intention-behavior gap in AI adoption is not primarily a technology problem but an organizational readiness and implementation support problem. The practical implications follow accordingly, differentiated by SME profile.

For early-stage adopters (Cases 1-3 profile (low to medium digital maturity, first AI attempt): The priority is expectation calibration. Assume implementation will require two to three times the effort suggested by vendor materials. Frame AI as augmentation requiring human expertise rather than autonomous replacement.

6 Limitations and future research

This study has several limitations. With six cases drawn predominantly from Danish service-sector SMEs, findings provide rich contextual insights but limited statistical generalizability. Future research should examine whether similar mechanisms appear in manufacturing, retail, or other national contexts.

The retrospective nature of data collection means we capture participants' recollections rather than real-time processes. Recruitment through professional networks likely selected for more digitally aware organizations. The challenges identified here may be more pronounced in less digitally oriented SMEs, suggesting that findings represent a lower bound on implementation difficulty for the broader SME population.

Finally, AI marketing tools are rapidly evolving. Some barriers identified, particularly around interface complexity and learning curves, may diminish as tools improve. Longitudinal research tracking how implementation challenges evolve as technologies mature would be valuable.

7 Conclusion

This study investigated why Danish SMEs that recognize AI marketing tools' value and express strong intentions to adopt them nonetheless fail to implement successfully. Through qualitative analysis of six cases representing failed, successful,

and partial implementation outcomes, we identified five mechanisms underlying the intention-behavior gap: expectation-reality mismatches, time paradoxes, organizational structure constraints, tool proliferation complexity, and support vacuums.

Successful cases differed from failed cases not in initial enthusiasm but in four organizational characteristics: realistic expectations about implementation scope; strategic resource commitment including protected time; access to contextually relevant external support; and structured implementation approaches rather than ad hoc experimentation.

We make three theoretical contributions. First, we extend TAM by demonstrating that implementation capability moderates the intention-behavior relationship and that perceived ease of use evolves dynamically during implementation. Second, we extend RBV by revealing temporal dynamics of resource constraint effects. Third, we introduce the intention-expansion gap as a theoretically distinct phenomenon requiring different mechanisms than those enabling initial adoption.

References

Alsheibani, S., Cheung, Y., & Messom, C. (2020). Artificial intelligence adoption: AI-readiness at firm-level. Proceedings of the 22nd Pacific Asia Conference on Information Systems (PACIS 2018), 37-49.

Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99-120.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. MIS Quarterly, 24(1), 169-196.

Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. American Economic Review, 106(5), 133-139.

Caldeira, M. M., & Ward, J. M. (2003). Using resource-based theory to interpret the successful adoption and use of information systems and technology in manufacturing small and medium-sized enterprises. European Journal of Information Systems, 12(2), 127-141.

Danish Business Authority. (2023). SME statistics Denmark 2023. Copenhagen: Danish Business Authority.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319-340.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., et al. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994.

Eisenhardt, K. M. (1989). Building theories from case study research. Academy of Management Review, 14(4), 532-550.

Ghobakhloo, M., Iranmanesh, M., Grybauskas, A., Vilkas, M., & Petraite, M. (2021). Industry 4.0, innovation, and sustainable development: A systematic review and a roadmap to sustainable innovation. Business Strategy and the Environment, 30(8), 4237-4257.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. American Psychologist, 54(7), 493-503.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. Advances in Experimental Social Psychology, 38, 69-119.

Levy, M., Powell, P., & Yetton, P. (2001). SMEs: Aligning IS and the strategic context. Journal of Information Technology, 16(3), 133-144.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage Publications.

Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. Information & Management, 58(3), 103434.

Patton, M. Q. (2015). Qualitative research & evaluation methods (4th ed.). Thousand Oaks, CA: Sage Publications.

Ramdani, B., Binsaif, A., & Boukrami, E. (2022). Business model innovation: A review and research agenda. New England Journal of Entrepreneurship, 22(2), 89-108.

Rogers, E. M. (2003). Diffusion of innovations (5th ed.). New York: Free Press.

Sheeran, P. (2002). Intention-behavior relations: A conceptual and empirical review. European Review of Social Psychology, 12(1), 1-36.

Statistics Denmark. (2024). ICT usage in enterprises 2024. Copenhagen: Statistics Denmark.

Thong, J. Y. (1999). An integrated model of information systems adoption in small businesses. Journal of Management Information Systems, 15(4), 187-214.

Thong, J. Y., Yap, C. S., & Raman, K. S. (1996). Top management support, external expertise and information systems implementation in small businesses. Information Systems Research, 7(2), 248-267.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425-478.

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). Thousand Oaks, CA: Sage Publications.

AI Adoption and Capability Gaps in Swiss Public Administration

Claudia Pedron, Hans-Dieter Zimmermann, Matthias Baldauf

OST – Eastern Switzerland University of Applied Sciences, Institute for Information and Process Management (IPM), St. Gallen, Switzerland
claudia.pedron@ost.ch, hansdieter.zimmermann@ost.ch, matthias.baldauf@ost.ch

DOI
https://doi.org/10.18690/um.fov.4.2026.9

ISBN
978-961-299-152-4

Artificial intelligence (AI) is widely regarded as an important technology in public administration, yet empirical evidence on the conditions that shape broader AI adoption remains limited. This paper examines digitalization, AI adoption, and implementation barriers across Swiss municipalities with populations of at least 10,000 and cantonal administrations in German-speaking Switzerland. The analysis is based on a 2025 anonymous survey, with one respondent per organization responsible for digitalization and/or AI-related initiatives. The study covers internal administrative processes and externally oriented administrative services, distinguishing between non-AI digital systems and AI-related applications. The findings suggest that digitalization is often comparatively advanced, while AI adoption remains concentrated in assistive applications and is still limited in analytical and decision-adjacent forms. Reported barriers include legal uncertainty, limited expertise, resource constraints, and fragmented responsibilities. The results indicate that broader AI adoption is driven not only by the availability of technology, but also by organizational and governance-related conditions.

Keywords:
artificial intelligence, public administration, digital maturity, dynamic capabilities, governance integration

University of Maribor Press

1 Introduction: AI Ambition Meets Institutional Reality

Driven by the broader digital transformation of the public sector, artificial intelligence (AI) is increasingly regarded as a key technology for improving administrative efficiency and service delivery (European Commission, 2020; 2024; Mergel et al., 2019; Vial, 2019). At the same time, public administrations operate under specific institutional conditions. Accountability, transparency requirements, and democratic legitimacy make the use of AI – especially in decision-relevant contexts – particularly sensitive with regard to explainability, legal compliance, and governance design (Bannister & Connolly, 2014; Janssen & van den Hoven, 2023; Wirtz et al., 2019; Zuiderwijk et al., 2021). AI adoption in public administration should therefore be understood not as a purely technical matter, but as an organizational and institutional process.

Although many public administrations have made substantial progress in digitalization, more advanced forms of AI use often remain limited and selectively integrated. Existing research has explored AI adoption and capabilities within public organizations, yet there remains a limited understanding of how widespread AI deployment aligns with existing levels of digitalization across administrative processes and which barriers constrain its further implementation in public administration (Mikalef et al., 2023). Against this background, this paper examines Swiss municipalities and cantonal administrations in German-speaking Switzerland[1] and focuses on two empirical dimensions: first, the degree of digitalization and AI adoption in externally oriented administrative services and internal administrative processes; second, the organizational, technical, legal, and cultural barriers constraining further implementation. AI adoption is understood as the reported organizational use and intended near-term implementation of AI-related applications. For the purpose of this study, AI encompasses applications such as text generation, chatbot-based interaction, classification, analysis, prediction, and decision-related support. By contrast, online forms, payment systems, appointment-booking systems, document management, enterprise resource planning systems, and cloud solutions are treated as elements of digitalization rather than AI applications.

[1] Switzerland's federal structure allocates administrative responsibilities across municipalities and cantons, creating heterogeneous resource capacities, legal interpretations, and coordination mechanisms that condition AI capability development.

The paper addresses the following research questions:

- RQ1: How developed are externally oriented administrative services and internal administrative processes in terms of digitalization and AI adoption in the surveyed administrations?
- RQ2: What organizational, technical, legal, and cultural barriers constrain the implementation of AI?

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature and introduces the analytical perspective. Section 3 outlines the research design and methodology. Section 4 presents the empirical results. Section 5 discusses the findings in relation to the literature. Section 6 discusses implications, limitations, and conclusions.

2 Literature Review and Theoretical Background

This paper does not treat AI adoption as the mere introduction of a new technology. Instead, it understands AI adoption in public administration as an organizational process shaped by institutional conditions and governance requirements (Janssen & van den Hoven, 2023; Mikalef et al., 2023). Established adoption frameworks such as the Technology-Organization-Environment perspective provide important reference points for studying technology adoption at the organizational level (Tornatzky & Fleischer, 1990). However, the present study does not seek to test a formal adoption model. Instead, the following review is organized around three analytical perspectives that guide the interpretation of the empirical findings. First, research on digital development in the public sector highlights the importance of established infrastructures, process integration, and the availability and quality of data (Mergel, 2019; Vial, 2019). Second, research on AI adoption points to additional requirements, including analytical expertise, coordination across units, and risk and governance management (Mikalef et al., 2023). Third, dynamic capabilities and organizational learning provide a perspective for understanding how organizations adapt structures and routines when moving from digitalization to more advanced forms of AI adoption (Argyris & Schön, 1978; Teece, 2007). Taken together, these perspectives help explain differences in digitalization, AI adoption, and implementation barriers across administrative contexts.

2.1 Digital Development in the Public Sector

Research on digital development in public administration commonly describes a shift from basic information provision to more transactional and integrated digital services (Mergel, 2019). More recent work also emphasizes interoperability, process integration, and data-driven management. In this paper, digital development refers to the degree to which administrative services and processes are digitally supported and integrated. Established systems such as enterprise resource planning or document management can improve efficiency and coordination, but they do not automatically imply more advanced forms of AI adoption or governance-related integration (Vial, 2019). Accordingly, this study treats digital development as an important empirical dimension and examines how it relates to reported patterns of AI adoption across administrative processes.

2.2 AI Adoption in Public Administration

AI adoption research highlights that the implementation of AI depends not only on technological availability, but also on organizational conditions and capabilities (Mikalef et al., 2023). AI requires high-quality data, analytical expertise, cross-departmental coordination, and institutionalized risk governance. In the public sector, AI is additionally shaped by requirements for legitimacy, transparency, and accountability (Janssen & van den Hoven, 2023). As a result, the implementation of AI in public administration is influenced not only by technical capacity, but also by legal, organizational, and governance-related conditions.

2.3 Dynamic Capabilities and Organizational Learning

Dynamic capabilities explain structural adaptation as the ability to recognize opportunities, seize them, and reconfigure organizational routines (Teece, 2007). Integrating AI into governance-relevant processes precisely requires such reconfiguring capabilities. Organizational learning approaches complement this perspective by distinguishing between incremental optimization (single-loop) and the fundamental transformation of decision-making logics (double-loop) (Argyris & Schön, 1978). Institutional frameworks can limit these profound transformations. In this paper, these approaches are used as interpretive perspectives for discussing

differences in AI adoption and implementation barriers across administrative contexts.

3 Method

This study employed a cross-sectional online survey to deliver a descriptive assessment of digitalization, AI adoption, and implementation barriers within public administration. The survey focused on both internal administrative processes and externally oriented administrative services and was conducted between September and October 2025.

3.1 Sample

The sample comprised Swiss municipalities with populations of at least 10,000 and cantonal administrations in German-speaking Switzerland. Invitations to participate in an anonymous online survey were distributed to 145 municipalities and 35 cantonal administrations. The survey was designed to elicit responses from a single representative per organization, specifically individuals responsible for digitalization and/or AI-related initiatives. In total, responses were obtained from 21 municipalities and 13 cantonal administrations. All results are reported in aggregated form to ensure anonymity. This design aimed to capture organization-level assessments from respondents with relevant knowledge of digitalization and AI-related activities. The moderate response rate, particularly at the municipal level, limits the generalizability of the findings. Accordingly, the results should be interpreted as indicative rather than representative.

3.2 Measurement and Analysis

The questionnaire consisted of four parts: general organizational information; digitalization and AI use in internal administrative processes; digitalization and AI use in externally oriented services for citizens, companies, and organizations; and perceived barriers to digitalization and AI projects. In both internal administrative processes and externally oriented services, respondents assessed the overall level of digitalization and AI use using five-point Likert scales. They also indicated the current and planned use of predefined non-AI digital systems, digital services, and AI-related applications through predefined categorical multiple-response items. For internal administrative processes, the questionnaire additionally assessed the current

level and perceived potential of digitalization across specific process categories. Implementation barriers were likewise recorded through predefined categorical multiple-response items covering organizational, technical, legal, cultural, personnel-related, and legitimacy-related constraints. The data were analyzed descriptively using frequency distributions and comparisons across administrative levels. Given the exploratory purpose of the study, no inferential statistical tests were performed. Appendix A provides a structured overview of the survey instrument, the measurement approach, and selected survey questions.

4 Digitalization, AI Adoption, and Implementation Barriers

This section presents the empirical results with respect to the two research questions. It begins by analyzing the extent of digitalization and AI adoption in externally oriented administrative services and internal administrative processes (RQ1) and subsequently examines the key barriers that constrain further implementation (RQ2).

4.1 Digitalization in Externally Oriented Administrative Services and Internal Administrative Processes (RQ1)

The reported degree of digitalization in externally oriented administrative services is predominantly medium to high at both administrative levels. Thirteen of 21 municipalities and 9 of 13 cantonal administrations place these services in the medium-to-high range, whereas lower levels are reported by 8 municipalities and 3 cantonal administrations. This pattern is consistent with the broad availability of core digital administrative services such as online forms, digital transactions, and payment functions. At the same time, more specialized or user-oriented functions are less consistently developed.

A similar but more differentiated pattern appears in internal administrative processes. Among municipalities, 12 of 21 respondents assess the degree of digitalization of internal processes as medium and 7 as high, while 3 rate it as low. Among cantonal administrations, 8 of 13 respondents report medium levels and 4 high levels, while 2 rate the degree of digitalization as low. Differences become more visible when individual process domains are considered. Infrastructure-oriented support areas such as IT and internal communication or financial management show comparatively high levels of digitalization. By contrast, leadership support, steering,

and quality-related processes remain less digitally developed. For example, 19 of 21 municipalities and 11 of 13 cantonal administrations rate leadership support and steering as low to medium, while the perceived digitalization potential of this area is high to very high in 20 of 21 municipalities and 8 of 13 cantonal administrations. A similar discrepancy is visible in quality management. Overall, the findings indicate that operational and support-oriented digitalization is more advanced than the digitalization of leadership- and governance-related internal processes.

4.2 AI Adoption in Externally Oriented Administrative Services and Internal Administrative Processes (RQ1)

Compared with digitalization, AI adoption is considerably less advanced across both externally oriented administrative services and internal administrative processes. In internal administrative processes, 16 of 21 municipalities and 12 of 13 cantonal administrations rate current AI use as very low to low, while only 1 municipality reports a high level of current use.

At the same time, the data show important differences between AI application types. Current use is concentrated in assistive forms of AI adoption. Generative AI systems are currently used by 11 of 21 municipalities and 9 of 13 cantonal administrations, while broader AI-assisted support systems are reported by 8 municipalities and 6 cantonal administrations. More advanced analytical and decision-related applications remain much less common. For example, document and information analysis is currently used by only 3 municipalities and 3 cantonal administrations, pattern or anomaly detection by 2 municipalities and 4 cantonal administrations, and decision-support tools by only 1 municipality and no cantonal administration. Process automation is not currently reported by any municipality or canton in the sample. At the same time, intended near-term implementation is substantially higher in many of these categories. Document and information analysis is planned by 12 municipalities and 8 cantonal administrations, process automation by 9 municipalities and 8 cantonal administrations, and decision support by 8 municipalities and 9 cantonal administrations. This points to a clear gap between current organizational use and expected expansion. Figure 1 illustrates that the gap between current use and planned implementation is particularly pronounced for analytical and decision-support AI applications, whereas assistive AI is already more widely used.

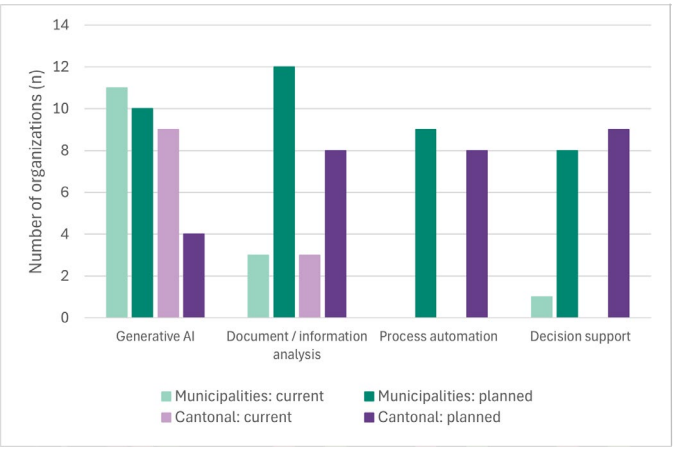

Figure 1: Current and planned AI adoption across application types in internal administrative processes by administrative level

Source: Own

A similar pattern can be observed in externally oriented administrative services. Here too, AI adoption remains limited overall: 16 of 21 municipalities and 10 of 13 cantonal administrations rate current AI use as very low to low. At the same time, selected assistive AI-related functions are already in use in a relevant minority of organizations. Chatbots or virtual assistants are reported by 10 municipalities and 8 cantonal administrations, and AI-supported form assistance or pre-checking by 10 municipalities and 5 cantonal administrations. More advanced analytical applications remain less common, for example forecasting of processing times (5 municipalities; 2 cantonal administrations), anomaly detection in entries (7 municipalities; 3 cantonal administrations), or feedback and sentiment analysis (5 municipalities; 1 cantonal administration).

Taken together, the results point to a consistent pattern across both administrative contexts. AI adoption is most visible where it supports relatively low-risk and assistive tasks, while analytically demanding or decision-adjacent applications remain rare. At the same time, the comparatively high level of intended near-term implementation in several categories indicates that many administrations see future relevance in broader AI deployment, even though current implementation remains selective.

4.3 Implementation Barriers (RQ2)

The main barriers to AI implementation are not primarily technological, but organizational, legal, and governance-related. The most frequently reported barriers include legal uncertainty regarding AI, lack of personnel and budget, fragmented responsibilities, and limited digital and AI-related skills among staff. Legal uncertainty regarding AI is reported by 18 of 21 municipalities and 6 of 13 cantonal administrations, while lack of personnel and budget is reported by 17 municipalities and 10 cantonal administrations. Fragmented responsibilities are mentioned by 14 municipalities and 8 cantonal administrations, and lack of digital skills by 16 municipalities and 8 cantonal administrations.

Some barriers show clear differences between administrative levels. Legal uncertainty and data protection requirements are especially pronounced among municipalities. Complex data protection requirements are reported by 17 of 21 municipalities, but only by 2 of 13 cantonal administrations. By contrast, legacy systems are more frequently emphasized by cantonal administrations (9 of 13) than by municipalities (8 of 21), suggesting that system integration and structural complexity are particularly relevant at the cantonal level. Other barriers, such as lack of model transparency or explainability, unclear user needs, or complex communication of AI, are present at both levels but less dominant than legal and organizational constraints.

5 Discussion: AI Adoption as a Governance-Mediated Capability Challenge

The empirical results reveal a clear discrepancy between the relatively advanced state of digitalization and the still selective use of AI in the surveyed administrations. This discussion interprets this discrepancy through the analytical perspectives introduced in Section 2. Rather than treating AI adoption as a linear extension of digitalization, it examines how existing digital infrastructures, organizational learning processes, and institutional constraints shape the development of AI-related capabilities in public administration.

5.1 An Asymmetrical Pattern of AI adoption

A first key finding is the uneven pattern of AI adoption across application types and process domains. Assistive AI applications are already used in several organizations, whereas analytical and decision-adjacent uses remain much less common despite considerable planned implementation. This pattern is reinforced by the lower level of digitalization in leadership support, steering, and quality-related processes, where AI use would require stronger data integration, governance capacity, and organizational coordination. The findings therefore point to an asymmetry between early assistive use and deeper AI integration.

5.2 Between Incremental Optimization and Broader Organizational change

A second key finding concerns the character of current AI adoption. Much of the observed use supports existing workflows rather than transforming administrative routines or decision-making structures. From the perspective of AI adoption research and organizational learning, the findings are therefore more consistent with selective, incremental adoption than with broader organizational transformation, especially where AI would need to be embedded in management, coordination, or governance-relevant processes.

5.3 Institutional Constraints on Broader AI Deployment

The reported barriers indicate that broader AI deployment is constrained less by the absence of digital tools than by legal, organizational, and governance-related conditions. Legal uncertainty, limited expertise, fragmented responsibilities, and resource constraints are more salient than purely technical obstacles. While the overall barrier profile is broadly similar across administrative levels, the relative weight of specific constraints differs. Municipalities report particularly strong legal and competence-related constraints, whereas cantonal administrations more often emphasize legacy systems and integration complexity. This suggests that the organizational adaptation required for broader AI adoption is shaped by institutional context and administrative structure.

5.4 Synthesis

Taken together, the findings suggest that the surveyed administrations have often moved beyond basic digitalization, but not yet toward structurally embedded AI adoption. The central challenge lies in translating digital infrastructures into organizational capabilities for more advanced and governance-relevant AI use. This points less to a linear stage model of AI development than to a capability gap shaped by organizational learning, institutional conditions, and governance requirements.

6 Implications, Limitations, and Conclusions

This section synthesizes the study's key findings by outlining their implications, acknowledging important limitations, and presenting the overall conclusions.

6.1 Implications and Contributions

The findings have implications for administrative practice, intergovernmental coordination, and research on AI adoption in public administration. For practice, they suggest that administrations should distinguish more clearly between digitalization and AI adoption. Established digital infrastructures may provide an important basis for AI use, but they do not automatically lead to analytically advanced or governance-relevant applications. Broader AI implementation therefore requires clearer responsibilities, stronger internal competencies, and governance arrangements that address legal certainty, data protection, transparency, and accountability.

For governance and policymaking, the results indicate that implementation strategies should be sensitive to administrative context. Municipalities appear particularly affected by legal uncertainty, resource constraints, and competence gaps, whereas cantonal administrations more often face legacy systems and integration complexity. This points to the value of shared guidance, exchange formats, and cooperative approaches to capability building across administrative levels.

This study contributes to the literature in three ways. First, it provides empirical evidence that relatively advanced digitalization does not necessarily translate into broader AI adoption in public administration. This highlights a gap between existing

digital infrastructures and the organizational conditions required for more advanced AI use. Second, it identifies a consistent adoption pattern in which AI use remains concentrated in assistive applications, while analytical and decision-adjacent applications remain limited, particularly where AI would require stronger integration into leadership, steering, and quality-related processes. Third, it shows that perspectives from AI adoption research, dynamic capabilities, and organizational learning are useful for interpreting this gap not primarily as a technological constraint, but as a capability and governance challenge shaped by organizational learning, institutional conditions, and administrative coordination.

6.2 Limitations

This study is subject to several limitations. First, it is based on a cross-sectional survey design and therefore captures digitalization, AI adoption, and implementation barriers only at a single point in time. Second, the results rely on self-reported assessments, meaning that perception bias cannot be excluded. Third, the moderate response rate and possible selection effects limit the generalizability of the findings. Fourth, the analysis is restricted to municipalities and cantonal administrations in German-speaking Switzerland. Finally, the study is descriptive in nature and does not allow causal conclusions.

6.3 Conclusions

Overall, the study shows that AI adoption in public administration cannot be understood as a simple extension of digitalization. While many surveyed administrations report comparatively advanced digitalization in operational and support-related processes, AI use remains selective and concentrated in assistive applications. More analytical and decision-adjacent forms of AI remain limited despite considerable planned implementation. The main challenge is therefore not the absence of digital infrastructure, but the capability gap between existing digital systems and the organizational, legal, and governance capacities required for broader AI deployment.

Acknowledgment

The authors gratefully acknowledge funding from Innosuisse, the Swiss Innovation Agency, within the framework of an InnoCheque project. The online survey and data analysis were conducted in cooperation with the Swiss consulting firm CSP AG (www.csp.ch).

References

Argyris, C., & Schön, D. A. (1978). Organizational learning: A theory of action perspective. Addison-Wesley.

Bannister, F., & Connolly, R. (2014). ICT, public values and transformative government: A framework and programme for research. Government Information Quarterly, 31(1), 119–128. https://doi.org/10.1016/j.giq.2013.06.002

European Commission. (2020). A European strategy for data. Publications Office of the European Union. https://digital-strategy.ec.europa.eu/en/policies/strategy-data

European Commission. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. http://data.europa.eu/eli/reg/2024/1689/oj

Janssen, M., & van den Hoven, J. (2023). Governing AI in the public sector: Challenges and directions. Government Information Quarterly, 40(1), 101739. https://doi.org/10.1016/j.giq.2022.101739

Mergel, I. (2019). Digital service teams in government. Government Information Quarterly, 36(4), 101389. https://doi.org/10.1016/j.giq.2019.07.001

Mergel, I., Edelmann, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. Government Information Quarterly, 36(4), 101385. https://doi.org/10.1016/j.giq.2019.06.002

Mikalef, P., Gupta, M., Boura, M., & Krogstie, J. (2023). Big data analytics capabilities and AI adoption in the public sector. Information Systems Frontiers, 25, 123–140. https://doi.org/10.1007/s10796-021-10154-6

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of sustainable enterprise performance. Strategic Management Journal, 28(13), 1319–1350. https://doi.org/10.1002/smj.640

Tornatzky, L. G., & Fleischer, M. (1990). The processes of technological innovation. Lexington Books.

Vial, G. (2019). Understanding digital transformation: A review and research agenda. The Journal of Strategic Information Systems, 28(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector: Applications and challenges. International Journal of Public Administration, 42(7), 596–615. https://doi.org/10.1080/01900692.2018.1498103

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance. Government Information Quarterly, 38(3), 101577. https://doi.org/10.1016/j.giq.2021.101577

About the authors

Claudia Pedron is Professor of Business Information Systems at the Institute for Information and Process Management (IPM), OST – Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland. She has extensive experience in business process management, process automation, IT and management systems, and digitalization. She has acquired and led numerous research and service projects in process management, digital transformation, real estate technology, and decision-support systems. She holds a Ph.D. from ETH Zurich and an Executive Master in Business Engineering from the University of St. Gallen. Her publications address process-oriented decision support, digital services, e-health, and technology-supported management systems.

Hans-Dieter Zimmermann is a senior lecturer of information systems at the University of Applied Sciences Eastern Switzerland (OST)'s School of Management in St. Gallen. He earned his doctoral degree from the University of St. Gallen (HSG) with a dissertation on the information society and holds a degree in business administration from the University of Erlangen-Nuremberg.
Dr. Zimmermann's research focuses on the ways in which digitalization transforms communication and interaction across societies, markets, and organizations. His areas of expertise include electronic markets and platforms, digital business, digital democracy, and smart cities. He also serves as co-editor of the scientific journal Electronic Markets: The International Journal on Networked Business.

Matthias Baldauf is Professor of Business Informatics with a focus on Human-Computer Interaction at the Institute for Information and Process Management (IPM), OST – Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland. His research interests include human-centered AI, human-AI collaboration, and smart government. He has published his work in peer-reviewed journals and at leading international conferences and regularly serves as a reviewer for renowned academic journals and conferences. He holds a Ph.D. in Computer Science and an MSc in Software Engineering from TU Wien.

Summary

This study examines AI capability gaps in Swiss municipalities and cantonal administrations in German-speaking Switzerland. Based on a 2025 survey of respondents responsible for digitalization and/or AI-related initiatives, it analyzes digitalization, AI adoption, and implementation barriers. The findings show comparatively advanced digitalization in several operational and support-related domains, while AI adoption remains concentrated in assistive applications and is still limited in analytical and decision-support uses despite considerable planned implementation. The main barriers are legal uncertainty, limited expertise, resource constraints, and fragmented responsibilities rather than technological availability alone. Interpreted through AI adoption research, dynamic capabilities, and organizational learning, the findings point to a capability gap between existing digital systems and the organizational, legal, and governance capacities required for broader AI adoption.

Appendix A: Survey Instrument and Measurement Approach

This appendix provides a structured overview of the survey instrument used in this study. Table A1 summarizes the main sections (A-D) of the questionnaire, the key topics covered, and the corresponding measurement approach. The following subsections describe how the main constructs were operationalized and provide selected survey questions to illustrate the items used in the analysis.

Appendix A1: Structure of the questionnaire

Table A1: Overview of survey sections and measurement approach

	Topic	Key question	Measurements
A	Organizational context	Administrative level, size, function of respondent	Predefined categorical items (single response)
B	Internal administrative processes	Overall digitalization level, AI use, process-specific digitalization, digital systems and AI applications (current and planned)	Five-point Likert scales and predefined categorical multiple-response items
C	Externally oriented administrative services	Overall digitalization level, AI use, and digital or AI-supported services for external users (current and planned)	Five-point Likert scales and predefined categorical multiple-response items
D	Implementation barriers	Organizational, technical, legal, cultural, personnel-related, and legitimacy-related barriers	Predefined categorical multiple-response items

Source: Own Elaboration

Appendix A2: Measurement of key constructs

Digitalization

Digitalization was measured as the reported degree to which administrative services and processes are digitally supported. Respondents assessed digitalization in two contexts: internal administrative processes and externally oriented administrative services.

For internal processes, respondents assessed both current digitalization and perceived digitalization potential in domains such as leadership support, quality management, process management, HR, IT and internal communication, financial management, document handling, and procurement.

Digitalization was measured using five-point Likert scales: Very low, low, medium, high, very high.

In addition, the questionnaire captured the current and planned use of non-AI digital systems, including document management systems, ERP systems, workflow tools, cloud solutions, collaboration software, BI/data warehouse systems, e-learning platforms, knowledge management platforms, and IT ticketing systems.

AI adoption

AI adoption was operationalized as reported current organizational use and intended near-term implementation of AI-related applications. The questionnaire distinguished between internal administrative processes and externally oriented services.

For internal processes, respondents indicated whether selected AI applications were currently used, planned within the next 3–5 years, or not used/not planned. Application categories included generative AI, AI-assisted support systems, document and information analysis, pattern or anomaly detection, process analysis and automation, decision support, decision automation, and real-time AI applications.

For external services, the questionnaire captured digital and AI-supported services such as citizen apps, chatbots, online transactions, online forms, AI-supported form assistance, appointment booking, processing-time forecasting, automated plausibility checks, anomaly detection, feedback analysis, translation tools, and speech recognition.

Implementation barriers

Implementation barriers were measured through predefined items. Respondents could select multiple applicable barriers. The categories covered organizational and strategic barriers, technological and data-related barriers, legal and regulatory barriers, cultural and personnel-related barriers, and user- or legitimacy-related barriers.

Measurement note

Likert scales were used to assess the degree or intensity of digitalization and AI use. Predefined categorical multiple-response items were used to capture the presence, type, and planned implementation of digital systems, AI applications, external services, and implementation barriers.

Appendix A3: Selected survey questions

Organizational context

Respondents were asked to provide information on their role/function, administrative level, municipality/city or canton, and organization size.

Internal administrative processes

How do you assess the current level of digitalization of internal administrative processes in your organization?

☐ Very low ☐ Low ☐ Medium ☐ High ☐ Very high

How do you assess the current use of AI solutions for supporting internal administrative processes in your organization?

☐ Very low ☐ Low ☐ Medium ☐ High ☐ Very high

Respondents then assessed the level of digitalization and the potential for further digitalization in selected internal process domains.

Digital systems for internal processes

Which technologies are currently used or planned for supporting internal administrative processes in your organization?

Table A2a: Current and planned use of non-AI digital systems in internal administrative processes

	Current use	Planned within 3-5 years	Not used / Not planned
A	☐	☐	☐
B	☐	☐	☐
C	☐	☐	☐
D	☐	☐	☐
E	☐	☐	☐
F	☐	☐	☐
G	☐	☐	☐
H	☐	☐	☐

Source: Own Elaboration

Legend:

A: Document management / records management; B: Collaboration software; C: Cloud solutions; D: ERP systems; E: Workflow management tools; F: Robotic process automation; G: BI / data warehouse systems; H: Knowledge management / IT support systems;

AI applications for internal processes

Which AI-based applications are currently used or planned for supporting internal administrative processes?

Table A2b: Current and planned use of AI-based applications in internal administrative services

	Current use	Planned within 3-5 years	Not used / Not planned
A	☐	☐	☐
B	☐	☐	☐
C	☐	☐	☐
D	☐	☐	☐
E	☐	☐	☐
F	☐	☐	☐
G	☐	☐	☐
H	☐	☐	☐

Source: Own Elaboration

Legend:
A: AI-assisted support systems; B: Document and information analysis; C: Pattern/anomaly detection and prediction; D: Process analysis or automation; E: Decision support or automation; F: Real-time AI applications.

Externally oriented administrative services

How do you assess the current level of digitalization in the provision of services for citizens, companies, and organizations?

☐ Very low ☐ Low ☐ Medium ☐ High ☐ Very high

How do you assess the current use of AI solutions in the provision of services for citizens, companies, and organizations?

☐ Very low ☐ Low ☐ Medium ☐ High ☐ Very high

Which digital and/or AI-supported services does your organization currently provide or plan to provide?

Examples included online forms and applications, online transactions, online payments, digital document ordering, register access, appointment booking, citizen

apps, chatbots, AI-supported form assistance, processing-time forecasting, anomaly detection, feedback analysis, translation tools, and speech recognition.

Implementation barriers

Which challenges do you experience or expect in the implementation of digitalization and AI projects in your organization?

Multiple answers were possible.

Examples included lack of personnel or budget, fragmented responsibilities, lack of strategy, lack of competencies, legacy systems, interoperability problems, data protection requirements, legal uncertainty regarding AI, resistance to change, lack of digital or AI skills, unclear user needs, low societal acceptance, and lack of participation or transparency in AI use.

Fifty Shades of Orange: Construal Levels, Cognitive Biases, and The Amber Dilemma in Cybersecurity Risk Communication

Gulet Barre,[1] Dinh Khoi Nguyen,[2] Tim Huygh,[1] Arno Nuijten[1, 3, 4]

[1] Open University of the Netherlands, Faculty of Science, Limburg, the Netherlands
gulet.barre@ou.nl, tim.huygh@ou.nl, arno.nuijten@ou.nl
[2] University of Jyväskylä, Faculty of Information Technology, Jyväskylä, Finland
khoi.d.nguyen@jyu.fi
[3] Erasmus University Rotterdam, Erasmus School of Accounting and Assurance, Rotterdam, the Netherlands
arno.nuijten@ou.nl
[4] University of Pretoria, Faculty of Economic and Management Sciences, Department of Auditing, Pretoria, South Africa
arno.nuijten@ou.nl

DOI
https://doi.org/10.18690/um.fov.4.2026.10

ISBN
978-961-299-152-4

Board-level decisions on cybersecurity risks are strongly shaped by how information is communicated. Traffic light reporting offers simplicity, but especially the amber light leads to inconsistent interpretations. This paper applies Construal Level Theory (CLT) to explore how psychological distance structures risk responses. Based on interviews with 11 CISOs, we focus on four exemplary biases arising from psychological distance in board-level risk perception. Our results indicate that these biases are not random flaws in judgement but patterned outcomes of directors' and executives' construal levels, influencing whether risks are seen in abstract or concrete terms. Our contribution lies in applying CLT to cybersecurity governance, demonstrating how psychological distance shapes board-level biases, and highlighting how CISOs can manage psychological distance through framing and narrative to support more balanced decision-making. Future work will conduct additional CISO interviews and design an experiment to test how psychological distance shapes board responses to amber risk signals.

Keywords:
cybersecurity communication, cognitive biases, psychological distance, construal level theory, cybersecurity governance, risk reporting

University of Maribor Press

1 Introduction

In the domain of cybersecurity governance, decision-making at the board level often depends on how risks are framed, communicated, and interpreted (Noh et al., 2025; Turel & Bart, 2014). More broadly, recent research highlights that organisational cybersecurity increasingly depends on behavioural and human-centred factors beyond purely technical controls (Barre, 2025). A widely used method for simplifying complex information, such as information about cybersecurity in an organization, is traffic light reporting, in which risks are color-coded as red (indicating high-risk domains), green (indicating low to no-risk domains), or amber (indicating domains whose risk levels are neither high nor low). While red and green are relatively clear, the amber light is more ambiguous, as it indicates that a risk is neither clearly urgent nor fully absent. This ambiguity creates a recurring interpretive tension in board-level decision-making. Construal Level Theory (CLT), developed by Trope and Liberman (2002, 2010), explains how people mentally interpret events at different levels of abstraction depending on psychological distance, that is, how far an event is perceived in time (temporal distance), space (physical distance), likelihood (hypothetical distance), or social connection (social distance). Events perceived as distant are processed more abstractly focusing on why they matter in terms of desirability, whereas events perceived as close are processed more concretely, focusing on how they can be implemented in terms of feasibility. These differences in construal affect not only how risks are understood but also lead to predictable distortions in judgement, known as cognitive biases (Kahneman, 2011).

In the context of cybersecurity governance, these dynamics are particularly important because board members operate at a greater psychological distance from technical operations than CISOs. Boards do not see incident logs, hear the SOC alerts, or manage the patch cycles. Instead, they rely on summaries, dashboards, and interpretations of security operations provided by CISOs. Further, prior research on information security investment demonstrates that board decision-making processes are heavily biased by psychological factors (Dor & Elovici, 2016). However, previous research has paid little attention to how systematic differences in risk interpretation occur, or how they lead to bias in board decision-making.

Our study focuses on four exemplary biases observed in interviews with CISOs, conducted to shed light on the communication of cyber risk to the executive level: optimism bias, pessimism bias, ambiguity bias, and confirmation bias. Each of these can be linked to variations in psychological distance. By applying CLT as a lens, we argue that these biases are not random flaws, but structured patterns shaped by the mental distance between the board and the reality of cyber threats.

This paper contributes to the literature on board-level cybersecurity decision-making by leveraging CLT as a theoretical lens for understanding cognitive biases in cybersecurity governance. Our first contribution is theoretical: we extend the application of CLT into the cybersecurity domain and show that psychological distance systematically structures biases in board decision-making. Second, our empirical contribution lies in analyzing interviews with 11 CISOs, which provide insights into how boards perceive and respond to cyber risks from the perspective of CISOs.

The central research question guiding this study is therefore: How do psychological distance and construal levels contribute to the presence of cognitive biases in board-level cybersecurity risk decisions? By answering this question, we contribute to both theory and practice, expanding CLT into a new organizational domain and offering practical insights for how CISOs can better tailor their communication to reduce bias and improve decision-making at the highest level.

2 Theoretical Background

Construal level refers to the degree of abstraction in how people mentally represent objects, events, or actions. Based on categorization theory (Rosch, 1975) and concept formation (Medin & Smith, 1984), high-level construals are abstract, decontextualized, and focused on goals or meaning, whereas low-level construals are concrete, detailed, and context specific. This distinction helps explain how individuals interpret and act on information differently depending on psychological distance, organizational role, and goals.

In the context of cybersecurity governance, these theoretical insights clarify why board members may perceive, communicate, and respond to cyber risks differently from other stakeholders. People employ higher levels of construal when dealing with

psychologically distant targets (Trope & Liberman, 2010). Psychological distance is a subjective construct, typically defined along four dimensions: temporal, spatial, social and hypothetical. Temporal distance refers to how far in time an event feels, with distant events perceived as less urgent and more abstract. Spatial distance is about how physically far a threat or event feels, making distant events seem less relevant. Social distance refers to emotional or hierarchical separation, with greater distance reducing a sense of personal responsibility. Hypothetical distance relates to how real or likely a situation feels, where uncertain scenarios are often underestimated or ignored (Wakslak et al. 2006).

Besides psychological distance, earlier CLT research has identified several other elements that characterize the effects of high- and low-level construal: desirability & feasibility, hypotheticality, category & exemplar, and pros & cons. These elements are widely recognized as fundamental components of CLT (Trope & Liberman, 2010).

People incorporate both desirability and feasibility considerations when evaluating future actions. According to CLT, actions may be construed in high-level terms, which link them to overarching purposes, or in low-level terms, which connect them to the concrete means of implementation. At higher level of construal, desirability tends to dominate, whereas at lower levels of construal, feasibility becomes more salient. Another dimension is hypotheticality, which concerns the perceived likelihood of future events (Armor and Sackett 2006). CLT suggests that improbable events are construed in abstract, high-level terms, while probable events are construed in concrete, low-level terms. In other words, construal levels become more concrete as the probability of a future event increases. Additionally, category & example addresses the way people mentally represent objects, actions, or events. Under abstract, high-level construal, individuals group these into broader categories, attending to commonalities across instances. Under concrete, low-level construal, they rely on narrower, more specific exemplars that emphasize uniqueness (Liberman et al. 2002). Finally, pros & cons refer to the evaluation of future action based on arguments for or against taking them. At higher levels of construal, positive arguments (pros) tend to become more salient, while at lower levels of construal, negative arguments (cons) are more prominent (Eyal et al. 2004).

Building on the literature on CLT and decision biases in cybersecurity governance, we propose a conceptual model in Figure 1 that links psychological distance to board-level decision-making in the context of cybersecurity governance. The model illustrates that psychological distance shapes the emergence of security-related biases, which subsequently influences how boards evaluate and decide on cybersecurity issues. Therefore, CLT provides a suitable lens to understand how psychological distance contributes to biased risk perception and decision-making at the board level.

Figure 1: Conceptual model

3 Methodology

In an earlier published paper, we conducted a literature review of biases relevant to cybersecurity decision-making (Barre et al., 2025). The literature review resulted in a longlist of decision-making biases relevant to cybersecurity governance. Subsequently, the perceived prevalence of these biases in practice was determined based on interviews with CISOs. Optimism bias was acknowledged by 8 of the 11 CISOs, pessimism bias by 7, ambiguity bias by 10, and confirmation bias by 7. We focus on these four most frequently mentioned biases, while acknowledging that other additional, less prominent biases were also identified during the interviews.

Prior to the main data collection, we conducted a pilot interview with a CISO. The purpose of this pilot was to test the clarity and structure of the interview questions, to assess whether the questions could be addressed within the planned time and to ensure that the terminology was well understood by participants. The pilot interview was not included in the final analysis but served to refine the interview protocol and strengthen the validity of the following interviews.

We conducted 11 semi-structured interviews with CISOs from across Europe. The interview protocol was developed based on existing literature on decision-making biases. Each bias was described in a neutral and non-leading way to avoid influencing responses. We applied the principle of data saturation to determine our stopping point (Guest et al., 2020). Each interview was analyzed for presence of new biases or new associations between biases and one or more CLT dimensions. Data saturation was achieved when the 10th and 11th interviews showed no new categories or linkages, thereby meeting the predefined stop-criterion. Interviews were conducted online and lasted approximately 45 minutes on average.

The interviews offered a balanced perspective from both local and international contexts. 7 CISOs from the Netherlands, 3 CISOs from Belgium and 1 CISO from the United Kingdom. All participants had senior cybersecurity roles in their organizations and had direct involvement in strategic decisions and communications with boards. The majority (10 out of 11) worked in large organizations (250+ employees), while 1 CISO represented a small enterprise with fewer than 50 employees. Professional experience also varied, ensuring perspectives from both seasoned and relatively new CISOs. 3 CISOs have 1-3 years of experience, 2 CISOs have 4-6 years of experience and 6 CISOs have 7+ years of experience. The gender distribution among participants included 3 females and 8 males. The interview protocol covered different components: participants' professional background and experience, their typical ways of reporting cybersecurity risks and communicating with executives, perceived obstacles and complexities in cyber risk communication and examples of how decision biases may manifest in practice. This structure enabled us to link the interview findings to dimensions of CLT.

4 Empirical Illustration of Biases Through the Lens of CLT

In Table 1 below, the rows represent the biases and reporting behavior that CISOs were asked to reflect upon (e.g., optimism bias, pessimism bias, ambiguity bias, confirmation bias and generic amber reporting). The columns represent the different dimensions of psychological distance. If a cell includes a reference such as "CISO#7", this indicates that this particular CISO provides an anecdote that reflects both the bias/reporting theme of that row and the dimension of psychological distance in that column. The table demonstrates that 9 out of 11 CISO provided

anecdotes that contained at least one form of psychological distance when discussing biases and reporting in practice.

Table 2: Mapping psychological distance across biases and reporting behaviours identified by CISOs

Bias	Spatial	Social	Hypothetic	Temporal
Optimism bias		CISO#11	CISO#6	CISO#1
Pessimism bias		CISOs#3; #10	CISOs#5; #10	CISOs#3; #5
Ambiguity bias			CISOs#6; #9	CISO#7
Confirmation bias		CISOs#7; #8	CISOs#7; #8; #9	
Generic (amber) reporting	CISOs#6; #9	CISO#10	CISOs#8; #11	CISOs#6; #7

4.1 Optimism bias

Optimism bias is defined as the tendency to underestimate the likelihood of unfavorable outcomes (Legoux et al., 2014). CISO#1 shared, "They (Board members) ask: 'Why prepare for what hasn't happened yet?'" This illustrates temporal distance. Because board members conceive higher temporal distance, their attention shifts towards broad strategic benefits rather than assessing whether a threat is realistic or urgent. As a result, security investments are forced to compete with other strategic initiatives that appear more immediate and their benefits are perceived more certain than those of abstract cybersecurity threats. Yet, in terms of resilience, trust, and sustained value creation, security investment can be recognized as strategic assets that strengthen organizational competitiveness rather than as purely defensive costs (Kwon & Johnson, 2014).

Similarly, CISO#6 stated, "Because we (the company) passed the audit, they (board members) said: 'I guess we're solid', pointing to hypothetical distance. Passing the audit made the board feel safe, as the risk had already been dealt with. That success made the board perceive the problem as less likely to recur and alleviate the board's uncertainty. While they might not intentionally ignore the risk, the sense of distance made them feel more secure than they probably should have.

In another example, CISO#11 reported: "They (board members) believe our external provider handles everything, so they don't worry much." By shifting responsibility to an external party, the threat appears less personally relevant to the board. Social distance rooted in CLT helps explain how people focus less on whether the threat is still relevant to them.

Taken together, optimism bias increases with greater psychological distance. Temporal distance, when risks are perceived as lying further in the future, and hypothetical distance, when risks are perceived as less likely, are both associated with optimism bias in board level perception.

4.2 Pessimism bias

Pessimism bias is defined as the tendency to overestimate negative outcomes and perceive risks as more severe than they actually are (Snow et al., 2007). Sometimes the board sees threats as too close, which makes them overly negative or even hopeless. CISO#3 said, "After observing a peer organization fall victim to ransomware, a board member remarked: 'It's just a matter of time for us.'" Because the affected organization is familiar, the social distance was reduced and the board suddenly recognized their exposure to the same risk. Consequently, the risk becomes urgent and unavoidable.

Another CISO#10 noted: "Following the Log4j incident, they (board members) were prepared to terminate all new technology projects." In this scenario, reduced hypothetical distance made a general threat appear concrete, while reduced social distance led the board to perceive the risk as directly relevant to their organization. As a result, they reacted with immediate measures, even if those responses were not always proportionate.

CISO#5 mentioned, "We'll never be secure, no matter what we do." This shows how perceiving the threat as imminent (reduced temporal distance) and almost certain to occur (reduce hypothetical distance) can foster extreme pessimism, making any effort seem pointless.

In short, pessimism bias happens when the threat feels psychologically close, either socially, in time, or in likelihood. That closeness makes board members zoom in on the risk and lose sight of the broader context. It can lead to fear-based decisions, overreactions, or even a disengagement from longer-term security initiatives. From the perspective of CLT, pessimism bias arises because of reduced psychological distance leads boards to think in concrete terms, making immediate threats feel more pressing while pushing broader and longer-term considerations into the background.

4.3 Ambiguity bias

Ambiguity bias is defined as the tendency to avoid or undervalue options where information is conflicting, incomplete, or uncertain (Bhandari et al., 2008). In this context, such ambiguity is often represented by 'amber risks' positioned between the clear danger of red and the relative safety of green. Boards only react to well-defined threats and often avoid those they do not fully understand. CISO#7 observed: "The board shows a tendency to postpone the discussion of 'amber risks' time after time." Such procrastination increases the temporal distance, which allows the issue to remain vague and less urgent.

CISO#6 said: "If I can't reduce it to traffic-light colors, they push it to the side." This demonstrates how visual clarity can reduce hypothetical distance and make engagement more likely. Conversely, risks without such framing remain psychologically remote and are deprioritized. A third example came from CISO#9 who recalled: "They (board members) don't like grey areas. They want yes/no answers." Here, the absence of certainty increases hypothetical distance, when risks are presented in shades of gray, the board perceives them as not concrete enough to warrant attention.

In short, ambiguity bias arises when unclear or complex risks get ignored, not because they're unreal, but because they're too messy to evaluate. If something feels too abstract, the board avoids it. Boards prefer simple over nuanced interpretations, which can result in blind spots in how emerging threats are evaluated and addressed.

4.4 Confirmation bias

Confirmation bias is defined as the tendency to favor information that supports pre-existing beliefs or hypothesis while disregarding or undervaluing contradictory evidence (Snow et al., 2007). Boards do not only seek clarity, but they also prefer familiarity. CISO#7 noted: "They always ask: 'Is this GDPR-related?' If not, they switch off." Here perceived relevance is framed through compliance such that issues linked to GDPR are seen as close or urgent, whereas other risks feel distant and are overlooked.

CISO#9 observed: "They want to see the same phishing report every month, even if it's no longer relevant." Because phishing is familiar, it feels close and manageable. Even when actual risks have shifted, the board remains focused on familiarity that reduces hypothetical distance and provides a sense of control.

CISO#8 noted, "They only engage when incidents matching what they already worry about" If a threat does not connect with past experiences or existing concerns, it is construed as abstract or unrelatable, showing increased social- and hypothetical distance and therefore receives little attention.

In short, confirmation bias shows up when boards stick to the risks they already recognize. Risks outside this frame are perceived as distant or difficult to interpret. As a result, emerging threats can be overlooked, simply because they don't fit the mental checklist the board already holds.

5 Psychological Distance and the Amber Light

Based on the logic of CLT and the way psychological distance shapes optimism, pessimism, ambiguity, and confirmation bias, we can also better understand how traffic light reporting is interpreted in board contexts. In our interviews, all eleven CISOs confirmed that they use traffic light reporting as their primary way of communicating cybersecurity risk to boards. The amber indicator is particularly revealing, as it embodies ambiguity: unlike green (no issues) or red (critical risk), amber lacks a clear interpretation and leaves room for subjective judgement. This ambiguity means that responses can vary across individual boards members, depending on their experience, risk tolerance, or intuition. In this way, the amber

light is not merely a reporting mechanism but a mechanism through which the role of psychological distance and cognitive bias in decision-making becomes visible.

Building on this perspective, our interviews further reveal practices and rationales behind the use of traffic light reporting. CISOs emphasized its clarity, visual simplicity, and often efficiency in communicating complex issues, as well as its alignment with the board's preference for abstraction and simplification.

From a temporal perspective, several CISOs observed that boards tend to postpone the discussion of amber risks. As CISO#7 explained: "The board shows a tendency to postpone the discussion of amber risks time after time." CISO#6 mentioned that when traffic lights remain unchanged, always green, amber, or red, board members often ask: "What do I do to change that". Social distance plays a similar role in shaping how amber is perceived. CISOs emphasized that board members often ignore ambiguous risks until an incident affects a peer organization. CISO#11 noticed that after a DDoS attack, the board immediately asked, "Does this also apply to us?" CISO#11 uses high-profile cases such as "NotPetya" in Rotterdam to make risks feel closer: "Then they say: okay, that's our ecosystem, that could be us."

The hypothetical dimension also plays a crucial role in understanding amber. Several CISOs mentioned that green and red are straightforward, but amber causes discussion. At times, reports are presented in abstract or generalized terms, which may unintentionally increase psychological distance and reduce engagement. In other instances, CISOs deliberately add detail, specificity, or urgency to lower the construal level, thereby making the risk appear more applicable. This illustrates how communicative alternatives can actively raise or lower psychological distance, aligning message framing with intended board behavior. Notably, narratives and scenarios are also used by some CISOs to reduce hypothetical distance. By describing real-life incidents, sector-specific breaches, or near-miss events, they aim to decrease abstraction and increase perceived feasibility. CISO#8 explained that "If you focus on the amber light, I think that's where my experience as the CISO has to come into play, because I can't offload my risk register to the board". Spatial distance influences how boards perceive and sometimes downplay risks associated with third parties. As CISO#9 explained: "If it's amber because of a supplier, they'll say it is outside their scope." CISO#6 noted: "We are already preparing, but the boards perceive it as far away. Until it turns red, it stays abstract for them."

The reflection above aligns with Eling et al. (2021), who emphasize the role of risk analysis stage in restructuring ambiguity. By systematically assessing cyber risks as a combination of probability and impact, CISOs can translate "amber" signals into more concrete evaluations that boards can act upon. Taken together, these patterns show that the way CISOs report to boards reflects not only their strategic priorities but also an active management of psychological distance. By varying abstraction, detail, and narrative, they influence how risks are construed and ultimately, how decisions regarding cybersecurity are made. In summary, the present research suggests that 'amber' is not a neutral middle state but a psychological distance-dependent signal: psychologically distant amber tends to be interpreted as almost green, whereas psychologically close amber is treated as almost red. This dynamic character of amber has been largely overlooked in cyber risk reporting research to date.

6 Discussion, Contribution, and Future Research

This study has explored how psychological distance and construal levels shape the way board members interpret and act upon cybersecurity, based on insights from CISOs. By applying CLT as a lens, we show that cognitive biases regularly observed in the boardroom are not random flaws in judgement but rather predictable consequences of distance-driven mental framing.

The potential contribution lies in extending CLT into the domain of cybersecurity governance. Our initial findings show that psychological distance provides a systematic explanation for how boards construe cybersecurity risks and why certain biases emerge in a governance context. Our work positions psychological distance as a central mechanism structuring board-level decision-making.

A central paradox emerging from our findings is that both too much distance and too little distance can distort decision-making. Optimism bias emerges when threats feel far away, but pessimism bias arises when they feel uncomfortable close. This raises the question of whether the goal should be to minimize psychological distance or rather to balance it. CISOs may need to adjust the way they present risks, keeping reports high-level when the boards need to discuss strategy, but making them more concrete when action is required.

The findings in this paper showcase an important first step by demonstrating how CLT offers a structured explanation for biases in cyber risk reporting. These insights highlight that psychological distance is not just a background factor but a central mechanism that influences how board members interpret and act upon cybersecurity risks. Several limitations should be acknowledged when interpreting these findings. The study is based on a small convenience sample of eleven CISOs, limiting generalisability, and is skewed towards participants from the Netherlands and Belgium and predominantly large organizations. In addition, the reliance on single-informant self-reported data from CISOs may introduce selection and social desirability biases. Finally, the qualitative design is exploratory in nature and does not permit causal claims regarding the relationship between psychological distance and board-level decision-making.

Conceptually, our study extends work on cyber risk communication, particularly the "amber light" category, by showing that this middle category is not interpreted as neutral but rather shifts depending on how psychological distance is framed. When risks are communicated as *distant in time or as unlikely to occur*, amber tends to be perceived as not yet urgent (i.e., the interpretation moving toward the green light), whereas when risks feel closer or more personally relevant, amber can be perceived as more concerning (i.e., the interpretation moving toward the red light). This insight positions the amber light as a dynamic communication point where psychological distance framing shapes how boards construe risks and decide whether to act.

As the next step, we will conduct additional interviews with CISOs to further refine these propositions and explore additional aspects of CLT, such as desirability and feasibility. A later stage of the research will empirically test these propositions through an experiment that systematically manipulates psychological distance and construal levels. Building on our exploratory findings, the experiment will contrast situations in which the CISO presents risks in a high-level construal, abstract way (focusing on strategic relevance) versus a low-level construal, concrete way (emphasizing technical details). This design may reveal the tension we observed, i.e., boards typically operate at higher construal levels, while CISOs often frame risks more concretely. Just as CISOs can adjust psychological distance in their reports to influence how boards construe risks, the experiment will test how different framing in reporting strategies shift boards' construal levels and shape cybersecurity decisions. In doing so, we aim to clarify whether and how adjusting psychological

distance through adaptive communication between CISOs and boards is part of effective cybersecurity governance.

References

Armor, D. A., & Sackett, A. M. (2006). Accuracy, error, and bias in predictions for real versus hypothetical events. *Journal of Personality and Social Psychology, 91*(4), 583-600.

Barre, G. (2025). From Breakers to Builders: The Role of Bug Bounty Hunters in Strengthening Organizational Cybersecurity. *Information*, *16*(3), 209.

Barre, G., Huygh, T., Nguyen, D. K., & Nuijten, A. (2025). Towards understanding cognitive biases in cybersecurity governance. *Proc. 38th Bled eConference*, 737-744.

Dor, D., & Elovici, Y. (2016). A model of the information security investment decision-making process. *Computers & Security, 63*, 1-13.

Eling, M., McShane, M., & Nguyen, T. (2021). Cyber risk management: History and future research directions. *Risk Management and Insurance Review, 24*(1), 93-125.

Eyal, T., Liberman, N., Trope, Y., & Walther, E. (2004). The pros and cons of temporally near and distant action. *Journal of Personality and Social Psychology, 86*(6), 781-795.

Guest, G., Namey, E., & Chen, M. (2020). A simple method to assess and report thematic saturation in qualitative research. *PloS one, 15*(5), 2-6.

Kahneman, D. (2011). Fast and slow thinking. *Allen Lane and Penguin Books, New York, 2*.

Kwon, J., & Johnson, M. E. (2014). Proactive versus reactive security investments in the healthcare sector. *Mis Quarterly, 38*(2), 457-467.

Legoux, R., Leger, P. M., Robert, J., & Boyer, M. (2014). Confirmation biases in the financial analysis of IT investments. *Journal of the Association for Information Systems, 15*(1), 33-52.

Liberman, N., Sagristano, M. D., & Trope, Y. (2002). The effect of temporal distance on level of mental construal. *Journal of Experimental Social Psychology, 38*(6), 523-534.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35*(1), 113-138.

Noh, H. H., Rim, H. B., & Lee, B. K. (2025). Risk preferences in decision-making: A construal level perspective. *Acta Psychologica, 252,* 2-4.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*(3), 192-233.

Snow, A. P., Keil, M., & Wallace, L. (2007). The effects of optimistic and pessimistic biasing on software project status reporting. *Information & Management, 44*(2), 130-141.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440-463.

Turel, O., & Bart, C. (2014). Board-level IT governance and organizational performance. European *Journal of Information Systems, 23*(2), 223-239.

Wakslak, C. J., Trope, Y., Liberman, N., & Alony, R. (2006). Seeing the forest when entry is unlikely: probability and the mental representation of events. *Journal of Experimental Psychology: General, 135*(4), 641-653.

Exploring the Relation between Information Security Policy Compliance and Cyber Resilience

Abigaïl de Rijp, Tim Huygh, Clara Maathuis, Rogier van de Wetering, Harald Vranken

Open Universiteit, Faculty of Science, Heerlen, the Netherlands
abigail.derijp@ou.nl, tim.huygh@ou.nl, clara.maathuis@ou.nl, rogier.vandewetering@ou.nl, harald.vranken@ou.nl

Contemporary digital environments necessitate an increased organizational commitment to safeguarding the confidentiality, integrity and availability of information assets. An indispensable element of this endeavor lies in issuing and enforcing Information Security Policies (ISP) that delineate appropriate handling and use of the organization's information assets. While prior research predominantly examined the impact of ISPs on employee cyber behavior, limited attention was given to the organizational-level effects of ISP compliance (ISPC). In particular, the extent to which ISPC and cyber resilience (CR) are interrelated is still insufficiently understood. In light thereof, this article explores mechanisms that can be implemented to increase ISPC, the general organizational-level effects of ISPC, and how ISPC relates to CR specifically. Drawing on an integrative literature review and a holistic case study, this article explores the role of ISPC in driving organizational CR.

DOI
https://doi.org/10.18690/um.fov.4.2026.11

ISBN
978-961-299-152-4

Keywords:
information security management, information security policy, compliance, AMC-framework, cyber resilience

University of Maribor Press

1 Introduction

A characteristic of the 21st century is the interwovenness of the digital and physical world, making online shopping, social media and digitalized business processes increasingly ordinary. This coalescence also provokes the transition of traditionally unwanted physical behavior to its digital counterpart. On this matter, the following applies: "But people themselves alter so much, that there is something new to be observed in them forever." – Jane Austen, Pride & Prejudice. Even though romance novels and an Information Security Management System (ISMS) are not necessarily relatable, both address the unpredictability of human behavior. In fact, within an ISMS, many security weaknesses arise from human behavior and actions (Ncubukezi, 2022). Thus, enforcing an Information Security Policy (ISP) can aid organizations in protecting their information assets (Delso-Vicente et al., 2025). An ISP is "a set of formalized procedures, guidelines, roles and responsibilities to which employees are required to adhere to safeguard and use properly the information and technology resources of their organizations" (Cram et al. 2019, pp. 525-526).

Even though there is a plethora of research and theories on the impact of ISPs on employee behavior (e.g., Xue et al., 2021; Mubarkoot et al., 2023), there seems to be a lack of a comprehensive account of the impact of ISP compliance (ISPC) on organization security (Cram et al., 2019). For instance, previous research focuses on the individual drivers of ISPC (Aggarwal et al, 2024; Balagopal et al., 2024; Sharma, et al. 2025). However, evaluation of ISPC necessitates consideration of both human behavioral as well as organizational factors (Alias et al., 2019; Zhang et al., 2025). This is particularly important as information security threats do not single out the employees whose non-compliant behavior might lead to an incident, but they affect the organization in its entirety (Delso-Vicente et al., 2025).

As seemingly critical elements of an organization's ISMS, pre-emptive planning and preparation for potential cyber-attacks requires a certain amount of cyber resilience (Calder, 2020; Alhidaifi et al., 2025). To this end, Ross et al. (2021, p. 1) describe cyber resilience (CR) as "The ability to anticipate, withstand, recover from, and adapt to adverse conditions, stresses, attacks, or compromises on systems that use or are enabled by cyber resources". Subsequently, this article addresses the relation and potential overlap between ISPC and CR. The objective of this paper is bifold. Firstly, it provides an understanding of which organizational mechanisms enable ISPC and

CR. Secondly, it addresses the extent to which ISPC (effects) relate to CR. To successfully achieve these objectives, the main research question is: *Which organizational drivers can enhance Information Security Policy Compliance and Cyber Resilience, and how are these phenomena related?* To answer this question, ISPC and CR are first assessed separately to avoid the assumption that ISPC always leads to CR. The research starts off with an integrative literature review that provides conceptual insight into ISPC and CR found in the literature. The resulting Theoretical Background is provided in section 2. Then, the empirical research – which is discussed in sections 3 and 4 – investigates the existence of the organizational impact of ISPC and its potential overlap with CR. Section 5 discusses the research results and addresses its limitations and aspects for further investigation, and section 6 concludes the article in its entirety.

2 Theoretical Background

2.1 Literature Review

2.1.1 ISPC-enhancing drivers

Literature on ISPs highlights a broad range of organizational drivers that shape how policies are enacted in practice. Albeit obvious, these policies are only effective if they are complied with. Additionally, poorly aligned or rigid policies may fail to deliver intended outcomes (Almuqrin, 2024; Zanke et al., 2024). In light hereof, many authors distinguish various behavioral theories that focus on individual attitudes and behavior relating to ISPC. However, as mentioned above, non-compliant behavior may affect the entire organization, not just the individual transgressor. To address this theoretical gap, this article identifies organizational-level drivers for ISPC by utilizing the awareness-motivation-capability (AMC) framework by Chen et al. (2016). This framework emphasizes an organization's proactive decision-making through a distinction between awareness-, motivation- and capability-enhancing measures, each addressing a different (cognitive) level for the end-user (Dong et al. 2023; Li et al., 2025). Thus, the framework distinguishes directional and behavioral variance, and allows for integration of behavioral, institutional, and resource perspectives. In doing so, it is particularly well suited for ISPC research because it accounts for both intentional and unintentional non-

compliance. Table 1 provides a non-exhaustive synthesis of most relevant measures, based on the following criteria:

- Awareness-enhancing: relates to employees' knowledge of the ISP; these drivers were selected based on reach, frequency, and ability to convey ISP's rationale.
- Motivation-enhancing: relates to reasons why employees would or would not comply with the ISP; these drivers were selected through assessing organizational/ environmental engagement, and perceived value of compliance.
- Capability-enhancing: relates to ease or difficulty of ISPC; these drivers were selected based on usability and ability of enabling performance of duties.

Table 1: AMC-measures for ISPC

Awareness measures	Motivation measures	Capability measure
Onboarding can introduce new employees to the ISP(C) (Vollenweider, 2023).	A **cyber security culture** aids the increase of ISPC (Amankwa et al., 2018; Alassaf et al., 2021; Zhang et al., 2025).	**Infrastructural constraints** can prevent execution of non-compliant behavior (Khando et al., 2021); (Mubarkoot et al., 2023).
Involving end-users in the development of the ISP leads to their active participation in and awareness of the organization's ISMS/ ISP (Amankwa et al., 2018).	**Involvement of managers** (Moody et al., 2018), and their **exemplary ISPC behavior** incentivize compliance by other staff (Whitman et al., 2018; Zhang et al., 2025).	**Monitoring and auditing** enable identification of deviant behavior (D'Arcy et al., 2019), highlighting areas for ISP(C) improvement (Alassaf et al., 2021, Mubarkoot et al., 2023).
Awareness campaigns aim to inform end-users about the existence, scope, and value of the ISP (Bagheri, 2020; Khando et al., 2021).	**Managerial control through sanctions and promotions** can incentivize motivation for complying with the ISP (Cram et al., 2019; Khando et al., 2021).	**Allocation of IS-related tasks and responsibilities** enables employee participation in IS-practices, enriching their self-efficacy (Moody et al., 2018; Ali et al., 2021).
Clear and accessible ISP provisions increase employee compliance with, and engagement in, security practices (Ali et al., 2021).	**Strategies for organizational ISP justice** prevent IS-related injustices that could demotivate end-users to comply with the ISP (Hovav et al., 2016; Ali et al., 2021).	**Cybersecurity exercises** like penetration testing (Balakrishnan et al., 2018) enable vulnerability- and incident preparedness and response (Bartes et al., 2017).
Gamification simplifies communication of ISPC content and importance (Khando et al., 2021).	Promotion, support, and liaising of ISMS activities by an **ISMS coordination team** (Zammani et al., 2019) can enhance motivation.	**Education** cultivates skills and confidence (Cram et al. 2019), and **training** enhances application of taught skills (Tsohou et al., 2018).
Security Education, Training and Awareness **(SETA)-Programs** facilitate knowledge of procedures, guidelines and actionable tasks (Moody et al., 2018; D'Arcy et al., 2019).	Giving employees **freedom** and control to make IS-related decisions can reduce non-compliant behavior, as restriction herein incentivizes noncompliance (Hovav et al., 2016; Stewart, 2021).	Strategically **designed** systems that promote compliant, **user-friendly**, easily comprehensible **interfaces**, facilitate users' understanding and adoption of compliant behavior (Falahati, et al. 2022).

Notably, the AMC-framework provides a diagnostic lens for addressing barriers that might cause non-compliance and subsequently identify necessary adjustments. However, the AMC-framework does not explicitly account for organizational ISPC effects, thereby requiring additional research that will highlight its relation to CR.

2.1.2 CR-enhancing drivers

Table 2: Organizational drivers for CR

Planning ***Parameter:*** knowledge and decision-making prior to an incident	Cyber resilience **awareness and training** contribute to long-term enhancement of organizational cyber resilience (Kott et al., 2019; Bagheri, 2020; Annarelli, 2021).
	A **cyber security culture** can facilitate cyber resilience planning through integration of cyber security practices in the business operations (Kott et al., 2019), emphasizing the crucial role of management and leadership (Bagheri, 2020).
	A **Response and recovery plan** is crucial for the efficacy and sustainability of a resilience program. Through established policies (pre-event), these programs prescribe the organization's approach for handling threats and incidents (Bagheri, 2020; Staves et al., 2022), as well as **appropriate roles and responsibilities** (Annarelli et al., 2021).
Detecting ***Parameter:*** recognition and interpretation of anomalies and incidents	**Governance and monitoring programs** facilitate structural documentation of CR procedures and issues (Bagheri, 2020; Petrenko, 2022), enabling identification of gaps and areas for improvement (Rieger, et al., 2021). Expansive analysis by for e.g. Artificial Intelligence can aid quick and effective detection and response (Balakrishnan et al., 2018; Klioskli et al., 2023).
	Cybersecurity exercises – for e.g. staged attacks (Balakrishnan et al., 2018) – create opportunities for identification and assessment of risks and threats, stakeholders' familiarity therewith, and the efficacy of existing response (Petrenko, 2022; Staves et al., 2022).
	Employee reporting plays a critical role in detecting anomalies; they should have access to **reporting mechanisms** for potential security incidents (Abrahams et al., 2024).
Recovery/ post-incident ***Parameter:*** reaction to and lessons learned from incidents	**Proactive response coordination** to threats and incidents is vital for organizations (Ross et al., 2021). This response should be based on established policies that prescribe actionable tasks (Annarelli et al., 2021) that can be executed and/ or coordinated by capable (external) IS-staff (Linkov et al., 2013; Bagheri, 2020), for e.g. a Cyber Incident Response Team (CIRT).
	Organizational **responses** to cyber incidents should be **evaluated** in order to allocate appropriate human resources and maintain up-to-date awareness of evolving cyber risks. In line herewith, addressing system issues and risks is vital and evaluating recovery durations will enable the organization to modify or replace non-recoverable components (Bagheri et al. 2023).

To commence, CR reflects an organization's capability to endure and adapt to disruptions in its information systems while continuing to achieve its operational objectives (Bagheri, 2021; Eling et al., 2021). Not to be confused with cyber security, CR is commonly regarded as a post-incident construct (Rieger et al., 2021). Cyber security, on the other hand, focuses on the pre-incident protection of electronic information (Calder, 2020). Even so, CR is strongly rooted in pre-incident measures, particularly the deliberate design of resilient systems that enable effective response and recovery from attacks or failures, making cyber security a core element of a CR-strategy. In essence, preparedness, detection, and recovery capabilities are key

enabling conditions for organizational CR (Annarelli et al., 2021; Ross et al., 2021). As such, CR can be divided into three phases, namely (1) planning, (2) detecting, and (3) recovery/post-incident (Linkov et al., 2013; Bagheri, 2020). Table 2 provides a non-exhaustive synthesis of relevant organizational drivers divided over each phase.

2.1.3 ISPC effects and correlation with CR

With regards to the ISPC effects, existing literature suggests that poorly aligned or rigidly implemented security frameworks may fail to deliver intended outcomes (Almuqrin, 2024; Zanke et al., 2024). Thus, organizational context – including culture, leadership and strategic goals – influence the effectiveness of these frameworks. Additionally, given the significant role employees have in security incidents (Fatoki et al., 2024), their involvement in risk management and open dialogue are critical to the effectiveness of an ISMS and ISPC (Ali et al., 2020).

Prior research suggests that ISPC may contribute to reducing cybersecurity incidents, safeguarding the Confidentiality, Integrity and Availability of information systems, and enhancing business continuity, organizational reputation, and risk management capabilities (Cram et al., 2019; Ali et al., 2020; Donalds et al., 2022; Mohamud, 2024). On the other hand, the notoriety of attributing security incidents to human behavior is contradicted by a lack of empirical evidence clearly linking ISPC to reduced security incidents. Thus, the validation of the presumed ISPC effects requires more controlled empirical work (e.g., longitudinal or experimental designs) that quantifies their magnitude.

Collectively, the literature indicates that well-established security policies support organizational security by protecting sensitive information and mitigating vulnerabilities, suggesting a meaningful convergence between ISPC and CR. Moreover, the implementation of policies can be considered the highest level of reactive defense since, for example, access management policies add more layers to information protection than physical and network security alone (Balakrishnan et al., 2018). Ultimately, well-formulated policies that are easily integrated into daily activities enable a better alignment between restrictive and protective security objectives within organizations (Qiu et al., 2020).

Both ISPC and CR encompass an organization's motivation to safeguard valuable information assets. While ISPC includes prescription of responsibilities for mitigating cyber-related risks, CR enables consistent delivery of the intended business outcome despite incidents (Eling et al., 2021). Notably, except for incident response, we also identified all CR-drivers in light of ISPC-drivers. Figure 1 presents the overlap between ISPC and CR.

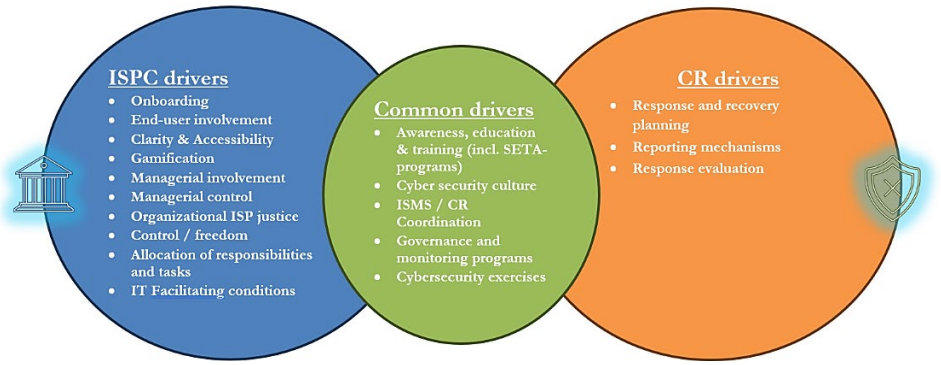

Figure 1: ISPC and CR overlap

In conclusion, effective use of IT systems – perhaps by enhanced ISPC capabilities – contributes to CR (Annarelli et. al, 2021), because consistently secure behavior aids in mitigating risks and enhancing organizational resilience (Oroni et al., 2025).

2.2 State of the art and perspectives for empirical research

A non-exhaustive set of mechanisms that may enhance ISPC and CR emerges from literature. For instance, regarding awareness, an employee may be first introduced to the ISP through onboarding. Subsequently, considering motivation, organizational security culture creates an environment for motivating their compliance with the ISP. Then, relating to capability, appropriate security training enables end-user's capability of utilizing obtained knowledge and necessary related skills. Despite this division, some mechanisms are complementary, and even fluid throughout the AMC-levels. For example, a security culture can also be an introduction to the ISP or enhance capabilities for ISPC for e.g. through security related conversations.

Furthermore, there are many notable similarities between ISPC and CR drivers, such as increased organizational awareness of security activities, security culture, and enhanced IT-capabilities (detection and monitoring). Basically, thus far, ISPC-enhancing drivers seem to be re-usable for CR as both elements are essentially aimed at safeguarding valuable organizational information assets. Nevertheless, our following empirical research further explores the effects of ISPC on the case organization, and how this potentially contributes to their CR.

3 Methodology

Our study adopts a qualitative case study design based on a single holistic case study that facilitated an exploration of the effects of ISPC and its relation to CR in practice. The methodological rigor and quality of this case study is ensured through several strategies (Yin, 2014). Construct validity was guaranteed using multiple sources of evidence – such as document analysis, semi-structured interviews and an exploratory survey – which enabled data triangulation. Internal validity was facilitated by comparing expected theoretical patterns with observed data (pattern matching) and by iteratively refining explanations based on evidence from multiple data sources (explanation building. Regarding external validity, the case organization, further anonymized as ALX, was selected based on its analytical relevance rather than statistical representativeness, thereby supporting analytical generalization. ALX provided a formally implemented and documented ISMS that is governed by a dedicated Information Security department. Finally, reliability was ensured by maintaining a chain of evidence and inviting key informants to review draft findings.

The analyzed documents applicable to ALX were: the ISP, an Annual Awareness Plan (2023-2024), a Risk Control Standard (2017), an Awareness Survey Report (2023), and an Information Security Baseline (2024). Furthermore, the interviewees consisted of the Chief Information Security Officer (CISO), two Information Security Officers (ISOs), an IT Security Officer (ITSO), and an Awareness Officer (AO). These interviewees have an average of 6+ years of experience in Information Security. Lastly, the survey only required employment by ALX. As just 27 of ± 5000 employees completed the survey, its results served an exploratory triangulation purpose, providing indicative insights rather than empirically generalizable findings.

4 Results: applying theory to practice

4.1 ISPC and CR drivers

The research findings suggest that ISPC at ALX is supported by mechanisms spanning all AMC levels, including formal training initiatives, awareness campaigns, and technical infrastructure. Figure 2 presents the identified ISPC and CR drivers within ALX, which were categorized using a three-level index: low (not or rarely identified), medium (partially identified), and high (frequently identified). Notably, despite their formal inclusion in ALX's strategic documentation, security culture and managerial involvement were perceived as relatively weak by interviewees and survey respondents. This can suggest a misalignment in design and implementation regarding leadership and community.

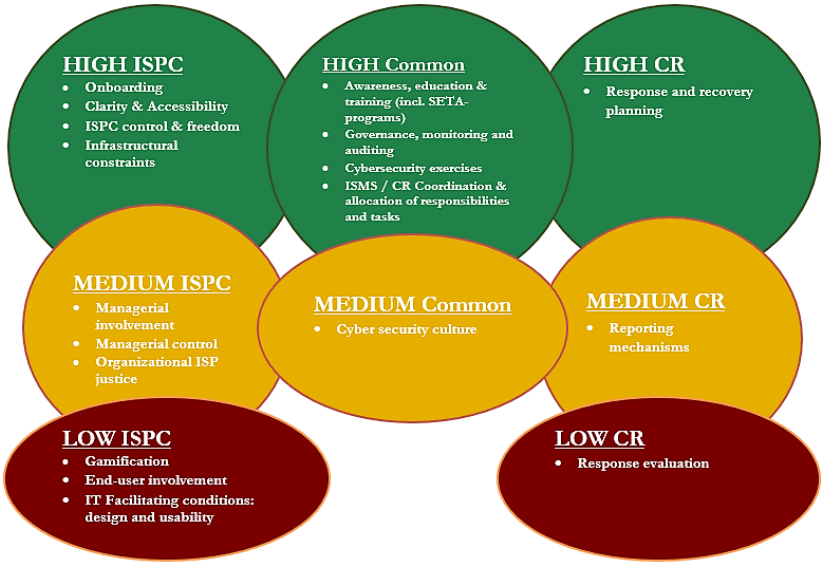

Figure 2: ISPC and CR overlap in ALX

4.2 ISPC effects

ALX's ISP prescribes the contours for how the organization should handle information systems, including processes for identifying, assessing, monitoring, reporting and controlling operational risks and incidents. Hence, through

implementation of its Information Risk Management Policy (and procedures), ALX is better equipped to detect vulnerabilities and threats prior to the maturing of these risks into incidents, providing opportunities for making informed security decisions. In light hereof, the CISO, ISO's and ITSO noted that risk management through ISPC enables ALX to become increasingly better at managing, identifying and mitigating risks. The ITSO mentioned that the ISP describes 'erroneous behavior'; therefore, if employees are aware of what they should comply with, they can report deviations therefrom. Subsequently, the ITSO and AO presume that heightened employee awareness may contribute to incident reduction by enabling earlier reporting and more secure information handling practices. They also infer that ALX's mandatory training "Digital Safety Online" and routine phishing simulations enable a combination of ISPC monitoring, auditing and exercising that facilitate the detection of AMC-related security gaps. Even so, the interviewees explicitly mentioned that ISPC effects can only be clear if ISPC is consistently and properly monitored. Nevertheless, the interviewees agreed that ISPC supports the maintenance of the confidentiality, integrity, and availability of information systems by strengthening risk, incident, and vulnerability management.

4.3 Potential ISPC and CR overlap

ALX's ISP explicitly identifies CR and business continuity as core information security objectives. The CISO and an ISO emphasized that ISPC and CR require a "human-centered approach" in which employees are proactively engaged in information security practices. One ISO stressed that the absence of policies can complicate the organization's chances of achieving CR. Another ISO adds that ISPC specifically helps with identifying and mitigating security risks, which, in turn increases CR. Additionally, the ITSO highlighted that proactive reporting by end-users – as prescribed by the ISP – is essential for an enhanced response to anomalies, as this will better equip ALX to deal with these on time, hence, aiding CR.

Furthermore, the findings highlight multiple organizational mechanisms associated with both ISPC and CR, including awareness, training, security culture, cyber exercises, governance, monitoring, coordination, and reporting. This follows ALX's adoption of the "Three Lines of Defense" approach wherein specific security-related tasks are divided amongst management, academic and supporting staff. In

conclusion, within ALX, reinforcing ISPC and CR relies on human-centered mechanisms such as security awareness, training, incident reporting, and response capabilities, supplemented by technical mechanisms, such as monitoring.

5 Discussion, limitations and future research

This study advances understanding of ISPC and CR by examining the mechanisms that support both phenomena and by theorizing their interconnection. Most literature treated ISPC and CR as separate domains. For instance, compliance is addressed as a dependent variable for behavioral change, and CR is treated as an outcome of systems, not behavior. Thus, the linkage between the two remains under-theorized. Hence, this study conceptually shifts ISPC from abiding by rules to addressing it as a dynamic capability that strengthens CR, conveying attention from whether compliance occurs to how it is organizationally enabled.

The findings further suggest that ISPC drivers can function as enabling mechanisms as well as observable indicators of compliance-related effects. For example, cyber exercises such as phishing simulations provide insight into employees' level of knowledge, willingness, and capability to respond to cyber threats. When such reporting behavior is institutionalized, organizations may develop a more coordinated approach to threat detection while simultaneously revealing socio-technical vulnerabilities across departments and infrastructures. Subsequently, misalignment within the organization's approach, both in terms of personnel (e.g., departmental differences), and technology (e.g., accessibility of the reporting database) can be identified and addressed, potentially creating evidence-based insights about the organization's overall CR posture.

Additionally, the findings contemplate that alignment of ISPC mechanisms with CR-oriented practices can aid positioning organizations better to withstand and adapt to the cyber threat landscape. Thus, the ISP could form a foundational governance instrument within broader CR strategies. However, the findings equally indicate that compliance alone is not sufficient to ensure CR. Rather, resilience is reinforced when compliance is embedded within an integrated socio-technical configuration spanning people, processes, and technology, whilst balancing prevention and response strategies. This underscores the importance of complementing policy enforcement with leadership engagement, security culture, and continuous capability building.

Limitations and future research

This study is subject to several limitations. First, the findings may reflect perceptual bias concerning the quality, enforceability and effectiveness of the ISPC as insights from the broader employee population were not extensively captured. Second, the relatively short research window of five months essentially implies a cross-sectional design and constrains the ability to observe longer-term longitudinal compliance dynamics. Third, the operationalization and measurement of ISPC effects remain difficult to isolate from broader information security governance outcomes. Finally, the study primarily focuses on organizational compliance mechanisms without explicitly accounting for how rapid technological advancements continuously reshape security requirements and user interactions and may underrepresent the dynamic interplay of innovation and security.

Recommendations for future research

Several avenues for future research emerge from this work. First, longitudinal research designs are needed to generate stronger empirical evidence regarding the organizational effects of ISPC. Such studies should account for dynamic contextual factors, including technological developments, evolving threat landscapes, and the dual retrospective and preparatory nature of information security policies. Moreover, measurable evidence of ISPC-effects could be provided through combining quantitative metrics, comparative analysis, and business impact linkage. Second, further exploration of the design and implementation of mechanisms for monitoring end-user ISPC could provide deeper insight into behavioral compliance and its associated organizational impacts. Third, further research is required to empirically assess the situational and organizational conditions and effectiveness of mechanisms aimed at enhancing ISPC and CR. Finally, considering the Network and Information Security Directive 2, it is particularly valuable to examine how Dutch Higher Education Institutions strengthen their CR strategy in compliance herewith.

6 Conclusion

Our work highlights the importance of human-centered mechanisms, such as awareness of security activities, security culture, and enhanced IT-capabilities (detection, monitoring, reporting and response) as critical links that influence both

ISPC and CR. This study also contributes to the extant literature by proposing ISPC as an organizational enabler of CR, rather than solely an individual-level behavioral outcome. In doing so, this research extends existing compliance-focused research through positioning CR as a potential complementary outcome of an ISMS. Thus, it demonstrates that ISPC and CR are closely related and mutually reinforcing, suggesting that ISPC can potentially be utilized as a foundational layer of CR, with its strongest contributions concentrated on preparedness. Taken together, this research reconceptualizes ISPC as a resilience-enabling capability, arguing that ISPC-enhancing mechanisms can play a critical role in shaping organizational CR.

References

Abrahams, T. O., Farayola, O. A., Kaggwa, S., Uwaoma, P. U., Hassan, A. O., & Dawodu, S. O. (2024). Cybersecurity awareness and education programs: a review of employee engagement and accountability. *Computer Science & IT Research Journal*, *5*(1), 100-119.

Aggarwal, A., & Srivastava, S. K. (2024). Synthesizing Information Security Policy Compliance And Non-compliance: A Comprehensive Study And Unified Framework. Journal of Organizational Computing and Electronic Commerce, 34(4), 338–369. https://doi.org/10.1080/10919392.2024.2381303

Alassaf, M., & Alkhalifah, A. (2021). Exploring the influence of direct and indirect factors on information security policy compliance: A systematic literature review. IEEE Access, 9, 162687-162705. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9634004

Alhidaifi, S. M., Asghar, M. R., & Ansari, I. S. (2024). A survey on cyber resilience: Key strategies, research challenges, and future directions. ACM computing surveys, 56(8), 1-48. https://doi.org/10.1145/3649218

Ali, R. F., Dominic, P. D. D., Ali, S. E. A., Rehman, M., & Sohail, A. (2021). Information Security Behavior and Information Security Policy Compliance: A Systematic Literature Review for Identifying the Transformation Process from Noncompliance to Compliance. Applied Sciences, 11(8), 3383–. https://doi.org/10.3390/app11083383

Alias, R. A. (2019). Information security policy compliance: Systematic literature review. Procedia Computer Science, 161, 1216-1224. https://doi.org/10.1016/j.procs.2019.11.235

Almuqrin, A. (2024). How About Enhancing Organizational Security. Journal of Global Information Management, 32(1). https://doi.org/10.4018/JGIM.358745

Amankwa, E., Loock, M., & Kritzinger, E. (2018). Establishing information security policy compliance culture in organizations. Information and Computer Security, 26(4), 420–436. https://doi.org/10.1108/ICS-09-2017-0063

Annarelli, A., & Palombi, G. (2021). Digitalization Capabilities for Sustainable Cyber Resilience: A Conceptual Framework. Sustainability (Basel, Switzerland), 13(23), 13065-. https://doi.org/10.3390/su132313065

Bagheri, S. (2020). Investigating organisational aspects of cyber resilience in large organisations (Doctoral dissertation, University Of Tasmania). https://web.archive.org/web/20210613091531id_/https://eprints.utas.edu.au/35959/1/Bagheri_whole_thesis.pdf

Bagheri, S., Ridley, G., & Williams, B. (2023). Organisational Cyber Resilience: Management Perspectives. Australasian Journal of Information Systems, 27. https://journal.acs.org.au/index.php/ajis/article/view/4183/1315

Balakrishnan, U. R., Mishra, G., & Bhatt, N. (2018). Strategies for Improving Cyber Resilience of Data-Intensive Business Information Systems. In Handbook of e-Business Security (pp. 285-312). Auerbach Publications. https://doi.org/10.1201/9780429468254

Balagopal, N., & Mathew, S. K. (2024). Exploring the factors influencing information security policy compliance and violations: A systematic literature review. Computers & Security, 147, 104062. https://doi.org/10.1016/j.cose.2024.104062

Calder, A. (2020). The cyber security handbook–Prepare for, respond to and recover from cyber attacks. IT Governance Publishing Ltd.

Cram, W. A., D'arcy, J., & Proudfoot, J. G. (2019). Seeing the forest and the trees: a meta-analysis of the antecedents to information security policy compliance. MIS quarterly, 43(2), 525-554. https://doi.org/10.25300/MISQ/2019/15117

D'Arcy, J., & Lowry, P. B. (2019). Cognitive-affective drivers of employees' daily compliance with information security policies: A multilevel, longitudinal study. Information Systems Journal (Oxford, England), 29(1), 43–69. https://doi.org/10.1111/isj.12173

Delso-Vicente, AT., Diaz-Marcos, L., Aguado-Tevar, O. et al. (2025) Factors influencing employee compliance with information security policies: a systematic literature review of behavioral and technological aspects in cybersecurity. Futur Bus J 11, 28 (2025). https://doi.org/10.1186/s43093-025-00452-7

Donalds, C., & Barclay, C. (2022). Beyond technical measures: a value-focused thinking appraisal of strategic drivers in improving information security policy compliance. European Journal of Information Systems, 31(1), 58-73. https://doi.org/10.1080/0960085X.2021.1978344

Dong, Y., Chen, S., & Wu, Y. (2023). Keeping up with the Joneses: The role of investee peers corporate environmental responsibility. *Corporate Social Responsibility and Environmental Management*, *30*(4), 1841-1855. https://onlinelibrary.wiley.com/doi/epdf/10.1002/csr.2458?saml_referrer

Eling, M., McShane, M., & Nguyen, T. (2021). Cyber risk management: History and future research directions. Risk Management and Insurance Review, 24(1), 93-125.

Falahati, A., Lapointe, L., & Beaudry, A. (2022). Employees' Compliance with ISP: A Socio-Technical Conceptual Model. https://doi.org/10.24251/HICSS.2022.840

Fatoki, J. G., Shen, Z., & Mora-Monge, C. A. (2024). Optimism amid risk: How non-IT employees' beliefs affect cybersecurity behavior. *Computers & Security*, *141*, 103812. https://doi.org/10.1016/j.cose.2024.103812

Hovav, A., & Putri, F. F. (2016). This is my device! Why should I follow your rules? Employees' compliance with BYOD security policy. Pervasive and Mobile Computing, 32, 35–49. https://doi.org/10.1016/j.pmcj.2016.06.007

Keys, B., & Shapiro, S. (2018). Frameworks and Best Practices. In Cyber Resilience of Systems and Networks (pp. 69–92). Springer International Publishing. https://doi.org/10.1007/978-3-319-77492-3_4

Khando, K., Gao, S., Islam, S. M., & Salman, A. (2021). Enhancing employees information security awareness in private and public organisations: A systematic literature review. Computers & security, 106, 102267. https://www.sciencedirect.com/science/article/pii/S0167404820300675

Kott, A., & Linkov, I. (Eds.). (2019). Cyber resilience of systems and networks. New York, NY: Springer International Publishing. https://link.springer.com/content/pdf/10.1007/978-3-319-77492-3.pdf

Li, W., Kang, X., He, A., & Zhou, M. (2025). CEO organizational identification and firm strategic change: an AMC framework. *Asia Pacific Journal of Management*, 1-34. https://link.springer.com/article/10.1007/s10490-025-10056-y

Linkov, I., Eisenberg, D. A., Bates, M. E., Chang, D., Convertino, M., Allen, J. H., Flynn, S. E., & Seager, T. P. (2013). Measurable Resilience for Actionable Policy. Environmental Science & Technology, 47(18), 10108–10110. https://doi.org/10.1021/es403443n

Mohamud, A. J. (2024). Impact of information security policies compliance (ISPC) on reducing the incidence of security breaches in organizations: Systematic Literature Review. https://www.preprints.org/manuscript/202409.1715

Moody, G. D., Siponen, M., & Pahnila, S. (2018). Toward a unified model of information security policy compliance. MIS quarterly, 42(1). https://doi.org/10.25300/MISQ/2018/13853

Mubarkoot, M., Altmann, J., Rasti-Barzoki, M., Egger, B., & Lee, H. (2023). Software Compliance Requirements, Factors, and Policies: A Systematic Literature Review. Computers & Security, 124, 102985-. https://doi.org/10.1016/j.cose.2022.102985

Oroni, C. Z., & Xianping, F. (2025). Evaluating the influence of cybersecurity policies and cybersecurity behavior on institutional security performance in remote learning: the moderating role of technological readiness. *Sustainable Futures*, *10*, 101554. https://doi.org/10.1016/j.sftr.2025.101554

Petrenko, S. (2022). Cyber resilience. CRC Press.

Qiu, J., Tian, Z., Du, C., Zuo, Q., Su, S., & Fang, B. (2020). A survey on access control in the age of internet of things. IEEE Internet of Things Journal, 7(6), 4682-4696. https://doi.org/10.1109/JIOT.2020.2969326

Rieger, C., Schultz, K., Carroll, T., & McJunkin, T. (2021). Resilient control systems—basis, benchmarking and benefit. IEEE Access, 9, 57565-57577. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9399144

Ross, R., Pillitteri, V., Graubart, R., Bodeau, D., McQuiad, R. (2021). NIST Special Publication 800-160, Volume 2 Revision 1. Developing Cyber-Resilient Systems: A Systems Security Engineering Approach. https://doi.org/10.6028/NIST.SP.800-160v2r1

Sharma, A., Koohang, A., & Singh, S. P. (2025). Information Security Policy Compliance: a structured review using scientometric analysis and topic modeling. Journal of Global Information Management (JGIM), 33(1), 1-32. https://doi.org/10.4018/JGIM.389715

Staves, A., Anderson, T., Balderstone, H., Green, B., Gouglidis, A., & Hutchison, D. (2022). A Cyber Incident Response and Recovery Framework to Support Operators of Industrial Control Systems. International Journal of Critical Infrastructure Protection, 37, 100505–. https://doi.org/10.1016/j.ijcip.2021.100505

Stewart, S. E. (2021). Managerial Control Effects on Information Security Policy Compliance Intentions: Considerations of Formal and Informal Modes of Control. https://digitalcommons.fiu.edu/cgi/viewcontent.cgi?article=6203&context=etd

Tsohou, A., & Holtkamp, P. (2018). Are users competent to comply with information security policies? An analysis of professional competence models. Information Technology & People (West Linn, Or.), 31(5), 1047–1068. https://doi.org/10.1108/ITP-02-2017-0052

Vollenweider, F., & Jahankhani, H. (2023, January). What Drives Generation Z to Behave Security Compliant? An Extended Analysis Using the Theory of Planned Behaviour. In Cybersecurity in the Age of Smart Societies: Proceedings of the 14th International Conference on Global Security, Safety and Sustainability, London, September 2022 (pp. 315-341). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-20160-8_17

Whitman, M. E., & Mattord, H. J. (2018). Management of Information Security. Cengage Learning.

Xue, B., Xu, F., Luo, X., & Warkentin, M. (2021). Ethical leadership and employee information security policy (ISP) violation: Exploring dual-mediation paths. *Organizational Cybersecurity Journal: Practice, Process and People*, *1*(1), 5-23. https://www.emerald.com/ocj/article-pdf/1/1/5/2066198/ocj-02-2021-0002.pdf

Yin, R. K. (2018). Case study research and applications (Vol. 6). Thousand Oaks, CA: Sage.

Zammani, M., Razali, R., & Singh, D. (2019). Factors contributing to the success of information security management implementation. International Journal of Advanced Computer Science and Applications, 10(11). https://thesai.org/Downloads/Volume10No11/Paper_53-Factors_Contributing_to_the_Success_of_Information_Security.pdf

Zanke, A., Weber, T., Dornheim, P., & Engel, M. (2024). Assessing information security culture: A mixed-methods approach to navigating challenges in international corporate IT

departments. *Computers & Security*, *144*, 103938. https://www.sciencedirect.com/science/article/pii/S0167404824002438

Zhang, Z., & Gao, Q. (2025). Practice what you preach: how and when does leaders' information security policy compliance influence subordinates' information sharing?. *Journal of Knowledge Management*. https://doi.org/10.1108/JKM-02-2025-0242

Students' Exploration of Elective Courses in Flexible Higher Education: Findings From a Digital Simulation Task

Suzan van Brussel,[1] Johan Smarius,[1]
Sami Heikkinen,[2] Esther van der Stappen[1]

[1] Avans University of Applied Sciences, Research Group Digital Education, Breda, the Netherlands
sjm.vanbrussel@avans.nl,jaw.smarius@avans.nl,ej.vanderstappen@avans.nl
[2] LAB University of Applied Sciences, Faculty of Business and Hospitality Management, Lahti, Finland
sami.heikkinen@lab.fi

DOI
https://doi.org/10.18690/um.fov.4.2026.12

ISBN
978-961-299-152-4

Flexible higher education requires students to compose individual learning routes by selecting elective courses. Such choice-making processes can induce stress and procrastination and thus demand effective guidance. This exploratory mixed-method study examines the students' use of a digital tool to support them during exploration-in-breadth, the initial phase of educational choice-making processes. Bachelor students (n = 49) used a prototype that enabled browsing, filtering, and shortlisting electives. Trace log data and sequence analysis captured interaction patterns, including time-on-task, filter use, and selection behaviour, while a survey assessed motives, usability, and support needs. Results show that students display diverse navigation behaviour and that such a tool can facilitate structured exploration, particularly through filtering before shortlisting favourites. Students valued the tool for orientation but indicated the need for personal guidance when making final decisions. These findings support the potential for integrating asynchronous digital exploration with synchronous mentoring (blended guidance). Design implications include adaptive technological support, integration of mentor dialogue, and explicit development of students' choice-making competencies.

Keywords:
flexible education, technology-enhanced learning, choice-making processes, exploration task, sequence analysis

University of Maribor Press

1 Introduction

Higher education (HE) is undergoing a significant shift towards flexibility, a trend illustrated by the prominence of flexible education in the EU Strategic Agenda for Higher Education (EC, 2017). Two main goals of flexible HE are 1) accessibility to better accommodate an increasingly diverse student population, and 2) adaptability to better prepare for future jobs (Jonker et al., 2020). Flexible education, such as the integration of elective courses in the curriculum, is seen as an approach that offers students the opportunity to take more responsibility for their own learning as well as the opportunity to choose education that matches their own interests (Wade, 1994). The ability to make choices based on individual interests is associated with higher academic performance and better well-being. Moreover, it is deemed necessary to remain motivated (Harackiewicz et al, 2008; Schulz et al., 2018).

Research into the needs of students regarding flexible education also shows that they wish for more flexibility than they experience (Huizinga et al., 2022). In addition, they indicate that they need guidance in making such choices (Huizinga et al., 2022; Bolhuis et al., 2020). Choice overload and decision-making anxiety can discourage young people to make a choice at all ("The paradox of choice"- Schwartz, 2016). To date, there is insufficient knowledge of how educational institutions implementing flexible education can support young people in making choices for elective courses to construct their own learning path (Van der Stappen, 2022).

2 Related work

2.1 Choice-making processes

The Dual Cycle Model outlines the process of identity formation of young adults and involves several choice-making tasks (Figure 1). This model shows which steps are part of a well-informed choice making process. Indeed, conscious choice-making activities contribute to identity formation: *I know what I want* (Luyckx et al., 2006; Zimmerman et al., 2015; Germeijs et al., 2006). Insight into your 'internal compass' provides direction for the choice-making process, and to the required information, to make a well-considered choice (Meens, 2020). A well-informed choice, in turn, leads to a satisfactory choice (Meens, 2020; Zimmermann et al., 2015). Having to make well-informed choices is difficult and might lead to stress and procrastination

(Meens, 2020; Schwartz, 2016). Therefore, students need support in exploring the set of options they can choose from.

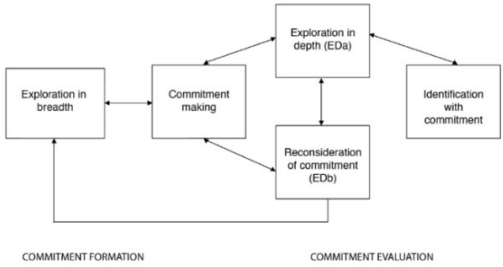

Figure 1: Choice-making process according to the Dual Cycle Model
Source: (Luyckx et al., 2006; Zimmerman et al, 2015)

Research on decision-making processes has identified five styles (Scott & Bruce, 1995). Students can adopt 1) a *rational style*, in which the student displays strong engagement and arrives at a clear choice with conviction; 2) an *intuitive style,* in which the student does not engage in extensive orientation, but still chooses with confidence; 3) an *avoidant style*, in which the student prefers to postpone or even avoid making a final decision; 4) a *dependent style*, where the student seeks extensive help and advice from others; or 5) a *spontaneous style*, characterised by spontaneous and impulsive choices. Which style – or combination of styles – a person adopts for a particular choice depends on their personality, the type of choice, and their previous experience with similar choices. Although the use of digital tools to support students' exploration and choice-making has increased, little is known about how students with diverse choice-making styles can be supported with such tools. This lack of insight hampers the development of effective support mechanisms, both within the digital environment and in face-to-face guidance.

2.2 Digital support tools

During the choice process to select an elective course, students evaluate their internal compass and seek and weigh information about the options they can choose from. Digital tools have the potential to support students, e.g., by scaffolding the process, navigating all options in a structured way, or accessing the required knowledge and insights independent of time and place (Smarius & Van der Stappen, 2022).

This study builds on two previous design-based studies (Smarius & Van der Stappen, 2022; Van Brussel et al., 2025). First, design propositions for digital support tools were developed, accompanied by two prototypes implementing these propositions, to support students in the early phase of the choice process known as *exploration-in-breadth* (see Figure 1; Luyckx et al., 2006; Zimmerman et al., 2015; Germeijs et al., 2006). Next, design propositions for blended guidance were developed and a prototype of an arrangement of guidance activities to support the entire choice-making process. The current study uses a software prototype that emerged from two prototypes of the study by Smarius & Van der Stappen (2022).

This study answers the following research question: *How do students use a digital support tool to explore elective courses, what motivates their selections, and what additional support do they need to make well-informed choices?* Based on the answers, we formulate design implications for technology-enhanced and personalised support of choice-making processes of students in flexible higher education.

3 Method

This study followed an explorative, mixed-method approach, in which students performed a simulated exploration and choice-making task (Dooley, 2002) and subsequently completed a survey. The survey consisted of 14 items (See Appendix B). From the prototype application described in the previous section, we retrieved the trace log data to investigate student interaction and navigation behaviour in the tool.

3.1 Participants and sample size

The study was conducted at a Dutch higher education institute that offers students the opportunity to compose 25% of their curriculum with elective courses (the first elective is scheduled in the third quarter of the second year). Participants were full-time students from six departments within the institution. Academic staff recruited these students; all students were active members of an academy council. The choice tool was provided online and students logged in anonymously. All participants signed an informed consent form, explaining why and how their data was collected, analysed and stored in a secure environment. Fifty-five students from different

domains logged into the tool, 49 completed the simulation task, of which 44 students completed a survey.

3.2 Procedure

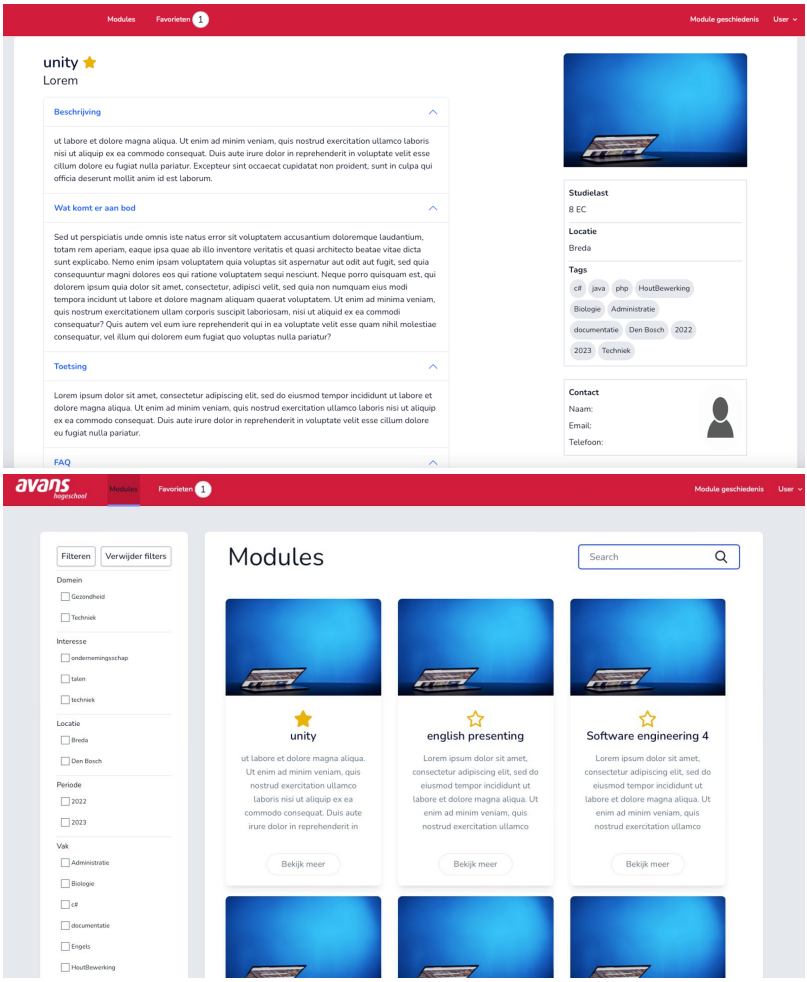

Figure 2: Screenshots of the tool used in this study

Tool design. The elective course portfolio of 211 elective courses that is offered in the upcoming years was imported into the digital tool (see Figure 2 for screenshots, displaying mock data). The simulation tool allowed students to browse, filter, and

mark elective courses (termed Modules). Students logged in and were able to choose between two tiles which were presented on the overview page: Modules and Favourites. In the Modules section, an overview of the 211 courses in the portfolio was shown (Figure 2, left). By clicking on the tile for a specific course, more detailed information about the course was shown (Figure 2, right). It was possible to mark a course as favourite both from the detailed information and from the overview page. Students were able to filter by single or multiple tags, with the possibility to add and remove filters during the session (see Appendix A for the filters).

Within the 'Favourites' tile it was possible to search, remove, and navigate the marked courses. It was also possible to see the courses that the student previously marked as favourite but that they later removed (history page).

The simulation task. Participants were asked to explore elective courses and compile a shortlist, as if they were choosing an elective for their next semester. The task was to first mark NLQF-5 level courses they would like to explore further, and subsequently NLQF-6 courses. In both tasks, students marked their favourite courses with a star, and this list could be adjusted during the whole task. There was no minimum, maximum or preferred number given for the favourite courses students could select.

Analysis. To examine students' interaction patterns with the digital simulation tool, we conducted a descriptive analysis of logged trace data. The dataset included clicking behaviour, filter usage, and marking and unmarking of favourites. We calculated mean, median, range, and standard deviation for the number of courses marked as favourite and time spent in the tool (time-on-task). We compiled sessions per student based on a time limit of 90 minutes; for five students, a second session was found, all consisting of at most 13 clicks. We removed these sessions as their timestamps did not correspond with the scheduled timing of the simulation tasks. Finally, we performed a sequence analysis in R (Saqr et al., 2024; Valle Torre et al., 2024) using the TraMineR package (Gabadinho et al., 2011), to explore variation in navigation behaviour and to identify clusters of how students engaged with the tool during the exploration-in-breadth phase.

Finally, students were asked to complete the survey consisting of 14 items (Appendix B). Categories were the number of items on each shortlist (Q1,2), motives to mark courses as favourites (Q3,4, 5,12,13), information needs during the choice-making process (Q6,7,8), usability of the choice tool (Q9,10,14), and the student's current academy (Q11). We analysed the survey data quantitatively by determining the mean and median (Q1,2), the current academy (Q11), and the total percentages of each answer option (Q3,4,6). We qualitatively analysed the open answers given to Q5 and 7–14. We analysed deductively, based on the relevant literature.

4 Results

On average, participants marked 5.6 elective NQLF-5 courses as favourite (*SD* = 2.85), and 5.8 NQLF-6 courses (*SD* = 2.99). The median in both levels was 5, and the range of marked courses was 0-20 for NQLF-5 (*SD* = 4.84), and 0-15 for NQLF-6 (*SD* = 3.69). Two students marked no favourites at the NQLF-5 level and four (different) students marked no favourites at the NQLF-6 level. From the 211 elective courses, 35 were never marked as favourite. The top three courses that were marked favourite were:

1. Introduction to Psychology (12 times – NQLF-5)
2. Data-driven decision-making with AI (10 times – NQLF-6)
3. Cybersecurity (NQLF-5) *and* International Project Management & Stakeholder Engagement (NQLF-6), both 8 times)

4.1 Filters and search terms

We cumulatively counted which filters students used. Most used were filters related to the 'domain' (440 times) and 'level' (421 times) of electives, followed by 'location' (270 times). The domain filter referred to the educational domains of the university, e.g., 'Technology & Engineering' (94) or 'Economics' (85). The filter for courses that transcend domains ('Cross-sectoral') was used 43 times. The NQLF-5 and NQLF-6 filters were used 256 and 161 times, respectively. The 'type of cooperation' filter (e.g., monodisciplinary or multidisciplinary) was used merely 12 times. See Figure 3 for an overview of the chosen filters.

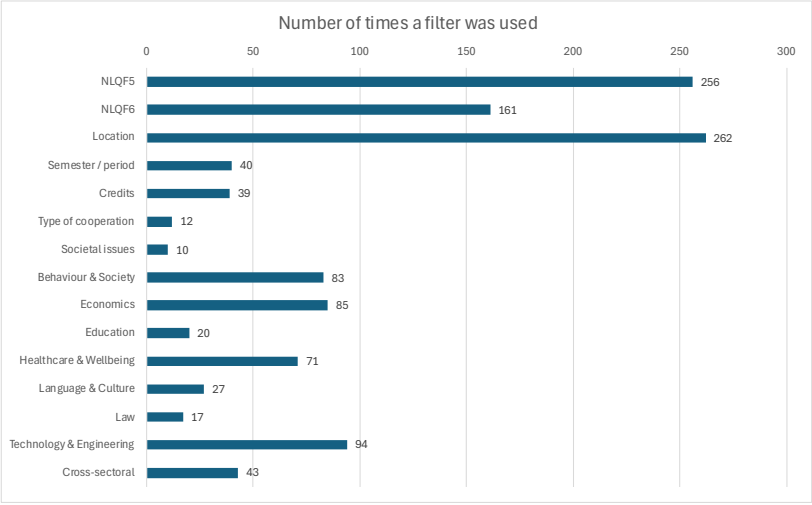

Figure 3: Overview of the chosen filters.

Search terms were used less than filters. An analysis of the search terms showed that 'care' and 'health' were used most (32 and 21 times, respectively), followed by 'innovation' (6), design, AI, research, and project (4), impact and management (3). Furthermore, 72 search terms were used only a single time (e.g., serious gaming, artistic, sustainability, and finance).

4.2 Trace data

Sequence analysis of the trace log data reveals the behavioural patterns that students displayed when interacting with and navigating in the tool. A total of 42 distinct sequences were analysed. On average, students spent 45.23 minutes in the online environment (SD = 20.4), with a median of 41.01 minutes, and a range spanning from 9.7 to 88.2 minutes. Overall, interaction with the tool was heavily centred on filtering operations and viewing the module overview. Across the entire dataset, the highest average time was spent in the 'Filter' state (mean = 89.28 seconds) and navigating to 'Filter -next page' (mean = 58.83s), followed by viewing the modules themselves (mean = 49.91s). We mapped the chronological sequence of eight distinct states (e.g., Filter, Show details, Favourites) for each student. Using optimal matching to calculate substitution costs and applying Ward's hierarchical clustering method (Ward, 1963; Murtagh & Legendre, 2014), three distinct behavioural profiles emerge (see Figure 4):

- **Cluster 1: Rapid Explorers (n = 20).** This largest group is characterised by the shortest average time-on-task (M = 36.5 minutes, SD = 17.8). These students spent a moderate amount of time setting filters (mean state time = 61.8s) and subsequently navigating module lists (mean state time = 26.8s) or viewing detailed descriptions (mean = 26.5s) before making a selection. Their sequences show straightforward progression from filtering to evaluating details, mostly resembling the intuitive choice-making style.
- **Cluster 2: Deep Evaluators (n = 15).** Students in this cluster spent significantly more time (M = 52.8 minutes, SD = 17.6) and used the filter options most (mean state time = 136.27s) and extensively reviewing module overviews (mean = 93.40s). Their behaviour mostly resembles the rational choice-making style where students view and adjust parameters across a wide breadth of options before adding items to their shortlist.
- **Cluster 3: Prolonged Browsers (n = 7).** This smallest cluster records the highest average time-on-task (M = 53.9 minutes, SD = 25.0). These students spent an overwhelming proportion of their session navigating to subsequent pages of search results ('Filter -next page'; mean = 140.71s). Furthermore, they spent less time reading specific module details (mean = 21.71s). This pattern shows superficial exploration behaviour, where students browse expansively through paginated lists, mostly resembling the avoiding style.

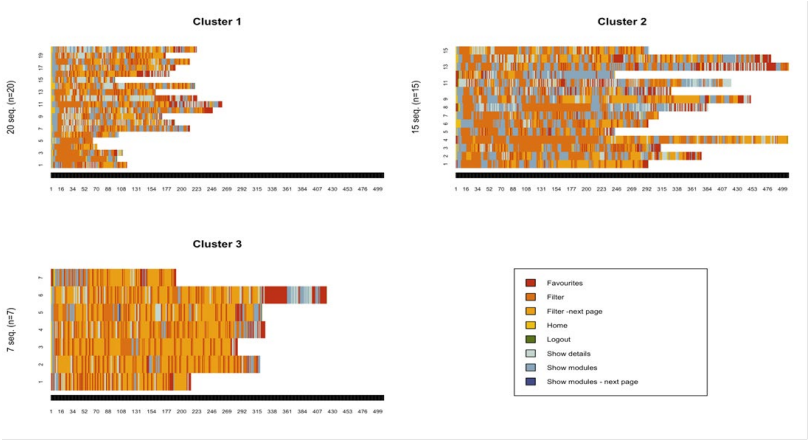

Figure 4: Overview of all distinct student sequences per cluster. Each row represents one student's session.

4.3 Survey insights

A total of 44 students completed the survey after using the tool for the simulated choice task. The participants were students from three different educational domains, namely Economy (n=19), Technology & Engineering (n=14), and Health & Education (n=11).

We asked participants to rank the reasons to choose the courses that they shortlisted based on importance (Q3). The most frequently ranked reason for adding an elective course to the shortlist was personal interest; 39% of students ranked this for their first choice and this reason remained influential across subsequent choices. Other prominent factors included professional development (18% for first choice) and perceived usefulness (14% for the first choice).

In addition, we asked participants to rank other important reasons to choose the courses they shortlisted (Q4). These included additional motivations such as broadening knowledge (55% for first choice) and deepening expertise (55% for second choice), suggesting that students value both diversification and specialization. Less common reasons included alignment with personal values, talents, and future study or career goals, which appeared in later choices. Only 2% of the students indicated that the possibility of following the course online influenced their decision to put a course on their short list. Elective courses were most often excluded because they did not align with students' interests, professional development needs, or learning objectives, or because the information provided was insufficient. Location and unclear descriptions were also mentioned as barriers.

Regarding the information needs to make a well-informed choice, 43% of the students requested more detailed course descriptions, including learning objectives, workload, assessment methods, and teaching formats. Furthermore, 23% indicated a need for personal guidance, such as conversations with a coach or course coordinator, while 20% wanted clearer information about the overall choice-making process. A small minority (7%) reported relying primarily on intuition and not requiring further details.

5 Discussion

Based on the qualitative and quantitative results, we see that students apply different interaction patterns during a simulated task related to exploration-in-breadth (Scott & Bruce, 1995; Meens, 2020); the three clusters we found aligned most with the intuitive, rational and avoiding style. However, it also might be that the simulation tool and the assignment stimulate certain behaviours over others, e.g., completely avoiding the choice process was not an option in the experimental setting. Moreover, the simulated exploration had no clear decision constraints, trade-offs, or consequences. Still, some students might rely on intuitive or dependent choice styles, indicating the need for adaptive tools that cater to different decision-making styles (Scott & Bruce, 1995; Meens, 2020). Further research might explore the relationship between choice style, information need and choice behaviour, which could lead to the development of a more adaptive choice tool that suits all types of choice makers.

Regarding what motivates students, the results show that the elective courses that are marked as favourite meet students' personal interests, professional development and deepening or broadening knowledge. It is therefore important that students have insight in their internal compass (their interests, values and goals – Meens, 2020). Discovering their internal compass should therefore be considered as part of a (blended) guidance program. Practical information must be concise but complete (e.g., location, credits, assessment description). These conclusions are in line with insights from previous research (Smarius & Van der Stappen, 2022). Building on these findings, future research might explore what information students need in later stages of the choice-making process (exploration in-depth and commitment).

Students indicate the need for additional support in the form of a conversation with their mentor about the shortlist. The tool provides overview and structure, but students indicate that they need in-person meetings to clarify choices and align them with their learning goals, supporting the blended approach. However, future research with students can unravel which content can only be discussed by an in-person meeting and which parts can be digitalised. For example, Virtual Reality (VR) might substitute some parts of face-to-face interaction by experiencing elective course settings beforehand. Compared to textual or oral information, a VR experience leads to increased immersion by creating the perception of physical

presence in a non-physical world (Tao, 2019; Mayer et al., 2022), which could encourage active exploration and reflection (Geriş & Kulaksız, 2025).

A clear limitation of this study lies in its simulated nature. Participants were not required to make real decisions with tangible consequences, which may have influenced their level of engagement, search behaviour, and use of filtering options. As such, the observed interaction behaviours and patterns do not fully reflect authentic decision-making behaviours. Moreover, the students were selected from a pool of active academy council members and are therefore probably not representative of the entire student population. Furthermore, while trace data provides valuable insights into user interactions with the tool, its interpretation remains complex. For instance, frequent clicking may indicate either active exploration or confusion. Future studies should therefore consider combining trace data analysis with methods such as think-aloud protocols, walkthroughs or interviews. Such methods may give more insight into students' considerations in the exploration-in-breadth phase.

6 Conclusion

This study investigated how students use a digital tool to explore elective courses, what motivates their selections, and what additional support they need to make well-informed choices. According to the Dual Cycle Model, exploration-in-breadth is a key phase in decision-making. The findings suggest that students engage in both broad scanning and focused selection, consistent with this model. Students could benefit from digital tools for initial exploration but express the need for personal guidance to make final decisions. This outcome supports a blended guidance approach to support students in their choice-making process for elective courses in flexible higher education.

From the results of the logged trace data as well as the subsequent survey, we can deduce that students differ in behaviour and do not proceed through the exploration-in-breadth phase in a straightforward manner: They alternate between general orientation and targeted selection, requiring tools that facilitate this behaviour. We identified three clusters of students that display different navigation behaviour: rapid explorers, deep evaluators and prolonged browsers. Many students

express the need for more information on both the courses and well-informed choice-making and for personal guidance from a coach or coordinator.

A blended guidance approach to support students in their choice-making process for elective courses in flexible higher education is promising. Mentors or teachers of elective courses can inform students about the content of the module in relation to the students' goals, what pedagogical approach is applied, and how the learning goals are tested. For scalable and rich support, immersive technologies and recommender systems can be implemented. Van Rossen et al. (2024) posit in their position paper that more research is needed to explore benefits and risks of recommender systems based on Artificial Intelligence in decision-making processes and the impact of these systems on students when selecting elective courses in flexible curricula.

The findings of this study offer valuable insights into students' decision-making behaviour for educational purposes and can contribute to the development of evidence-informed tools to support students' choice-making process for elective courses in flexible higher education. Despite growing interest in the current topic, little is still known about, for example, contextual factors and their influence on choice behaviour and whether a digital choice support tool actually leads to more satisfaction with the choice or to better performance.

7 Implications for practice and policy

Institutions that are flexibilising their curriculum should support students in choosing elective courses, and this study contributes to knowledge about how this can best be facilitated with technology. Training mentors to guide students effectively in making well-informed choices is also essential. The current study provided insights in one part of the choice-making process, namely exploration-in-breadth supported by a digital tool. An important outcome is that students appreciate the technological support and that they differ in their behaviour in the tool. Importantly, students indicate that interaction with their coach (or mentor) is essential for them to able to make an informed choice for an elective course that fits their professional interests and goals.

Acknowledgements

The authors thank the students and teachers that participated in and/or contributed to our study.

References

Bolhuis, E., van der Kaap, G., Petter-Mikx, E., Jansen, P., te Wierik, M., & de Graaf, S. (2020). Mag ons hoger beroepsonderwijs wat meer flexibel? *Tijdschrift Voor Hoger Onderwijs, 38(*1), 41–60. https://tvho.nl/article/view/13308

Dooley, K. (2017). Simulation research methods. In S. R. Clegg, C. Hardy, T. B. Lawrence, & W. R. Nord (Eds.), *The Blackwell companion to organizations* (pp. 829–848). Blackwell. https://doi.org/10.1002/9781405164061.ch36

European Commission. (2017). *EU agenda for higher education*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2017%3A247%3AFIN

Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, *40*, 1–37. https://doi.org/10.18637/jss.v040.i04

Geriş A., & Kulaksız T. (2025). Predicting teachers' intentions to use virtual reality in education: a study based on the UTAUT-2 framework. *Research in Learning Technology*, *33*. https://doi.org/10.25304/rlt.v33.3429

Germeijs, V. & Verschueren, K. (2006). High school students' career decision-making process: Development and validation of the Study Choice Task Inventory. *Journal of Career Assessment,* 14(4), 449-471. https://doi.org/10.1177/1069072706286510

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. Journal of Educational Psychology, 100, 105–122. https://doi.org/10.1037/0022-0663.100.1.105.

Huizinga, T., de Vries, S., de Vos, M., te Lintelo, L., & Stuijver, S. (2022). *Flexibilisering in het hoger onderwijs: studentbehoeften.* Versnellingsplan Onderwijsinnovatie met ICT. https://www.versnellingsplan.nl/publicaties/flexibilisering-in-het-hoger-onderwijs-studentbehoeften/

Jonker, H., März, V., and Voogt, J. (2020). Curriculum flexibility in a blended curriculum. *AJET*. https://doi.org/10.14742/ajet.4926

Mayer, R. E., Makransky, G., & Parong, J. (2022). The promise and pitfalls of learning in immersive Virtual Reality. *International Journal of Human–Computer Interaction*, *39*(11), 2229–2238. https://doi.org/10.1080/10447318.2022.2108563

Meens, E. (2020). *Een leven lang kiezen* (1st ed.). Boom.

Ministerie van Onderwijs, Cultuur en Wetenschap. (2019). *Strategische agenda hoger onderwijs en onderzoek*. https://www.rijksoverheid.nl/documenten/publicaties/2019/12/02/strategische-agenda-hoger-onderwijs-en-onderzoek

Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, *31*(3), 274–295. https://doi.org/10.1007/s00357-014-9161-z

Saqr, M., López-Pernas, S., Helske, S., Durand, M., Murphy, K., Studer, M., & Ritschard, G. (2024). Sequence analysis in education: Principles, technique, and tutorial with R. *Learning Analytics Methods and Tutorials: A Practical Guide Using R*, 321–354. https://doi.org/10.1007/978-3-031-54464-4_10

Schulz, P., Schulte, J., Raube, S., Disouky, H., & Kandler, C. (2018). The role of leisure interest and engagement for subjective well-being. *Journal of Happiness Studies, 19*(4), 1135–1150.

Schwartz, B. (2016). *The Paradox of Choice: Why More Is Less* (Revised ed.). HarperCollins Publishers Inc.

Scott, S. G. & Bruce, R. A. Decision-Making Style: The Development and Assessment of a New Measure. *Educational and Psychological Measurement* 55, 818–831 (1995).

Smarius, J., & Van Der Stappen, E. (2022). Designing a choice support system for students in flexible curricula. *Proceedings of the 11th Computer Science Education Research Conference*, 39–45. https://doi.org/10.1145/3569173.3569178

Smith J., & Oliver M. (2005). Exploring behaviour in the online environment: student perceptions of information literacy. *Research in Learning Technology*, *13*(1). https://doi.org/10.3402/rlt.v13i1.10972

Tao, X. (2019). Virtual and augmented reality enhanced by touch. *Nature*, *575*(7783), 453–454. https://doi.org/10.1038/d41586-019-03506-3

Valle Torre, M., Oertel, C., & Specht, M. (2024). The Sequence Matters in Learning—A Systematic Literature Review. *Proceedings of the 14th Learning Analytics and Knowledge Conference, LAK '24*, 263–272. https://doi.org/10.1145/3636555.3636880

Van Brussel, S., Smarius, J., & van der Stappen, E. (2025). Blended begeleiding in flexibel onderwijs: Het ontrafelen van ontwerpprincipes voor de begeleiding van studenten bij inhoudelijke keuzevrijheid in het hoger onderwijs. *Tijdschrift OnderwijsPraktijk Studies*. https://doi.org/10.54657/TOPS.19623

van Rossen, S., Kluijtmans, E., van Brussel, S., van Harsel, M., Smarius, J., & van der Stappen, E. (2024). Recommender Systems for Students in Flexible Education: An Exploration of Benefits and Risks. In A. Bramwell-Dicks, A. Evans, M. Winckler, H. Petrie, & J. Abdelnour-Nocera (Eds), *Design for Equality and Justice* (pp. 196–204). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-61698-3_18

Wade, W., Hodgkinson, K., Smith, A., & Arfeld, J. (1994). Introduction. In W. Wade et al. (Eds.), *Flexible learning in higher education* (pp. 12–17). Routledge. https://doi.org/10.4324/9781315041575

Ward Jr., J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Zimmermann, G., Lannegrand-Willems, L., Safont-Mottay, C., and Cannard, C. (2015). Testing New Identity Models and Processes in French-speaking Adolescents and Emerging Adults Students. *J Youth Adolescence 44*(1), 127–141. DOI: 10.1007/s10964-013-0005-7

Appendix A

Filters used in the tool.

Domain	Location	Societal issues	Level	Credits	Period	Type of cooperation
Economics Behaviour & society Healthcare Education Law Language & culture Technology Cross-sectoral	4 Dutch cities	Biobased economy Broad prosperity and new entrepreneurship Arts Healthcare & wellbeing Safe and resilient city	NQLF-5 NQLF-6 NQLF-7	15EC 30EC 45EC 60EC	1 2 3 4	Interdisciplinary Monodisciplinary Multidisciplinary Transdisciplinary

Appendix B

Survey questions and answer categories.

Q1	How many items were on your first shortlist (NQLF 5)	Answer 1 – 10
Q2	How many items were on your second shortlist (NQLF 6)	Answer 1-10
Q3	What were important reasons for you to choose the courses you shortlisted? Please rank.	8 ranking options
Q4	What other important reasons were for you to choose the courses you shortlisted? Please rank.	6 ranking options
Q5	What are the top 3 reasons why courses did not make it onto your shortlist?	Open answer field
Q6	Drawing up a shortlist of elective courses was a first step in the process of choosing courses. What information do you still need in the follow-up process to make a well-considered choice for one course for which you would like to register?	6 options, multiple answers possible.
Q7	If you indicated that you wanted more detailed information about the course, can you indicate what information you would like and why this information?	Open answer field
Q8	If you have indicated that you would like to have a conversation about the courses on your shortlist, who would you like to have a conversation with and what should this conversation be about? Can you also indicate why this person(s)?	Open answer field
Q9	What is your opinion about the usability of the tool when compiling your shortlist? What could be better?	Open answer field
Q10	Which search options in the tool are you still missing? Why that one in particular?	Open answer field
Q11	Which course are you following and which to which academy does this course belong?	Open answer field
Q12	Were there any courses on your shortlist that were offered outside your own program but within your own domain?	Open answer field
Q13	Were there any courses on your shortlist that were offered outside your own domain?	Open answer field
Q14	What else would you like to say about this optional task and the possibility of choosing courses during your training at XX? For example, do you have any advice?	Open answer field

Structural Governance Shortcomings in Business Model Canvas Use

Jos van Capelleveen,[1] Guy Janssens,[1]
Rogier van de Wetering,[1] Natasha Alechina,[2]
Gideon Maillette de Buy Wenniger[2]

[1] Open Universiteit, Department of Information Science, Heerlen, the Netherlands
jos.vancapelleveen@ou.nl, guy.janssens@ou.nl, rogier.vandewetering@ou.nl
[2] Open Universiteit, Department of Computer Science, Heerlen, the Netherlands
natasha.alechina@ou.nl, gideon.maillettedebuywenniger@ou.nl

The Business Model Canvas (BMC) is widely used to structure strategic dialogue, yet recurring problems with the organisation and continuity of BMC initiatives remain under-theorised. This study examines structural governance shortcomings in BMC practice and their implications for organisational learning. Drawing on twenty-three semi-structured interviews across multiple sectors, the analysis identifies recurring breakdown patterns characterised by episodic use, fragmented ownership, limited feedback routines, weak validation practices, and insufficient technological support. These shortcomings reflect governance misalignments rooted in coordination dynamics, accountability structures, and interpretive processes rather than in the artefact's technical limitations. The study contributes an empirically grounded typology of governance shortcomings and outlines a minimal governance configuration to embed BMC processes more sustainably within organisational routines.

DOI
https://doi.org/10.18690/um.fov.4.2026.13

ISBN
978-961-299-152-4

Keywords:
business model canvas, process governance, business model innovation, governance mechanisms, organisational learning

University of Maribor Press

1 Introduction

Contemporary organisations operate in dynamic environments (Oberländer et al., 2025; Teece, 2010). To sustain competitiveness, they must continuously adapt strategic orientations, organisational processes, and structural configurations (Ahokangas & Myllykoski, 2014; Krüger & Teuteberg, 2018). In many sectors, these adaptations result in digital transformation, as digital technologies reshape value creation logics and coordination mechanisms (Krüger & Teuteberg, 2018; Liu et al., 2025; Vial, 2019). Organisations address this by using conceptual models that provide structure and orientation (Caccamo et al., 2023; Holtmann et al., 2025). In information systems (IS) research, such models are represented as business models and enterprise models. Business models capture the external logic of value creation (Teece, 2010; Wirtz, 2020). Enterprise models describe the capabilities, processes, and governance that realise these choices and, where relevant, represent inter-organisational interfaces and ecosystem dependencies (Lantow et al., 2022; Tsai et al., 2022). They link external logics with internal structures and cross-boundary coordination (Tsai et al., 2022). In organisational practice, one of the most widely used tools for articulating and discussing business models is the Business Model Canvas (BMC) (Osterwalder & Pigneur, 2010). Purposeful adaptation of business models to changing environments is referred to as business model innovation (BMI). BMI is widely regarded as a key mechanism for sustaining competitiveness in dynamic settings. Within IS research, and particularly in the domain of BMI, the BMC is frequently referenced and widely applied in both research and practice (George & Bock, 2011; Kamariotou & Kitsios, 2022). Besides modelling, it structures discussions and legitimises decisions (De Man & Luvison, 2019; Perkmann & Spicer, 2010). Yet studies report recurring weaknesses: canvases are often applied as static templates with limited iteration, fragmented ownership, and weak feedback loops (Demil & Lecocq, 2010; Massa & Tucci, 2013; Wirtz & Daiser, 2018). At the same time, such patterns are not necessarily problematic in themselves, as BMC use is often inherently episodic and strategically oriented (Teece, 2010; Wirtz & Daiser, 2018). However, they become analytically relevant when they constrain organisational learning, continuity across modelling iterations, and cross-organisational coordination (Van Capelleveen et al., 2025). Despite its widespread adoption, limited research has systematically examined which structural shortcomings characterise the governance of BMC use in organisational practice. This gap concerns the governance of BMC use, understood as the routines,

ownership structures, information flows, and informal dynamics through which meaning, power, and legitimacy are negotiated during modelling activities (Cabantous & Gond, 2011; Kaplan, 2011). Prior studies describe BMC use and business model innovation processes but provide limited explicit conceptualisation of how these processes are structured and governed over time (Andreini et al., 2022; Foss & Saebi, 2017). Recent synthesis work has begun to consolidate recurring patterns, including undocumented iteration, fragmented ownership, and weak feedback mechanisms, yet it has not explicitly conceptualised their governance implications (Van Capelleveen et al., 2025). Rather than assuming that the identified patterns reflect inherent limitations of the BMC as an artefact, this study interprets them as manifestations of governance arrangements that shape how modelling activities are organised and sustained. This study addresses this gap by analysing recurring shortcomings in the organisational embedding of BMC use from a governance perspective. The guiding research question is: *What structural shortcomings characterise the governance of Business Model Canvas (BMC) use in organisational practice?*

2 Related work

The related work is organised into two streams: (1) fragmentation and process variation in business model innovation (BMI) research, and (2) recurring shortcomings in Business Model Canvas (BMC) use and their governance implications. Research on business models has produced numerous conceptual approaches across IS and management studies (Wirtz, 2020; Zott & Amit, 2010). This diversity reflects the conceptual fragmentation of the field across strategic, organisational, technological, and ecosystem perspectives (Foss & Saebi, 2017). As a result, the business model has become a central concept for articulating how organisations create and capture value (Teece, 2010). Subsequent research therefore shifted towards processes of business model innovation (BMI) (Chesbrough, 2007; Massa & Tucci, 2013). BMI research shows considerable variation in how innovation processes are conceptualised. Some studies describe structured phases such as analysis, ideation, prototyping, and implementation, whereas others emphasise the situated and negotiated character of BMI practices (Ahokangas & Myllykoski, 2014; Foss & Saebi, 2017; Wirtz & Daiser, 2018). These perspectives contrast rational design models with practice-based accounts highlighting improvisation, sensemaking, and political dynamics (Perkmann & Spicer, 2010; van Nistelrooij & Homan, 2019). However, limited attention has been paid to how specific business

model tools used in BMI processes are organised and governed in practice. Business model tools not only represent strategic choices but also structure participation, documentation, and follow-up in business model development. Within this context, the Business Model Canvas (BMC) builds on the business model ontology originally proposed by Osterwalder (2004). The BMC has since emerged as one of the most widely adopted tools for representing and analysing business models in both research and practice (George & Bock, 2011; Kamariotou & Kitsios, 2022). Conceptually, the BMC can be understood as a boundary object that coordinates heterogeneous actors around a shared representation of value creation. However, this coordinating function depends in part on how the tool is embedded within organisational routines rather than on the canvas structure alone. Empirical studies report recurring shortcomings in how the BMC is applied. Research identifies deficits such as undocumented iteration, fragmented ownership, and missing feedback mechanisms (Demil & Lecocq, 2010; Massa & Tucci, 2013). A recent review consolidates these into four structural shortcomings: undocumented iteration, absent feedback loops, fragmented ownership, and insufficient ecosystem reflection (Van Capelleveen et al., 2025). These shortcomings directly affect governance of BMC use by limiting transparency, accountability, and organisational learning across modelling iterations. From a governance perspective, modelling practices often exhibit ambiguous ownership structures and limited rationale documentation, echoing broader concerns in enterprise modelling regarding accountability and traceability (Koç et al., 2018). Governance mechanisms in modelling environments remain comparatively immature (Pappas et al., 2023). In addition, performative perspectives emphasise that management tools such as the BMC do not merely represent organisational realities but actively shape discussions and decisions (Cabantous & Gond, 2011; Kaplan, 2011). However, BMC research has rarely conceptualised these recurring shortcomings as structural governance challenges embedded in organisational routines. While deficiencies such as fragmented ownership and missing feedback loops have been documented, they have rarely been analysed as a coherent configuration of governance mechanisms shaping how the BMC is organised and sustained in practice. The literature identifies recurring weaknesses in BMC use but does not explain how they jointly shape governance and continuity. An empirically grounded analysis of governance shortcomings in BMC use is therefore required to understand how these breakdown patterns emerge and how they constrain continuity and organisational learning in practice.

3 Methods

3.1 Study design

This study sought detailed, practice-based testimonies on practitioners' experiences of governing BMC use, including concrete routines for decision rights, ownership, collaboration, iteration, feedback, and documentation, as well as their interpretations of recurring shortcomings. Given the research focus, people rather than documents were the primary information sources. Internal documents rarely captured day-to-day governance breakdowns, and such shortcomings were often articulated retrospectively by those involved. Therefore, we targeted experts with direct BMC application experience, specifically consultants who facilitated BMC work and organisational stakeholders (e.g., product owners, innovation leads, strategy managers) who participated in or sponsored BMC-driven projects. To uncover coordination and governance shortcomings not yet fully articulated, a qualitative research approach was deemed most appropriate. Because we sought to document and explain recurring structural shortcomings in the governance of BMC use, semi-structured interviews were preferred to obtain rich, contextualised accounts and allow probing where necessary (Myers & Newman, 2007). Open-ended questions elicited concrete experiences and examples. The study adopts an abductive orientation, iteratively moving between empirical material and sensitising theoretical concepts to refine and stabilise categories of governance shortcomings. The coding procedure was structured according to the Gioia methodology (Gioia et al., 2013).

3.2 Participant selection and data collection

Participants were selected using a purposive sampling strategy to ensure that all respondents were actively involved in strategic or innovation projects in which the Business Model Canvas (BMC) played a central role. The sample comprised twenty-three professionals, including consultants, innovation managers, and project or programme leaders with 8–20 years of experience in business model design and digital transformation. Participants represented diverse sectors, including energy, ICT, financial services, agritech, creative industries, and public administration, spanning corporate organisations, public agencies, start-ups, and independent consultancies. Sampling continued until conceptual sufficiency was reached, meaning that additional interviews no longer contributed substantively new

governance-related themes relevant to the research question. This assessment was made during iterative coding and team discussions rather than defined a priori. The interview guide was derived from prior literature on BMC practice and documented governance shortcomings (Andreini et al., 2022; Kamariotou & Kitsios, 2022; Wirtz & Daiser, 2018) and was used to structure questions on decision-making, ownership, documentation, iteration, and feedback coordination (Demil & Lecocq, 2010; Wirtz & Daiser, 2018). The guide consisted of open-ended questions prompting respondents to describe the process flow, key decision points, collaboration mechanisms, and the rationale for documentation. Questions covered the governance chain from scoping and ownership through iteration, documentation, validation, and feedback, and examined breakdowns, perceived causes, consequences, and mitigation practices. Given the need to elicit concrete experiences while allowing analytical probing, a semi-structured interview format was used, balancing consistency across interviews with flexibility to explore contextual variation (Myers & Newman, 2007). The guide was pilot-tested once for clarity, and minor wording adjustments were made before the full application. Interviews lasted between 45 and 60 minutes and were conducted either via videoconference or on-site. All sessions were recorded with consent, transcribed verbatim, anonymised, and stored in a secure research environment prior to systematic coding.

3.3 Data analysis

The analysis followed a structured qualitative approach aligned with the Gioia methodology (Gioia et al., 2013). All transcripts were reviewed iteratively to identify statements related to BMC use, encountered shortcomings, and perceived consequences. During first-order coding, meaningful fragments were labelled to capture concrete actions, coordination routines, and perceived obstacles. Codes captured recurring coordination breakdowns, including missing ownership ("no one felt responsible for follow-up"), disrupted iteration ("each session restarted from scratch"), and information discontinuities ("lost notes" or "missing post-its"). Codes were continuously refined and merged across multiple rounds of analysis to ensure conceptual consistency. From the transcripts, 485 first-order fragments were identified and progressively abstracted into 67 first-order codes. These were then consolidated into ten second-order themes that captured structural shortcomings in the governance of BMC use. The coding process was iterative and abductive, moving

between the empirical material and emerging conceptual categories until stable thematic distinctions were achieved. After consolidation, the coding framework was systematically applied across all twenty-three interviews to ensure analytical consistency. A second researcher was involved in validating code definitions, category boundaries, and interpretation, rather than performing full independent recoding. This validation process ensured shared understanding of inclusion criteria and thematic distinctions. Reflexive memos documented analytical decisions and interpretive assumptions throughout the process. This data structure is reflected in the Results section, where second-order themes are presented as empirically grounded categories, and further synthesised into higher-level governance dimensions in Appendix 2. The trustworthiness of the study was strengthened through complementary strategies addressing credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985). Credibility and dependability were supported by transparent coding procedures, iterative category refinement, and researcher triangulation. An audit trail documented coding iterations, category revisions, and inclusion decisions. All participants provided informed consent, and transcripts were anonymised and securely stored in accordance with institutional ethical standards.

4 Results

This section presents empirical findings from twenty-three semi-structured interviews with consultants, innovation managers, project leaders, and policy professionals involved in Business Model Canvas (BMC) initiatives. From the transcripts, 485 first-order fragments were identified and assigned descriptive labels during first-order coding. This iterative process yielded 67 distinct codes, which were continuously refined through cross-case comparison and researcher reflection. The fragments were consolidated into ten second-order themes that capture structural governance shortcomings in BMC use. The analysis shows how individual experiences converged into recurring patterns across cases. The resulting data structure is reflected in the second-order themes below and further synthesised into higher-level governance dimensions in Appendix 2.

4.1 Overview of structural shortcomings in the governance of BMC use

The inductive coding process revealed ten recurring themes that capture structural shortcomings in how organisations govern the use of the Business Model Canvas (BMC). Appendix 1 (Table A1) summarises these shortcomings and their relative occurrence across the dataset. Fragment counts indicate the number of coded excerpts associated with each theme and reflect analytical salience within the dataset rather than statistical generalisability. Themes are presented in descending order of fragment counts. Themes 9* and 10* are treated as emerging observations because they were mentioned by a single respondent and therefore do not constitute structurally recurring patterns within the dataset. Together, these themes represent empirically grounded categories that link first-order observations to broader governance-related patterns. **Theme 1** highlights the most salient shortcoming: the absence of cumulative learning and structured feedback loops. Respondents reported that BMC workshops rarely produced cumulative learning cycles. Once completed, canvases were rarely revisited or compared with earlier versions. As one respondent explained, "We make one every year, but we never look back." As a result, modelling activities were frequently restarted rather than building cumulative organisational understanding. **Theme 2** concerns fragmented ownership and weak governance arrangements. Ownership of the BMC process was often diffuse, and no single actor was responsible for maintaining versions, ensuring follow-up, or linking the canvas to ongoing strategy execution. One respondent noted that "after the consultant leaves, nobody feels responsible for it anymore." Consequently, the BMC often remained an event-based artefact rather than an institutionalised coordination mechanism. **Theme 3** shifts attention to the internal quality of modelling discussions, revealing an emphasis on form over substance. Participants often prioritised completing the template over critically examining its underlying assumptions. As one participant remarked, "It looks nice on the wall, but it doesn't really change anything." In such cases, the canvas served more as a procedural ritual than as a space for analytical reflection. **Theme 4** reveals considerable variability in BMC application across organisational contexts. Teams frequently adapted terminology and the model's interpretation to local project language or departmental practices. While this flexibility enabled contextual tailoring, it reduced comparability across initiatives and weakened opportunities for cross-organisational learning. The absence of shared interpretive standards limited the development of consistent governance routines. **Theme 5** highlights power and consultancy-driven process

dynamics. Workshops were sometimes dominated by consultants or senior managers, limiting broader participation in modelling discussions. One respondent described such sessions as "consultant theatre; we just nodded along." These dynamics reinforced dependence on external facilitation and constrained internal ownership of modelling outcomes. **Theme 6** identifies insufficient validation and limited knowledge sharing as structural weaknesses. Canvases were often developed through internal reasoning rather than through empirical testing with customers, partners, or external data. As one respondent admitted, "We never check if what we wrote is actually true." Consequently, assumptions remained largely unchallenged, and experiential knowledge stayed confined to individual projects. **Theme 7** concerns the operational drift of modelling efforts. In many organisations, BMC work shifted rapidly from exploration to execution. Once initiatives moved into implementation, the canvas often lost visibility and relevance. This weakened the link between modelling activities and subsequent learning from implementation. **Theme 8** addresses the limited technological support surrounding BMC processes. Participants reported relying primarily on analogue tools, such as sticky notes or static PowerPoint templates, without systematic version tracking. The lack of structured documentation and version control contributed to fragmentation and information loss. Digital tools, where available, were rarely integrated into formal governance routines. **Theme 9** identifies a performative use of the BMC as a legitimising ritual. In isolated instances, the canvas served primarily to signal strategic competence rather than to support collective learning. In such situations, modelling functioned more as managerial theatre than as a mechanism for structured organisational reasoning. Finally, **Theme 10** concerns limited stakeholder involvement and ecosystem engagement. External actors, such as customers or partners, were rarely structurally included in modelling activities. When modelling remained internally oriented, the canvas could not serve as an ecosystem coordination device. Governance arrangements in such cases appeared weakly embedded within broader stakeholder networks.

5 Discussion

This study examined structural governance shortcomings in the use of the Business Model Canvas (BMC). The findings show that coordination, validation, and learning remain weakly institutionalised across organisations. Building on the ten shortcomings summarised in Appendix 1 (Table A1), the analysis indicates that

modelling practices are often episodic and weakly embedded within organisational routines. Taken together, the shortcomings reveal a recurring configuration rather than isolated deficiencies, pointing to a broader governance pattern that shapes how BMC activities are organised and sustained in practice. This pattern can be understood through governance arrangements that combine formal coordination mechanisms, such as accountability and traceability (Koç et al., 2018), with informal processes of influence and meaning-making (Cabantous & Gond, 2011; Kaplan, 2011). Viewed through this lens, the identified shortcomings cluster around two interrelated dimensions. The first concerns formal and procedural mechanisms oriented towards control, traceability, and accountability. The second concerns social and interpretive processes centred on meaning-making, influence, and legitimacy. **Figure 2 in Appendix 2** offers a conceptual synthesis of these governance dimensions, as emerging from the empirical findings. When these dimensions are weakly aligned, modelling processes remain episodic and fail to support cumulative organisational learning. The formal–procedural dimension concerns mechanisms that ensure continuity over time, such as version control, rationale documentation, and feedback routines. The absence of such mechanisms leaves modelling practices weakly connected to experimentation and organisational learning, echoing earlier observations in enterprise modelling about accountability and traceability (Koç et al., 2018) and the need for structured design principles (Aier et al., 2011). Limited feedback loops reduce validation and follow-up, causing modelling to become detached from implementation and learning (Demil & Lecocq, 2010; Wirtz & Daiser, 2018). As a result, modelling outputs remain weakly embedded in organisational memory. The social–interpretive dimension captures the interactional processes by which participants negotiate meaning and legitimacy during modelling activities. Within this dimension, ownership is often diffuse, facilitation may substitute for commitment, and external perspectives are rarely integrated. Power asymmetries and local vocabularies shape what becomes visible on the canvas and what remains implicit (Culot & Battistella, 2024; Tsai et al., 2022). Respondents frequently described modelling sessions as producing convincing yet transient outcomes, sometimes resembling consultant theatre (Perkmann & Spicer, 2010). Taken together, the findings indicate that recurring shortcomings cannot be explained by artefact design or facilitation style alone, but are embedded in governance arrangements that structure continuity, accountability, and reflection. Prior studies have repeatedly observed limited iteration and feedback in business model innovation but have provided little explanation of their organisational origins.

While Demil and Lecocq (2010) and Wirtz and Daiser (2018) characterised these shortcomings as procedural weaknesses, the present study suggests that they reflect structural governance misalignments that fragment coordination, accountability, and learning over time. The findings also align with performative perspectives on modelling, which argue that management tools not only represent organisational realities but also actively shape behaviour and decision-making (Cabantous & Gond, 2011; Perkmann & Spicer, 2010). In practice, the act of "completing the canvas" may symbolise rational planning and control, substituting for sustained reflection and illustrating how modelling can function as a legitimising practice rather than a cumulative learning mechanism. Building on these insights, recurring deficiencies in BMC use can be understood as governance-related rather than artefact design limitations alone. By distinguishing structural coordination mechanisms from interpretive dynamics, the study develops a governance perspective that clarifies how formal routines and social processes jointly shape modelling outcomes. Structural mechanisms concern ownership, traceability, documentation, and continuity across iterations. Interpretive dynamics concern dialogue, participation, influence, and negotiated meaning. When these dimensions are weakly aligned, modelling practices remain weakly embedded in organisational learning processes. The study's contribution is to conceptualise these recurring shortcomings as a governance configuration underpinning BMC use. This configuration can be understood as the minimal governance framework required for sustained BMC use, comprising explicit ownership structures, traceability and documentation routines, feedback and validation mechanisms, and scheduled review cycles. This shifts analytical attention from artefact adequacy to organisational embedding and helps explain the persistence of discontinuous BMC use despite its widespread diffusion. A recurring theme was the limited technological support for BMC processes. Participants noted the absence of structured documentation systems, version control mechanisms, and shared repositories linking canvases across iterations. These observations focused primarily on continuity, traceability, and coordination rather than technological innovation itself. Participants framed digital and AI-based tools as exploratory support mechanisms for documentation, transparency, and analytical reflection, rather than as autonomous remedies. The findings indicate that improving business model innovation processes requires explicit attention to governance arrangements for using the Business Model Canvas. Organisations should clarify ownership structures, define responsibilities for follow-up, and establish routines for revisiting and validating assumptions over time. Without such mechanisms, the BMC remains

an event-based artefact rather than an institutionalised coordination device. Sustained innovation capability depends on institutionalising governance routines that integrate structural discipline with reflective dialogue.

6 Conclusion, limitations and future research

This study examined structural shortcomings in the governance of Business Model Canvas (BMC) use within organisational practice. Drawing on twenty-three interviews, it identified ten governance-related shortcomings that constrain continuity, accountability, and organisational learning. The findings show that recurrent deficiencies in BMC practice are less attributable to artefact design than to governance misalignments that shape ownership, documentation, validation, and temporal embedding. These shortcomings are consistently characterised by fragmented ownership, missing feedback loops, weak documentation, and limited validation, which together reflect a broader governance misalignment. By conceptualising these patterns as structural governance conditions, the study shifts analytical focus from facilitation techniques and artefact design to organisational embedding and governance arrangements. This suggests that BMC use is not primarily a modelling problem but a governance problem, in which weakly defined roles, coordination routines, and feedback structures impede cumulative organisational learning. The study has several limitations. First, the findings are based on qualitative interviews and therefore do not permit conclusions about statistical prevalence. Although the respondents represented diverse sectors and organisational roles, the results reflect interpretive accounts of practice rather than direct observation of modelling processes. Future research should examine how governance routines around BMC use evolve over time and under which organisational conditions modelling becomes embedded as a sustained learning process. Further research should also explore how digital infrastructures and collaborative modelling environments support traceability, coordination, and collective sensemaking in business model innovation, and how such mechanisms can operationalise governance structures across iterative modelling cycles.

Acknowledgment

We gratefully acknowledge Anne Lhoest, Maarten Hermans, Daan Knobbe, Vinesh Doekharan, and Ouiam Bousalam, Master of Science students at Open Universiteit, whose graduation projects provided valuable empirical and conceptual input for this research.

References

Ahokangas, P., & Myllykoski, J. (2014). The Practice of Creating and Transforming a Business Model. *Journal of business models*, *2*(1), 6. https://doi.org/10.5278/ojs.jbm.v2i1.719

Aier, S., Fischer, C., & Winter, R. (2011). *Construction and evaluation of a meta-model for enterprise architecture design principles*. 10th International Conference on Wirtschaftsinformatik (WI 2011),

Andreini, D., Bettinelli, C., Foss, N. J., & Mismetti, M. (2022). Business model innovation: a review of the process-based literature. *Journal of Management and Governance*, *26*(4), 1089-1121. https://doi.org/10.1007/s10997-021-09590-w

Cabantous, L., & Gond, J.-P. (2011). Rational decision making as performative praxis: Explaining rationality's Éternel Retour. *Organization Science*, *22*(3), 573-586. https://doi.org/10.1287/orsc.1100.0534

Caccamo, M., Pittino, D., & Tell, F. (2023). Boundary objects, knowledge integration, and innovation management: A systematic review of the literature. *Technovation*, *122*, 102645. https://doi.org/10.1016/j.technovation.2022.102645

Chesbrough, H. (2007). Business model innovation: it's not just about technology anymore. *Strategy & leadership*, *35*(6), 12-17.

Culot, G., & Battistella, C. (2024). Future ecosystem business model tool: Design science and field test in the efuel ecosystem towards the sustainability transition. *Technological Forecasting and Social Change*, *208*, 123682.

De Man, A.-P., & Luvison, D. (2019). Collaborative business models: Aligning and operationalizing alliances. *Business Horizons*, *62*(4), 473-482.

Demil, B., & Lecocq, X. (2010). Business model evolution: In search of dynamic consistency. *Long range planning*, *43*(2-3), 227-246. https://doi.org/10.1016/j.lrp.2010.02.004

Foss, N. J., & Saebi, T. (2017). Fifteen years of research on business model innovation: How far have we come, and where should we go? *Journal of management*, *43*(1), 200-227. https://doi.org/10.1177/0149206316675927

George, G., & Bock, A. J. (2011). The business model in practice and its implications for entrepreneurship research. *Entrepreneurship theory and practice*, *35*(1), 83-111.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational research methods*, *16*(1), 15-31.

Holtmann, J., Horkoff, J., Wohlrab, R., Vu, V., Kasauli, R., Maro, S., Steghöfer, J.-P., & Knauss, E. (2025). Using boundary objects and methodological island (BOMI) modeling in large-scale agile systems development. *Software and Systems Modeling*, *24*(1), 183-207. https://doi.org/10.1007/s10270-024-01193-x

Kamariotou, M., & Kitsios, F. (2022). Bringing Digital Innovation Strategies and Entrepreneurship: The Business Model Canvas in Open Data Ecosystem and Startups. *Future Internet*, *14*(5).

Kaplan, S. (2011). Strategy and PowerPoint: An inquiry into the epistemic culture and machinery of strategy making. *Organization Science*, *22*(2), 320-346.

Koç, H., Sandkuhl, K., Stirna, J., & Kuhr, J.-C. (2018). Capability as a service: method and tool support for context-aware business services. *International Journal of Information Systems in the Service Sector (IJISSS)*, *10*(3), 64-84.

Krüger, N., & Teuteberg, F. (2018). Consulting Business Models in the Digital Era. MKWI,

Lantow, B., Sandkuhl, K., & Stirna, J. (2022). Enterprise modeling with 4EM: perspectives and method. In *Domain-specific conceptual modeling: Concepts, methods and ADOxx tools* (pp. 95-120). Springer. https://doi.org/10.1007/978-3-030-93547-4_5

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. sage.

Liu, J., Fang, Y., & Ma, Y. (2025). The Driving Impact of Digital Innovation Ecosystems on Enterprise Digital Transformation: Based on an Interpretable Machine Learning Model. *Sustainability*, *17*(13), 5898. https://doi.org/10.3390/su17135898

Massa, L., & Tucci, C. L. (2013). Business model innovation. *The Oxford handbook of innovation management*, *20*(18), 420-441. https://doi.org/10.1093/oxfordhb/9780199694945.013.002

Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and organization*, *17*(1), 2-26.

Oberländer, A. M., Karnebogen, P., Rövekamp, P., Röglinger, M., & Leidner, D. E. (2025). Understanding the influence of digital ecosystems on digital transformation: The OCO (orientation, cooperation, orchestration) theory. *Information Systems Journal*, *35*(1), 368-413. https://doi.org/10.1111/isj.12425

Osterwalder, A. (2004). *The business model ontology: A proposition in a design science approach (Doctoral dissertation, Université de Lausanne).*

Osterwalder, A., & Pigneur, Y. (2010). Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers.

Pappas, D., Braunholtz-Speight, T., Hannon, M., Webb, J., Fuentes González, F., & Sharmina, M. (2023). Business models for smart local energy systems—A triple layered perspective. *Frontiers in Sustainable Energy Policy*, *1*.

Perkmann, M., & Spicer, A. (2010). What are business models? Developing a theory of performative representations. In *Technology and organization: Essays in honour of Joan Woodward* (pp. 265-275). Emerald Group Publishing Limited. https://doi.org/10.1108/S0733-558X(2010)0000029020

Teece, D. J. (2010). Business Models, Business Strategy and Innovation. *Long range planning*, *43*(2), 172-194. https://doi.org/10.1016/j.lrp.2009.07.003

Tsai, C. H., Zdravkovic, J., & Stirna, J. (2022). Modeling digital business ecosystems: a systematic literature review. *Complex Systems Informatics and Modeling Quarterly*(30), 1-30. https://doi.org/10.7250/csimq.2022-30.01

Van Capelleveen, J., Janssens, G., Van de Wetering, R., Alechina, N., & Maillette de Buy Wenniger, G. (2025). Business Model Canvas development process: A scoping review of process structures, weaknesses, and governance dynamics.

van Nistelrooij, A., & Homan, T. (2019). "You call it a structure taboo!" Relating consultant experiences to changing dynamics in healthcare. *Journal of Organizational Change Management*, *32*(2), 182-193.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems*, *28*(2), 118–144.

Wirtz, B., & Daiser, P. (2018). Business model innovation processes: A systematic literature review. *Journal of business models*, *6*(1), 40-58. https://doi.org/10.5278/ojs.jbm.v6i1.2397

Wirtz, B. W. (2020). *Business model management: design-process-instruments*. Springer Nature. https://doi.org/10.1007/978-3-030-48017-2

Zott, C., & Amit, R. (2010). Business Model Design: An Activity System Perspective. *Long range planning*, *43*(2), 216-226.

About the authors

Jos van Capelleveen is a doctoral researcher in the Department of Information Science at Open Universiteit. Guy Janssens is an Assistant Professor in the Department of Information Science at Open Universiteit. Rogier van de Wetering is Full Professor of Digital-Driven Transformation and Vice-Dean of Education in the Department of Computer Science at Open Universiteit. Natasha Alechina is Full Professor of Safe and Responsible AI in the Department of Computer Science at Open Universiteit. Gideon Maillette de Buy Wenniger is an Assistant Professor in the Department of Computer Science at Open Universiteit.

Appendix 1

Appendix 1 provides a comprehensive overview of the coded fragments associated with each identified theme. Table A1 summarises the ten structural shortcomings in the governance of Business Model Canvas (BMC) use, identified through the inductive coding process across the interview dataset. Fragment counts indicate the number of coded excerpts associated with each theme and reflect analytical salience within the dataset rather than statistical prevalence. Themes marked with an asterisk (*) represent emerging observations mentioned by a single respondent and therefore do not constitute structurally recurring patterns within the dataset.

Table A1: Recurring governance shortcomings in BMC use identified across interviews

Nr.	Second-order theme (n = 23 interviews, 485 coded fragments)	No. of coded fragments	No. of respondents mentioning theme
1	Lack of learning and feedback loops	154	18
2	Fragmented ownership and weak governance arrangements	70	18
3	Emphasis on form over substance in model use	51	16
4	Inconsistent and context-dependent BMC application	48	17
5	Power and consultancy-driven process dynamics	47	16
6	Insufficient validation and knowledge sharing	45	10
7	Operational focus with little reflective learning	44	11
8	Low level of technological support	19	9
9 *	Performative use of the BMC as a legitimising ritual	6	1
10 *	Low stakeholder involvement and ecosystem engagement	1	1

Appendix 2

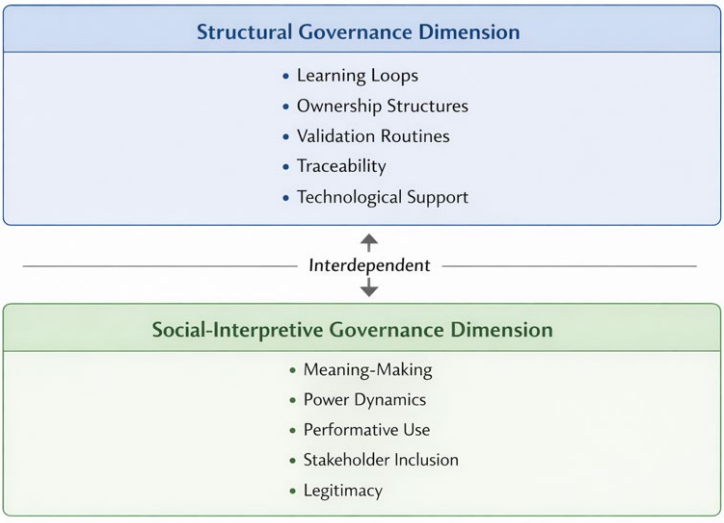

Figure 2: Conceptual synthesis of governance dimensions emerging from the ten identified themes

Source: Authors

An Identity-Oriented Framework for AI-Supported mHealth

Esther Federspiel

Eastern Switzerland University of Applied Sciences, School of Management, St. Gallen, Switzerland
esther.federspiel@ost.ch

Metabolic chronic conditions require sustained lifestyle change, yet mobile health (mHealth) interventions typically show modest and short-lived effects. A central unresolved question is how behavior change becomes personally meaningful. Beyond meaningful gamification, a growing body of literature emphasizes identity processes as crucial for long-term change: health behaviors persist not merely through motivation or prompts, but when narratively integrated into the self-concept. In the age of large language models (LLM)-based conversational systems, this perspective becomes urgent. While LLMs can scaffold reflection at scale, it may normatively script users' developmental trajectories. Existing work has addressed related questions around identity, reflection, and behavior change, but an integrative framework for identity-oriented mHealth design remains underdeveloped. This paper proposes an identity-oriented framework for designing mHealth systems that support narrative identity development and sustained behavior change. The paper organizes empirically reported design patterns along four narrative identity processes (interpretation, temporal integration, episodic configuration, and self-representation) to offer a design-relevant conceptual lens for LLM-supported mHealth systems.

DOI
https://doi.org/10.18690/um.fov.4.2026.14

ISBN
978-961-299-152-4

Keywords:
identity narratives, meaningful gamification, mHealth, intrinsic motivation, self regulated behavior

University of Maribor Press

1 Introduction

Healthcare systems are facing steadily increasing costs, largely driven by chronic conditions that require long-term management rather than episodic treatment. In Switzerland alone, total healthcare costs were estimated at 97.1 billion Swiss francs (CHF) in 2024, corresponding to CHF 10,783 per capita (Swiss Health Observatory Obsan, 2025). Many metabolic conditions, including type 2 diabetes and obesity, are at least partly preventable through sustained lifestyle patterns such as regular physical activity, healthier eating, and stress regulation (Chen et al., 2025). Mobile health (mHealth) applications promise scalable support for lifestyle change, and thousands of behavior-change apps are available across platforms. However, recent reviews indicate that their effects on weight and metabolic indicators are generally modest, heterogeneous, and often limited in long-term follow-up (Da Couto et al., 2025; Leuzzi et al., 2026; Pujia et al., 2025). High dropout rates and relapses into previous behavioral patterns remain persistent problems. This suggests that engagement alone is insufficient; rather, sustained change may require deeper psychological integration.

Meta-analytic evidence suggests that identity is both associated with and predictive of health-related behaviors (De Hoog & Pat-El, 2024; Di Zhu et al., 2024). Empirical studies on physical activity show associations between behavior and activity identity, yet findings on identity change are inconsistent. Wierts et al. (2024) report only small and unstable identity gains among beginners, whereas Priebe et al. (2020) found significant increases in running identity correlated with training frequency. Rhodes (2025) therefore emphasizes that identity development is dynamic and does not automatically result from repeated behavior. This evidence suggests that identity should be conceptualized not merely as an outcome, but as an actively supported developmental process.

This gains renewed significance with the rise of LLM-based systems, which enable scalable, dialogical exploration and co-construction of identity narratives (Meyer & Elsweiler, 2025; Rahsepar Meadi et al., 2025). Context-sensitive artificial intelligence (AI) journaling deepens meaning-making by prompting linguistic interpretation rather than mere documentation (Nepal et al., 2024), while narrative progress framing and immersive reminiscence systems may support broader autobiographical integration (Murnane et al., 2020; Tran et al., 2022; Xu et al., 2025). Without a solid psychological and conceptual foundation, however, such developments risk giving

rise to significant safety and ethical concerns (ibd.). At the same time, such developments raise important safety and ethical concerns. If reflective processes are overly structured or tied to optimized behavioral pathways, AI-supported systems may unintentionally constrain users' interpretive autonomy and self-authorship (Chandra et al., 2025).

In this paper, LLM-supported mHealth refers primarily to user-facing conversational and reflective features, especially LLM-enabled journaling, coaching, and narrative prompting. Backend analytics, prediction-only AI, and non-interactive personalization are outside the main scope. The framework is further intended for contexts in which sustained lifestyle change and reflective sense-making are central, rather than for brief or purely transactional interventions.

Against this backdrop, the paper makes three contributions. First, it proposes narrative identity as a useful mediating lens between mHealth design and sustained health behavior, complementing existing behavior change perspectives with a stronger focus on long-term internalization. Second, it offers a conceptual framework that organizes identity-oriented design knowledge along four narrative processes: interpretation, temporal integration, episodic configuration, and self-representation. Third, it presents an initial taxonomy of empirically reported design patterns that may inform the development of mHealth systems intended to scaffold identity work rather than prescribe predefined optimization trajectories.

Methodologically, the paper does not present a systematic review but a structured conceptual synthesis of empirical work across Human-Computer Interaction (HCI), mHealth, lifelogging, digital reminiscence, and LLM-supported journaling. Studies were considered insofar as they described or evaluated interactive features related to reflection, narrative reconstruction, temporal sense-making, or identity-relevant self-representation. The aim is not exhaustive coverage, but theory-informed integration and design abstraction.

2 Theoretical Background

To theoretically ground identity-oriented persuasive design, behavior change must be situated within broader models of self-construction, particularly considering emerging LLM-based conversational systems that increasingly mediate reflective

health interactions. This section therefore integrates four complementary perspectives: identity theories, behavior change models, meaningful gamification, and narrative identity in the context of LLM-based systems.

2.1 From Trait Models to Narrative Identity: A Theoretical Rationale

Several psychological frameworks conceptualize identity and its developmental trajectories across the lifespan. Trait- and status-based models, such as the Big Five and Marcia's identity status model, reify relatively stable characteristics or categorical outcomes (Costa & McCrae, 1992; Marcia, 1966) and thus underemphasize the situational, socially negotiated, and technologically mediated dynamics through which identities are enacted and revised. Social identity and interactionist perspectives (Mead, 1934; Goffman, 1959) illuminate belonging and contextual performance but are less suited to explaining how experiences are integrated into biographical coherence over time. Developmental models like Erikson's psychosocial theory (Erikson, 1968) offer a strong life-span heuristic yet remains difficult to operationalize empirically and only indirectly address digitally mediated identity construction.

Against this backdrop, McAdams' narrative identity framework is adopted (1996, 2001; McAdams & McLean, 2013), which conceptualizes identity as an internalized, evolving life story integrating past, present, and future into a coherent self-understanding. This approach is particularly suitable because it (1) treats identity as dynamic and reconstructive, (2) specifies analytically distinct narrative processes, such as meaning making and temporal integration, which can be translated into design relevant mechanisms, and (3) foregrounds interpretation and authorship, aligning with interactive systems that scaffold reflection rather than prescribe behavior.

2.2 An Identity-Based Perspective on Behavior Change

An identity-based perspective on behavior change does not replace established behavioral models but reinterprets them through the lens of self-construction. Fogg's Behavior Model explains how behavior is triggered in digital environments through the convergence of motivation, ability, and prompts, yet motivation itself already encompasses socially grounded drivers such as belonging and social

acceptance (Fogg, 2009). While identity is not explicitly modeled, it shapes the stability and qualitative character of motivation by influencing whether a behavior is experienced as self-relevant rather than situationally attractive.

Self Determination Theory deepens this view by distinguishing forms of regulation and demonstrating that behavior becomes durable when it is internalized and integrated into the self (Deci & Ryan, 1985; Ryan & Deci, 2000). Behavioral persistence is most likely when behavior aligns with core values and self-understanding. Narrative identity can, in this context, be linked to the concept of self-defining action. Michie's COM-B model (2011) complements these accounts by offering a systemic framework for diagnosing behavioral determinants (capability, opportunity, and motivation). Although identity is not explicitly theorized, it is embedded in psychological capability (beliefs about competence and role) and reflective motivation (values, intentions, self-evaluation). Identity thus operates as a mediating layer linking evaluative processes to self-definition.

Taken together, these models provide a multilevel account of behavioral enactment, motivational quality, and systemic intervention design. An identity-oriented framework builds on this foundation by conceptualizing narrative identity as the mechanism through which behavior becomes biographically integrated rather than merely situationally enacted.

2.3 Reframing Meaningful Gamification Through Narrative Identity

McAdams conceptualizes identity as an internalized and evolving life story through which individuals organize experiences, values, and goals into a temporally coherent self-understanding (McAdams, 1996, 2001; McAdams & McLean, 2013). This processual and interpretive account of identity provides a deeper theoretical foundation for analyzing gamified systems than approaches that focus primarily on motivation. Nicholson's (2014) concept of meaningful gamification already moves beyond reward-based mechanics by emphasizing personal relevance, autonomy, and reflection. However, its conceptualization of meaning remains theoretically underdeveloped with respect to identity formation and self-construction.

Narrative identity theory extends this perspective in three significant ways: First (1), it reframes meaning as temporally structured self-interpretation: beyond asking whether a goal is important "to me," it examines how repeated actions become integrated into an unfolding life story (McAdams, 1996, 2001; McAdams & McLean, 2013; Ricoeur, 1984). Sustained behavior depends on biographical integration, not momentary relevance. Second (2), it conceptualizes reflection as longitudinal and reconstructive, involving the integration of turning points, setbacks, and revised self-ascriptions rather than situational motive clarification. Identity work thus becomes an ongoing interpretive process. Third (3), it reveals the normative assumptions embedded in gamified systems. By privileging growth and optimization, many designs reinforce narrow self-improvement scripts while marginalizing alternative trajectories such as stability or care. Narrative identity theory makes these exclusions analytically visible.

Meaningful gamification can thus be linked to the four narrative processes proposed here: reflection and personal relevance support interpretation, journey framing supports temporal integration, quest or chapter structures support episodic configuration, and role-related framing supports self-representation. From this perspective, narrative identity extends meaningful gamification by framing persuasive systems not only as motivational architectures, but as infrastructures for narrative self-construction shaped by temporality, moral positioning, and sociocultural context.

2.4 Narrative Identity in the Age of LLM-Based Systems

LLM-driven agents are increasingly used in mobile health to provide adaptive coaching and reflective dialogue (Meyer & Elsweiler, 2025; Rahsepar Meadi et al., 2025). Unlike traditional gamification mechanisms, LLMs may function as infrastructures for identity work: they can scaffold meaning-making, support temporal integration, and assist in interpreting setbacks within a broader personal trajectory (Kvietkute & Ness, 2026). However, this capability also introduces a new form of normativity. Because LLMs are trained on large-scale cultural corpora, they may implicitly reproduce dominant scripts such as self-optimization and continuous improvement (Dokic et al., 2025). If conversational systems over-structure reflection or privilege specific develop-mental arcs, they risk shaping identity rather than supporting its autonomous construction (Chandra et al., 2025).

Taken together, these perspectives play different roles in the paper: narrative identity provides the core explanatory lens; behavior change models specify how sustained enactment becomes possible; meaningful gamification clarifies how motivational design can support personally relevant engagement; and LLM-based systems introduce both new scaffolding opportunities and new ethical risks.

3 Operationalizing Narrative Identity in HCI Design

In this section, the conceptual foundations underlying the taxonomy of narrative identity in HCI design are first systematically elaborated. Subsequently, representative examples from empirical literature are presented for each category to illustrate and substantiate the proposed framework.

3.1 Taxonomy of Narrative Identity in HCI Design

The proposed taxonomy builds on the assumption that identity is not a stable trait but an ongoing, constructed life story (McAdams, 1996, 2001; McAdams & McLean, 2013). Narrative identity theory conceptualizes identity as an internalized and evolving self-narrative through which individuals integrate experiences over time into a coherent sense of self (ibd.). This integrative process involves intertwined operations: interpreting and assigning meaning to events, structuring them temporally, and positioning the self as an acting figure within the story.

From this processual understanding, four design levers can be derived that translate core narrative operations into interactive systems. (1) **Prompting interpretation** supports meaning-making: events become identity-relevant only through reflective processing. Context-sensitive prompts or dialogical reminiscence systems can actively scaffold this interpretive work (Adler et al., 2016; Zimmerman et al., 2007). (2) **Structuring time** enables temporal integration by organizing past, present, and future into visible developmental trajectories, thereby supporting biographical continuity (McAdams, 1996; 2001; McAdams & McLean, 2013; Habermas & Bluck, 2000). (3) **Composing episodes** addresses configurational integration: fragmented experiences are condensed into meaningful episodes. Systems that curate artifacts or semi-automatically generate stories scaffold the process of narrative identity construction (Ricoeur, 1984). (4) **Representing the self**-targets narrative agency. Because life stories center on an acting protagonist, features such as identity-

congruent self-representation or avatar customization can reinforce autonomy and perceived efficacy (McAdams, 1996, 2001; McAdams & McLean, 2013; Yee & Bailenson, 2007).

In summary, the taxonomy does not categorize interface functions but core operations of narrative identity construction: **interpretation, temporal integration, episodic configuration, and agentic self-positioning**. It thus clarifies how interactive systems can scaffold identity work by deliberately shaping narrative processes rather than merely collecting data.

3.2 Empirically Grounded Design Patterns for Narrative Identity

To operationalize narrative identity in interactive systems, empirical research from HCI, lifelogging, digital reminiscence, LLM-supported journaling, immersive media, and mHealth interventions were synthesized, including studies that investigate reflective scaffolding and identity processes in digitally mediated health behavior change. We treated the studies in Table 1 as empirical instantiations of broader design patterns and abstracted them into four higher-level narrative processes based on their primary identity-related function.

Across these studies, a common pattern emerges. Systems become identity-supportive not when they merely capture behavioral data, but when they scaffold the transformation of experience into interpretable narrative units.

Context-sensitive reflection prompts, for example, encourage users to articulate personal significance rather than passively log events (Nepal et al., 2024). Timeline-based elicitation methods and revisitable journey views make developmental continuity visible and allow reinterpretation across time (Haimson et al., 2023). Narrative framing of progress transforms metrics into story-like trajectories, positioning the user as an active protagonist (Murnane et al., 2020). Similarly, lifelog curation tools and semi-automated story composition features help consolidate fragmented traces into coherent episodes (Byrne & Jones, 2009). Immersive reminiscence systems and narrative VR prompts further deepen episodic reconstruction by coupling emotional re-experiencing with guided interpretation (Tran et al., 2022; Xu et al., 2025). Finally, dialogic photo-based AI systems and structured memory scaffolds support elaboration and social sharing, enabling

identity to be co-constructed in interaction (Wang et al., 2024). By mapping each feature to its narrative contribution and empirical grounding, we provide a structured foundation for designing systems that actively support identity work.

Table 1: Design for Narrative Identity

Design-Feature	Contribution to Narrative Identity	Empirical Basis
Reflection (Context-sensitive reflection prompts (LLM-supported sensing-based prompts) (Nepal et al., 2024))	Foster meaning-making and coherence by prompting interpretation rather than mere logging.	MindScape: MindScape: contextual journaling prompts; users report enhanced self-reflection.
Temporal Structuring (Timeline/phase elicitation for identity transitions (structured retrospection) (Haimson et al., 2023))	Strengthen temporal coherence and agency through reinterpretation of turning points.	Technology-mediated elicitation of reflection on past identity transitions.
Temporal Structuring / Self-Representation (Narratively framed progress feedback (chapters/quests instead of metrics) (Murnane et al., 2020))	Support continuity (progress as storyline) and agency (self as protagonist).	WhoIsZuki: activity goals visualized as multi-chapter quests; users experience data as story.
Temporal Structuring (Revisitable journey/flashback views ("Previously on…") (Murnane et al., 2020))	Maintain biographical coherence by enabling reinterpretation of earlier episodes.	Narrative tracking setup emphasizes longitudinal story access.
Episodic Composition (Semi-automated story composition from lifelog artifacts + user curation (Byrne & Jones, 2009))	Transform fragments into meaningful episodes; support coherence.	Lifelog reduction/composition into narrative retrospectives.
Episodic Composition (Immersive virtual reality (VR) reminiscence (spatial re-experiencing) (Tran et al., 2022))	Facilitate episodic reconstruction and narrative reconsolidation.	VR reminiscence prototype: user study reports helpful yet sometimes overwhelming experience.
Reflection (Narrative identity prompts in VR story frames (e.g., "Who am I?") (Xu et al., 2025))	Trigger identity exploration and active meaning-making.	Someone else's Lifelog in Virtual Reality (SLIVeR): narrative VR with existential prompts and lifelog fragments.
Reflection (LLM-supported photo-based memory dialogues (Wang et al., 2024))	Support narrative elaboration and social identity through storytelling.	GoodTimes: multimodal photo album with dialogical AI.
Episodic Composition / Self-Representation (Memory scaffolding via dialogical photo AI + structured cues (who/what/when/where) + sharing preparation (Wang et al., 2024))	Enhance coherence through structured reconstruction and social narratability.	GoodTimes: development of LLM-supported multimodal photo album.

Source: Own elaboration

4 A Framework for Identity-Oriented mHealth Design

Building on the theoretical foundation outlined in Section 2 and the operationalization of narrative identity in Section 3, a conceptual framework for identity-oriented persuasive design in mHealth can be derived. Figure 1 illustrates this framework, linking identity-oriented design features in mHealth applications to sustained health behavior through processes of narrative identity development. It synthesizes these abstractions into a three-level conceptual model linking design features, identity processes, and downstream behavioral outcomes.

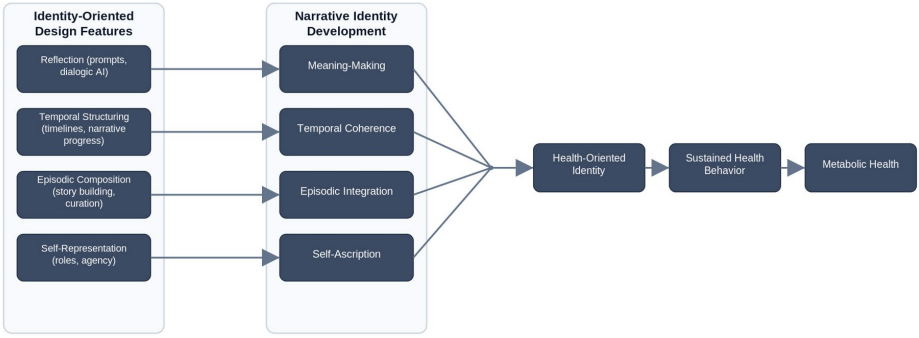

Figure 1: A Framework for Identity-Oriented mHealth Design
Source: Own Elaboration

A person using an identity-oriented mHealth app is supported in making sense of setbacks, linking present behavior to past and future, and viewing health actions as part of an evolving personal story. In this way, the system scaffolds a health-oriented identity rather than merely tracking behavior. The framework assumes that long-term behavior change does not primarily emerge from external reinforcement mechanisms, but from the integration of health-related behaviors into a person's self-understanding. Accordingly, the model differentiates between three analytical levels: (1) identity-oriented design features, (2) narrative identity development processes, and (3) behavioral and health-related outcomes.

At the design level, four groups of features are distinguished. Reflection features (e.g., context-sensitive prompts or dialogical LLM-systems) stimulate active meaning-making. Temporal structuring features (e.g., timelines or narrative progress

framing) support the integration of experiences across time. Episodic composition features (e.g., curated story-building from lifelog data) enable the transformation of fragmented behavioral traces into coherent episodes. Finally, self-representation features (e.g., role framing or identity-congruent self-positioning) reinforce agency and the experience of oneself as an acting subject within an unfolding narrative.

These feature groups scaffold four corresponding processes of narrative identity construction: meaning-making, temporal coherence, episodic integration, and self-ascription. Together, these processes contribute to the emergence of a health-oriented identity, that is, the perception of oneself as a person for whom health-related behavior is self-defining and value congruent.

The model further proposes that once health-related behavior becomes identity-congruent, it is more likely to be regulated through integrated forms of motivation rather than external pressure. This identity-based internalization increases the likelihood of sustained lifestyle behaviors, such as regular physical activity, balanced nutrition, and stress regulation, which in turn support long-term metabolic health. Importantly, the framework conceptualizes mHealth systems not as instruments that prescribe identities, but as environments that scaffold identity work. The effectiveness of identity-oriented design therefore depends on preserving user autonomy and interpretive flexibility, allowing individuals to construct personally meaningful health narratives rather than conforming to predefined optimization scripts.

5 Discussion: Identity Integration in mHealth Design

Many mHealth interventions continue to show limited long-term effectiveness despite increasingly sophisticated tracking and personalization capabilities (Puja et al., 2025). This paper argues that an important explanatory gap lies at the level of identity integration. Health behaviors are more likely to persist when they become integrated into an individual's evolving narrative understanding of the self rather than remaining externally prompted actions. The proposed framework therefore reconceptualizes persuasive mHealth as narrative scaffolding and positions narrative identity as mediating layer linking design features to sustain behavior change. It organizes identity supportive design knowledge around four narrative processes: interpretation, temporal integration, episodic configuration, and self-representation.

From this perspective, interaction design features such as reflective prompts, narrative progress representations, lifelog curation, and identity congruent forms of self-representation can help individuals interpret behavioral data as meaningful episodes within a broader life trajectory.

This perspective is particularly relevant in the context of emerging LLM-based conversational systems. Such systems enable longitudinal interaction and reflective dialogue around behavioral change. Research on narrative interfaces indicates that technological framing shapes how users position themselves developmentally (Murnane et al., 2020). Generative systems may therefore not only scaffold meaning making but also privilege narratives of self-optimization, productivity, or moralized health behavior (Dokic et al., 2025). The framework thus invites critical reflection on identity-oriented design, including which identity trajectories are amplified by technological framing and how interpretive openness and user authorship can be preserved. At the same time, the importance of narrative identity integration likely varies across behavioral contexts. Identity level processes may be particularly relevant for complex and sustained lifestyle changes such as long-term metabolic health management, where individuals must repeatedly reinterpret progress, setbacks, and turning points (Hajdarevic et al., 2025). In contrast, more routine or time limited interventions may require less identity level integration (Verplanken & Sui, 2019).

This work has several limitations. The framework is based on a structured synthesis of literature across HCI, mHealth, lifelogging, and digital reminiscence rather than a fully systematic review. While the taxonomy organizes existing design knowledge conceptually, further empirical research is needed to evaluate how identity-oriented design features influence long-term behavior change in real world interventions. Future work should therefore investigate how narrative scaffolding mechanisms operate in longitudinal mHealth deployments and how LLM mediated reflection shapes identity processes over time. Mixed method approaches combining narrative coding, validated psychosocial measures, and behavioral and clinical tracking such as glycated hemoglobin (HbA1c), a long-term blood glucose indicator and weight trajectories could assess identity development as a process level mechanism rather than merely an outcome. Ultimately, the long-term effectiveness of mHealth may depend less on optimizing behavioral prompts than on supporting the narrative integration of health behaviors into a person's evolving sense of self.

References

Adler, J. M., Lodi-Smith, J., Philippe, F. L., & Houle, I. (2016). The incremental validity of narrative identity in predicting well-being: A review of the field and recommendations for the future. *Personality and Social Psychology Review, 20(2)*, 142–175. https://doi.org/10.1177/1088868315585068

Byrne, D., & Jones, G. J. (2009). Creating stories for reflection from multimodal lifelog content: An initial investigation. In *Designing for reflection on experience workshop at CHI 2009*. https://doras.dcu.ie/16136/

Chandra, M., Naik, S., Ford, D., Okoli, E., De Choudhury, M., Ershadi, M., Ramos, G., Hernandez, J., Bhattacharjee, A., & Warreth, S. (2025). From lived experience to insight: Unpacking the psychological risks of using AI conversational agents. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25)*. https://doi.org/10.1145/3715275.3732063

Chen, S., Cao, Z., Chen, W., Zhao, J., Jiao, L., Prettner, K., Kuhn, M., Pan, A., Bärnighausen, T. W., & Bloom, D. E. (2025). The global macroeconomic burden of diabetes mellitus. *Nature Medicine, 32(1)*, 126–138. https://doi.org/10.1038/s41591-025-04027-5

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI R) and NEO Five Factor Inventory (NEO FFI) professional manual*. Psychological Assessment Resources.

Da Couto, F. F. S., & Almeida, C. P. B. de. (2025). Mobile and web apps for weight management in overweight and obese adults: An updated umbrella review and meta-analysis. *International Journal of Environmental Research and Public Health, 22(7)*, Article 1152. https://doi.org/10.3390/ijerph22071152

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.

Dokic, K., Pisker, B., & Radisic, B. (2025). Mirroring cultural dominance: Disclosing large language models' social values, attitudes and stereotypes. *Societies, 15(5)*, 142. https://doi.org/10.3390/soc15050142

De Hoog, N., & Pat-El, R. (2024). Social identity and health-related behavior: A systematic review and meta-analysis. *Social Science & Medicine, 344*, 116629. https://doi.org/10.1016/j.socscimed.2024.116629

Di Zhu, Al Mahmud, A., & Liu, W. (2024). Design requirements for a digital storytelling application for people with mild cognitive impairment (MCI). *Digital Health, 10*, 20552076241282237. https://doi.org/10.1177/20552076241282237

Erikson, E. H. (1968). *Identity: Youth and crisis*. W. W. Norton.

Fogg, B. J. (2009). A behavior model for persuasive design. In S. Chatterjee & P. Dev (Eds.), *Proceedings of the 4th international conference on persuasive technology (pp. 1–7)*. Association for Computing Machinery. https://doi.org/10.1145/1541948.1541999

Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday.

Habermas, T., & Bluck, S. (2000). Getting a life: The emergence of the life story in adolescence. *Psychological Bulletin, 126(5)*, 748–769. https://doi.org/10.1037/0033-2909.126.5.748

Haimson, O. L., & Marathe, M. (2023). Uncovering personal histories: A technology-mediated approach to eliciting reflection on identity transitions. *ACM Transactions on Computer-Human Interaction, 30(2)*. https://doi.org/10.1145/3504004

Hajdarevic, S., Norberg, A., Lundman, B., & Hörnsten, Å. (2025). Becoming whole again—Caring for the self in chronic illness—A narrative review of qualitative empirical studies. *Journal of Clinical Nursing, 34(3)*, 754–771. https://doi.org/10.1111/jocn.17332

Kvietkute, D., & Ness, I. J. (2026). Encountering generative AI: Narrative self-formation and technologies of the self among young adults. *Societies, 16(1)*, 26. https://doi.org/10.3390/soc16010026

Leuzzi, G., Job, M., Cosentino, C., Testa, R., Scafoglieri, A., & Testa, M. (2026). Mobile apps to improve health parameters in healthy adults: Systematic review. *JMIR mHealth and uHealth, 14*, e66881. https://doi.org/10.2196/66881

Marcia, J. E. (1966). Development and validation of ego-identity status. *Journal of Personality and Social Psychology, 3(5)*, 551–558. https://doi.org/10.1037/h0023281

McAdams, D. P. (1996). Personality, modernity, and the storied self: A contemporary framework for studying persons. *Psychological Inquiry, 7(4)*, 295–321. https://doi.org/10.1207/s15327965pli0704_1

McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology, 5(2)*, 100–122. https://doi.org/10.1037/1089-2680.5.2.100

McAdams, D. P., & McLean, K. C. (2013). Narrative identity. *Current Directions in Psychological Science, 22(3)*, 233–238. https://doi.org/10.1177/0963721413475622

Mead, G. H. (1934). *Mind, self, and society from the standpoint of a social behaviorist* (C. W. Morris, Ed.). University of Chicago Press.

Meyer, S., & Elsweiler, D. (2025). LLM-based conversational agents for behaviour change support: A randomised controlled trial examining efficacy, safety, and the role of user behaviour. *International Journal of Human-Computer Studies, 200*, 103514. https://doi.org/10.1016/j.ijhcs.2025.103514

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science, 6*, 42. https://doi.org/10.1186/1748-5908-6-42

Murnane, E. L., Jiang, X., Kong, A., Park, M., Shi, W., Soohoo, C., Vink, L., Xia, I., Xin, Y., Yang-Sammataro, J., Young, G., Zhi, J., Moya, P., & Landay, J. A. (2020). Designing ambient narrative-based interfaces to reflect and motivate physical activity. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. Association for Computing Machinery. https://doi.org/10.1145/3313831.3376478

Nepal, S., Pillai, A., Campbell, W., Massachi, T., Choi, E. S., Xu, O., Kuc, J., Huckins, J., Holden, J., Depp, C., Jacobson, N., Czerwinski, M., Granholm, E., & Campbell, A. T. (2024). Contextual AI journaling: Integrating LLM and time series behavioral sensing technology to promote self-reflection and well-being using the MindScape app. In *Extended abstracts of the 2024 CHI conference on human factors in computing systems*. Association for Computing Machinery. https://doi.org/10.1145/3613905.3650767

Nicholson, S. (2014). A RECIPE for meaningful gamification. In T. Reiners & L. C. Wood (Eds.), *Gamification in education and business* (pp. 1–20). Springer. https://doi.org/10.1007/978-3-319-10208-5_1

Obsan. (2025). Health care costs (MonAM indicator). *Swiss Health Observatory*. https://ind.obsan.admin.ch/en/indicator/monam/health-care-costs

Priebe, C. S., Beauchamp, M. R., Wunderlich, K. B., & Faulkner, G. (2020). I'm a runner not a smoker: Changes in identity as predictors of smoking cessation and physical activity. *Psychology of Sport and Exercise, 49*, 101702. https://doi.org/10.1016/j.psychsport.2020.101702

Pujia, C., Ferro, Y., Mazza, E., Maurotti, S., Montalcini, T., & Pujia, A. (2025). The role of mobile apps in obesity management: Systematic review and meta-analysis. *Journal of Medical Internet Research, 27*, e66887. https://doi.org/10.2196/66887

Rahsepar Meadi, M., Sillekens, T., Metselaar, S., van Balkom, A., Bernstein, J., & Batelaan, N. (2025). Exploring the ethical challenges of conversational AI in mental health care: Scoping review. *JMIR Mental Health, 12*, e60432. https://doi.org/10.2196/60432

Rhodes, R. E., Rebar, A. L., & Strachan, S. (2025). Physical activity identity as an axis of dual process motivation and self-regulation processes: Current evidence and future research directions. *Psychology of Sport & Exercise, 80*, 102923. https://doi.org/10.1016/j.psychsport.2025.102923

Ricoeur, P. (1984). *Time and narrative* (Vol. 1, K. McLaughlin & D. Pellauer, Trans.). University of Chicago Press.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55(1)*, 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Tran, L. D., Kennedy, D., Zhou, L., Nguyen, B., & Gurrin, C. (2022). A virtual reality reminiscence interface for personal lifelogs. In *MultiMedia Modeling* (Lecture Notes in Computer Science, Vol. 13142, pp. 479–484). Springer. https://doi.org/10.1007/978-3-030-98355-0_42

Verplanken, B., & Sui, J. (2019). Habit and identity: Behavioral, cognitive, affective, and motivational facets of an integrated self. *Frontiers in Psychology, 10*, 1504.

Wang, X., Li, J., Liang, T., Hasan, W. U., Zaman, K. T., Du, Y., Xie, B., & Tao, C. (2024). Promoting personalized reminiscence among cognitively intact older adults through an AI-driven interactive multimodal photo album: Development and usability study. *JMIR Aging, 7*, e49415. https://doi.org/10.2196/49415

Wierts, C. M., Kroc, E., & Rhodes, R. E. (2024). The role of intention, behavioral regulation, and physical activity behavior in the prediction of physical activity identity across time. *Behavioral Sciences, 14(10)*, 886. https://doi.org/10.3390/bs14100886

Xu, L., Jia, S., Gurrin, C., Ward, M., & Tran, A. (2025). Through someone else's eyes: Lifelogging meets narrative virtual reality. In *MM '25: Proceedings of the 33rd ACM international conference on multimedia* (pp. 8988–8996). Association for Computing Machinery. https://doi.org/10.1145/3746027.3755833

Yee, N., & Bailenson, J. N. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research, 33(3)*, 271–290. https://doi.org/10.1111/j.1468-2958.2007.00299.x

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 493–502). Association for Computing Machinery. https://doi.org/10.1145/1240624.1240704

About the author

Esther Federspiel, psychologist, studies meaningful personalization and identity-oriented persuasive mHealth.

Summary

Metabolic chronic conditions require sustained lifestyle change, yet mHealth interventions often produce only modest and short-lived effects, with high dropout and relapse rates. This paper argues that a central explanatory gap in persuasive mHealth design lies at the level of identity integration. Health behaviors are more likely to persist when they are not merely performed but become narratively integrated into an individual's evolving self-understanding.

To address this gap, the paper reconceptualizes persuasive mHealth as narrative scaffolding instead of behavioral control. Drawing on McAdams' narrative identity theory and extending the concept of meaningful gamification, it introduces an identity-oriented framework that conceptualizes narrative identity as the mediating mechanism between design features and sustained behavior change.

The framework organizes identity supportive design knowledge along four core narrative processes: interpretation, understood as meaning making; temporal integration, referring to biographical continuity; episodic configuration, which transforms behavioral traces into coherent episodes; and self-representation, emphasizing agency and protagonist positioning. Based on a structured synthesis of empirical HCI and mHealth research, a taxonomy of design patterns such as reflective AI prompts, narrative progress framing, lifelog curation, reminiscence interfaces, and identity congruent self-representation is mapped onto these processes.

The framework is particularly relevant in the context of LLM-based conversational systems. While such systems can scaffold reflection and longitudinal meaning making at scale, they also risk reproducing culturally dominant self-optimization scripts and over structuring identity development. The paper therefore contributes both generative design guidance and critical evaluation criteria and advocates for mHealth systems that preserve interpretive openness and user authorship while supporting the gradual integration of health behaviors into narrative identity over time.

Review of Disruptive Digital Solutions in the Function of Maritime Transport and Safety Matters

Zdravko Paladin,[1] Mira Šorović,[2]
Nexhat Kapidani,[1] Žarko Lukšić,[1] Rasim Mujalović[3]

[1] Maritime Safety Administration (MSA), Bar, Montenegro
zdravko.paladin@pomorstvo.me, nexhat.kapidani@pomorstvo.me,
zarko.luksic@pomorstvo.me
[2] University of Donja Gorica, Humanistic Studies, Podgorica, Montenegro
mira.sorovic@udg.edu.me
[3] Jadroagent Bar, Bar, Montenegro
rasi@t-com.me

The development and improvement of a vast range of disruptive information technologies (IT), including revolutionary Artificial Intelligence (AI) applications, requires the continuous collection of data from ship navigation and transport, transforming it into meaningful information for end-users. Contemporary AI-driven applications in the maritime industry are significantly advanced for processing data directly at the source, ensuring resilience, minimizing latency and errors, and reducing overall costs. This paper aims to present and investigate the features of the latest disruptive IT and AI use case technologies with their real impact on maritime trade, transport, surveillance, and safety based on data sharing and exchange platforms. The innovative solutions review wrapped most recent AI tools and features, concluding with the analysis of the practical improvements in the maritime industry and operating systems. By embracing AI, the maritime industry makes remarkable success with regards to competitive edge, better user and customers' experiences, with an aim to achieve a higher operational efficiency.

DOI
https://doi.org/
10.18690/um.fov.4.2026.15

ISBN
978-961-299-152-4

Keywords:
maritime transport,
AI,
big data,
ADRIREP,
data sharing

University of Maribor Press

1 Introduction

Maritime transport undergoes revolutionary technological changes that involve a shift in various competencies, redefining how ships, ports, and logistics systems operate in the digital era (Kumar, 2025). Transportation, movement, and surveillance are transforming from purely technological considerations into digital services, thanks to Big Data, AI, and Internet of Things (IoT) sensor technologies serving to increase efficiency, accuracy, and reduce costs and errors (Schulz *et al.*, 2021). The digital transformation and optimization of core processes in shipping and logistics require the adoption of a wide range of modern technologies such as IoT, automation, cloud computing, and AI-supported intelligent decision-making (Raza *et al.*, 2023). Other disruptive technologies in the maritime navigation sector are digital twins, autonomous maritime vehicles, and metaverse integration, whose main purpose is to support maritime operational centres, in compliance with standards for data management and cybersecurity (Abu Hantesh, 2024; Tejwani, 2024).

The paper unfolds as follows: After methodology Chapter 2, the recent development in a ship mandatory reporting system, mostly used for dangerous cargoes in the Adriatic region (ADRIREP – ADRIatic traffic monitoring REPorting), and enhanced technological advancements in maritime and logistic physical internet, are summarized in Chapter 3, reviewing two outstanding projects, the EUREKA and the ePIcenter. Chapter 4 provides a study of advanced digital platforms for managing the cargo flows in maritime transport and logistics, and the most important recently developed AI solutions in the maritime industry, focusing on the maritime safety, surveillance, and security issues, while Chapter 5 gives discussion and analysis of the application of AI-supported modules in maritime safety and transport efficiency.

2 Methodology

The methodological approach of this paper is a qualitative structured research design, providing a comprehensive overview of contemporary maritime transport, safety, and logistics IT solutions involved in recent digital transformation within the maritime industry. This research has reviewed the most recent developments in the maritime sector, focusing on the analysis of technical features, including presentation of two innovative projects in maritime and transport digitalization. Several cases of

implementation of AI in maritime transport systems were analyzed by criteria of data collection, sharing, and cost-effectiveness in operational and safety functions.

3 Presentation of recent Maritime Digitalisation Projects technologies

3.1 Achievements of the Project EUREKA regarding ADRIREP system

The EUREKA project represents a coordinated regional initiative (between Italy, Slovenia, Croatia, Bosnia and Herzegovina, Montenegro, Albania, and Greece) focused on improving maritime safety, surveillance efficiency, and institutional cooperation among Adriatic-Ionian countries. Conducted within the INTERREG ADRION framework, the project addressed long-standing challenges in maritime monitoring systems, data exchange, and regulatory alignment by introducing integrated technical and organizational solutions. Its central objective is to increase navigational safety through systematic coordination among national maritime administrations, harmonized legal frameworks, standardized Vessel Traffic Service (VTS) structures, improved data exchange, and the development of unified training standards for operators (Šorović *et al.*, 2023).

A key achievement of the project is the modernization of the ADRIREP ship reporting system, originally adopted in 2003 under International Maritime Organization (IMO) regulations (Figure 1), as an obligatory reporting mechanism for vessels carrying dangerous cargo onboard or meeting specific tonnage thresholds. The EUREKA introduced an updated technical design to automate reporting processes, reduce administrative workload and enhance surveillance accuracy through digitalization and integration. These improvements allow ship data to be submitted, processed, and shared electronically, significantly reducing redundant reporting and manual communication (Šorović *et al.*, 2023).

Also, this project introduced a centralized reporting architecture based on an Integrated Report Distribution (IRD) platform, which serves as a shared repository for maritime data. Through this infrastructure, ship reports can be exchanged automatically or manually among competent coastal authorities, enabling faster information flow with improved situational awareness. The system supports interoperability with existing national and EU maritime databases and ensures that previously submitted information can be reused rather than repeatedly transmitted.

Such integration is particularly important for Vessel Traffic Monitoring and Information Systems (VTMIS), whose effectiveness depends on accurate, timely, and shared navigation data among multiple stakeholders (Komadina, 2013; Kuma, 2015). Thus, beyond technical modernization, the EUREKA project produced significant institutional outcomes. One of its main contributions is the establishment of a Maritime Safety Permanent Transnational Network, intended to maintain long-term cooperation among participating countries and ensure continuity of safety improvements after project completion. In addition, this project supports the harmonization of VTS organizational structures and develops standardized training programs and competence requirements for operators, promoting consistent professional qualifications across the region. These measures strengthen interoperability, facilitate joint operations, and support coordinated responses to maritime incidents.

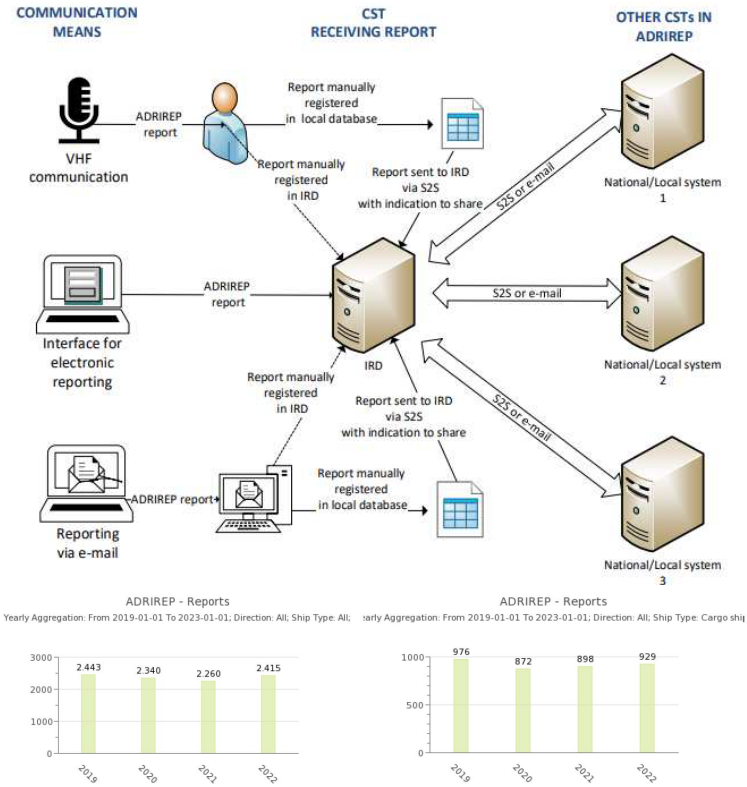

Figure 1: The ADRIREP system context diagram with output ship statistics in Port of Bar
Source: EMSA, 2022; Šorović *et al.*, 2023

Another important innovation is the planned testing and evaluation of Sea Traffic Management (STM) integration within national monitoring systems. This concept aims to optimize logistics, communication, and traffic management processes in ports and surveillance areas, contributing to more efficient vessel movement and safe navigation conditions (Šorović *et al.*, 2023). By combining modern technologies with coordinated governance mechanisms, the project demonstrates how regional cooperation can address complex maritime challenges more effectively than isolated national approaches. Overall, the results of the EUREKA project show that integrated digital platforms, harmonized procedures, and sustained institutional collaboration can significantly enhance maritime safety and operational efficiency. The modernization of ADRIREP, reduction of administration burdens, elimination of redundant data collection, and improvement of real-time monitoring collectively form a comprehensive framework for safer navigation in the Adriatic-Ionian region. As such, the project serves as a model for future multinational initiatives seeking to modernize maritime governance through technological innovation.

3.2 Achievements of the Project ePIcenter towards logistic optimization

Modern global logistics chains operate under conditions of high interdependence, dynamic flows, and increased uncertainty. In such an environment, disruptions in the supply chain (delays, transport disruptions, port congestion, unplanned schedule changes, and capacity constraints) have a ripple effect across multiple system links, which may lead to broader operational and economic consequences (Christopher & Peck, 2004). Additionally, the growing complexity of trade flows, reflected in multimodality, the large number of stakeholders, and the geographical dispersion of processes, further complicates the planning and coordination of transport activities. A particular challenge is the presence of information asymmetries among stakeholders, as different participants in the supply chain often have access to inconsistent, incomplete, or delayed information, which reduces decision-making quality and increases the risk of misalignment in logistics operations.

In this context, the ePIcenter project developed and tested IT solutions that enhance daily logistics activities, including planning, coordination, and monitoring of cargo flows along the supply chain. One of the project's key contributions is the improvement of connectivity and coordinated data exchange among different actors in the logistics system. Instead of a fragmented approach, where each participant

uses isolated information systems, ePIcenter promotes collaboration through a shared ecosystem of tools and methods based on cloud infrastructure. This approach reduces information fragmentation and ensures more consistent data exchange, which is particularly important in multimodal logistics, where even minor delays or untimely updates can significantly affect planning decisions (ePIcenter GA, 2020).

In maritime logistics, the ePIcenter project's case study highlights the importance of advanced data-sharing platforms and harmonized digital processes as a foundation for more efficient coordination among stakeholders. Examples of structured information exchange in port and shipping environments include initiatives such as Maritime Single Window (MSW), Port Community Systems (PCS), smart containers, Transporeon multimodal visibility solution and others systems based on mandatory IMO and UN regulations on data sharing (Paladin *et al.*, 2024). Availability of information on schedule changes, shipment status, and port events enables early detection of deviations and timely adjustment of plans before delays accumulate and cause broader operational disruptions. This is particularly important in conditions of frequent disruptions in global supply chains, as improved visibility and data exchange enhance resilience and the ability to respond quickly. The application of the same principle of digital connectivity also has a significant impact on road and rail transport, as planning in land-based segments often directly depends on events synchronization in the maritime stage. In this regard, the approach developed within the ePIcenter project enables faster coordination and rescheduling of dispatch and land transport operations when changes occur. By relying on harmonized data exchange and shared decision-support workflows, updated information from ports and maritime arrivals can be integrated earlier into land transport planning. This reduces timing mismatches between maritime and land segments, mitigates the effects of information asymmetries among stakeholders, and lowers the risk of congestion at transfer points between sea and land transport (ePIcenter GA, 2020).

Viewed in a broader context, the contribution of the ePIcenter project is not limited to technical improvements of digital tools but also includes strengthening the coordination capacity of logistics chains under conditions of global complexity and disruption. Therefore, system connectivity, standardized data exchange, and shared decision-support mechanisms represent key prerequisites for greater resilience, flexibility, and efficiency of modern transport and logistics systems.

4 Synthesis of an advanced digital technologies framework to ensure Maritime Transport, Surveillance, and Cargo Safety

4.1 Advanced digital solutions in maritime transport and logistics

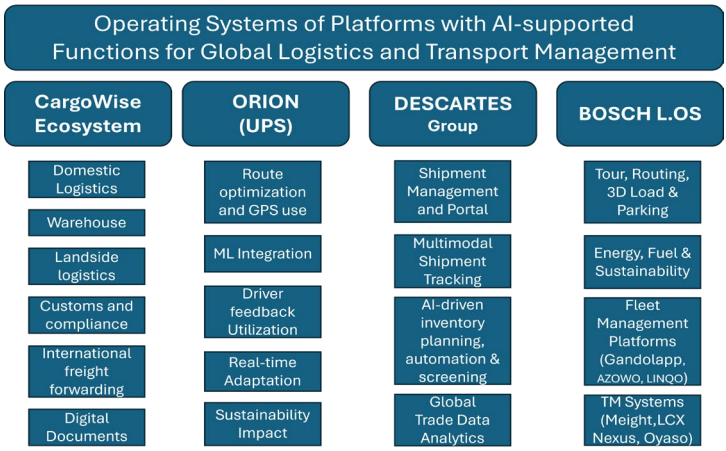

Figure 2: General overview of main operating ecosystem functions
Source: Own elaboration based on websites references

As an outstanding contemporary platform widely used in logistics and supply chain management, we elaborate on CargoWise software that brings together all key freight forwarding functions into a single platform used for global operations. Among the advanced capabilities of this tool, the end-to-end integration of functions such as bookings management, shipments, consolidations, customs clearance, and billing should be pointed out. What is especially important is the integrated AI-powered assistant CargoWise Ace that provides answers to any question about the platform, with clear, step-by-step guidance, making the use of CargoWise easier. As a software for intelligent logistics execution, Ace represents a step towards deeper AI integration in CargoWise, enabling the automation of routine tasks, uncovering real-time insights, flagging exceptions, and leading to a high level of productivity. Following the mentioned advanced solutions in logistics, the global delivery company United Parcel Service, Inc. (UPS), leverages AI to optimize its package delivery routes and operations and to reduce fuel consumption, by using an AI-driven optimization engine called ORION (On-Road Integrated Optimization and Navigation), which analyses factors such as delivery volume, requirements, and

traffic patterns to generate efficient routes for drivers (Soumpenioti, 2023). Other relevant platforms in international logistics are DESCARTES and BOSCH L.OS. The general overview of their operating ecosystem functions is shown in Figure 2.

4.2 Maritime Safety Monitoring Technologies and Reporting

Of paramount importance to comprehensive maritime surveillance and continual management of vessel traffic is establishing a reliable monitoring system for accurate navigation and safety evaluation that will enable reaching a high level of situational awareness and preparedness at sea. This comprehends a determined position, speed, course, weather conditions, and other relevant technical factors that make the navigation safe and efficient, using the equipment and complex devices based on reliability, integrity, interference, and failure resilience. These devices leverage the advanced features of integrated sensors and monitoring systems for maritime navigation and situational awareness, including interference suppression, multisensor-processing, monitoring of integrity, back-up storage, data plausibility, and Automatic Identification System (AIS)/Radar fusion (Noack *et al.*, 2017). These technologies are widely based on operational Global Navigation Satellite Systems (GNSS), satellite navigation devices (GPS), functions known as PNT (Positioning, Navigation and Timing), and specifically used by the EU, the GALILEO, together with the EGNOS satellite-based augmentation system (SBAS) to enhance the accuracy of the global GNSS. The system's high-level architecture for multisensor ship data processing and outputs is illustrated in Figure 3.

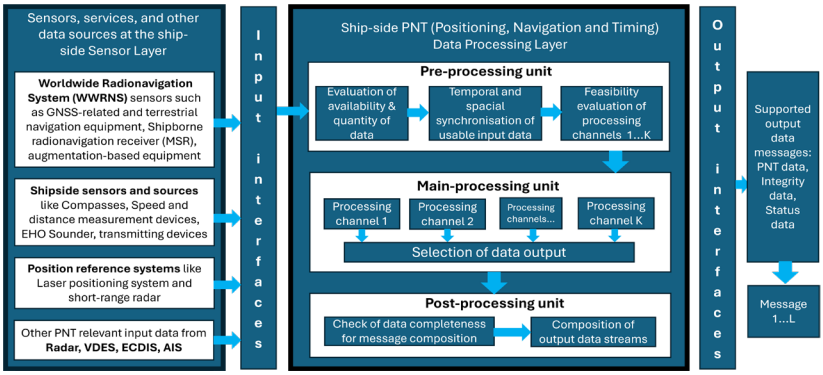

Figure 3: The high-level architecture of GNSS and PNT data processing framework
Source: Our elaboration from Noack *et al.* (2017)

Given that maritime safety relies on technological systems for vessel surveillance and rapid exchange of information through national maritime authority systems, EU countries established the concept of a Common Information Sharing Environment (CISE). The idea for this stems from EUCISE2020, a project standard network for sharing and exchanging relevant maritime data and information between cooperating agencies. Based on the European Maritime Safety Agency (EMSA) Guidelines for CISE, the entire information sharing/exchange/interpretation process is managed through the CISE Data & Services Model, which is structured in five main building blocks: the legacy system (LS) of the agency participating in the maritime data exchange, the EU/regional/national CISE node, the CISE adapter, the CISE node/gateway, and the CISE network (Paladin *et al.*, 2022; 2023).

Another important recently developed platform for sharing of data in the maritime sector is the Data Space large database, utilized as a new approach to decentralized and sovereign data sharing, consisting of several architectural components: The Connectors, the Identity Provider, the Data Broker, and Governance. Such a design of the Maritime Connectivity platform integrated within the VTS and Traffic Management System (TMS) offers advanced features in the sovereign and secure data provision, connection of heterogeneous data sources, integrity and certification, finding, and availability of maritime data (Höhn *et al.*, 2025). Concerning the VTS, AI enhances its functions by predicting ship arrival times, optimizing berthing schedules, and managing congestion, by using the Machine Learning (ML) models that analyze real-time AIS data and weather reports, to ensure safe and efficient vessel movement. The outstanding example for this in the EU is the Port of Rotterdam, which already employs AI-driven analytics to predict ETA and EDA (Estimated Time and Date of Arrival) with high accuracy, reducing idle times and port congestion (Kumar, 2025).

4.3 Advanced AI modules in maritime surveillance and logistics

In recent years, AI advancements at all segments of the maritime sector, especially in driving predictive maintenance, enhancing safety and security, and streamlining compliance with the needs of port and shipping operators, as well as both national and international regulations (Abu Hantesh *et al.*, 2024). AI is also widely applied in various aspects of logistics, such as shipping demand forecasting, route optimization, predictive models for congestion management in ports, ship collision risk

assessment, automation of customs clearance processes, and maritime surveillance. One of the most important features of AI in maritime navigation is that, based on historical data and real-time conditions analysis from Big Data repositories, AI algorithms find the optimal route that minimizes cost function, respecting the constraints like fuel prices, traffic conditions, and weather patterns, which significantly reduces transportation costs and improves delivery times and reliability (Amadasu, 2024).

Table 1: Maritime and logistics AI projects/products examples by tasks performed

Sector	Sensing		Thinking		Acting	
	Company	AI project / product	Company	AI project / product	Company	AI project / product
Intra-logistics	Fraunhofer (DE)	Identification based on natural features	VTT (FI)	KnowlEdge	Fraunhofer (DE)	Real autonomous robot navigation
			Baobab (ES)	Production Planning		
			TU/e (NL)	AI Planner		
Ports	Port de Barcelona (ES)	Goclever-Sensor, Cloud, and API Integration	UD (ES)	LOGISTAR	TUDelft & TNO (NL)	SWARM-PORT (Agent-based model simulation for decision support tool for the port nautical services, developed in PySeidon)
			Port of Algeciras (ES)	PortXChange And IDOM & AWS		
	Captain AI (NL)	AI based radar tracker for inland shipping	Fundación Valenciaport (ES)	E2RM		
			Circle Group (IT)	Milos Intelligence		
	ICCS (EL)	5G-LOGINNOV	IMEC (BE)	COOCK Smart Port 2025		
Transport Management	DB Schenker Enterprise Lab (DE/AT) and Fraunhofer IML (DE)	Video Analytics in Landtransport (automated yard access control, real-time allocation of (un)loading docks using MLCVZoo and OCR algorithms)	CIMNE (ES)	TOOL: Traffic fOrcasting fOr Logistics	ITAINNOVA (ES)	SmartDriving (Telemetry processing to update speeds based on fuel consumption)
			FH Steyr (AT)	Demand prediction in retail		
			TU/e & UT (NL)	DYNAPLEX and Digital Twins for City Logistics	Utwente (NL)	ICCOS (using AI to reduce empty mileage, by Python and Google Colab)
			Rulex (IT)	Data driven dashboard		

Source: Adapted according to the Murrenhoff *et al.*, (2023) ALICE, with companies' seats countries codes

Utilizing the Big Data repositories, the analysis of a vast amount of maritime and ship data, reflects in main groups: supervised, semi-, and unsupervised algorithms (Vu *et al.*, 2024). Supervised analysis is based on linear and logistic regression, decision trees, and support vector machines, used for predictive purposes and requiring historical data, while unsupervised analysis is based on clustering, used for data structuring. Following these facts, there has already been an increasing trend for the implementation of ML methods as part of AI applications in the maritime industry, whose examples follow (Gomez-Ruiz *et al.*, 2023):

1. Prediction of ship motion, possible paths, and providing optimal trajectory.
2. Evaluation of navigation, oil spill event risk, and ship fuel consumption rates.
3. Development, application, and eco-design optimization of intelligent vessels.
4. Estimation of the berthing state of autonomous maritime surface ships.

An overview of the most relevant AI applications is given in Table 1.

An important field of transformative and developing AI platform applications is reflected in autonomous shipping and navigation tasks in the form of Maritime Autonomous Surface Ships (MASS). In this sense, the set of AI-driven systems for MASS integrate data from sensors, radars, Light Detection and Ranging (LIDAR), RGB and infrared cameras, Motion Reference and Inertial Measurement Units (IMU/MRU), AIS, weather and depth data sensors and satellite communication networks (GNSS) to support decision-making, situational awareness, ship and shore recognition, and predictive navigation (Kumar, 2025; Koch, 2022). A classic example of system integration of all the mentioned features is the Mayflower Autonomous Ship, produced in 2022. Therefore, a high level of intelligent assistance systems provides the ship autonomy and the new methods of ship handling to use existing infrastructure more efficiently. This will make expensive infrastructure projects unnecessary, such as dredging. It is important to mention that the highly automated ships already exist, e.g., "Yara Birkeland" of the Norwegian shipping company Kongsberg, and the highly automated ship "C-worker-7", managed by a British shipping company (Hesse *et al.*, 2019). A recent market analysis of the most disruptive AI tools utilized in maritime navigation, done by Stach *et al.* (2024), covered 26 products, among which the highly evaluated were Avikus NeuBoat and

HiNAS 2.0, Groke Pro, Kongsberg Maritime Intelligent Awareness, MTU NautIQ Copilot, Sea Machines SM300, Seasight and Wärtsilä Voyage Autonomy Solutions, which correspond to safety and quality requirements from the IMO, and ISO/IEC.

5 Discussion on the application of AI-supported modules in maritime safety and transport efficiency findings

Two essential conditions to enable the maritime transport flows proper are safety and risk mitigation. Safety in maritime transportation refers to the measures and practices implemented to prevent accidents and incidents that could endanger human life, damage property, or harm the marine environment, while the risk management in maritime transportation involves the systematic identification, assessment, and mitigation of risks associated with maritime operations. To manage these challenges, some global maritime multinational companies have introduced advanced AI modules such as Kongsberg's KognifAI Maritime Ecosystem for Predictive Maintenance, Maersk's AI-Powered Risk Management, DNV GL Veracity Platform for risk analysis, Wärtsilä's Fleet Operations Solution and ABB Ability™ Marine Pilot Vision for the safety of navigation systems. With the mentioned AI tools, Maersk increased the ability to ensure the safety of its vessels and crew while maintaining efficient and safe operations. On the other side, the Veracity platform enables continuous monitoring of critical systems, providing insights into their condition and performance. The Wärtsilä system's predictive capabilities enhance situational awareness and reduce the risk of accidents, contributing to safer and more efficient maritime operations (Durlik *et al.*, 2024).

As stated in Amadasu (2024), the combined use of AI-supported modules and blockchain in maritime applications contributed to higher safety and reduction the operational costs in port by 15%. Also, it is important to mention that sectors with the biggest opportunities for AI, based on the report of McKinsey (Chui *et al.*, 2018), are transport and logistics industries, having an annual value of US$ 400-500 billion, which is a great potential to capitalise on this business value. Therefore, AI potential in the freight logistics sector is based on various use cases, depending on the research project for the particular use or process, as follows (du Plessis *et al.*, 2025):

- Freight opportunity platform (matching the transport supply and demand);
- Production, regulation, or control of goods to prevent under-production;

- Resilience of logistical system (risk mitigation, adjustment of safety levels);
- Sustainability initiatives, energy or fuel consumption, and waste reduction;
- Ship and cargo safety (security hotspot identification, accident detection);
- Transport management system (booking/allocation system optimisation);
- Efficiency improvement (loading/unloading and routing of vehicles/ships).

Due to fierce international shipping competition, the industry is under cost pressure and must innovate, but highly automated systems promise relief. According to the study of Hesse *et al.* (2019), a highly automated freighter can save US$4.3 million over 25 years compared to a traditionally manned vessel.

An interesting approach for an integrated use of AI tools and digital twins (a digital projection of the ship) is provided by the Marine Digital GmbH company in the field of ship performance optimization. This combination offers modular digitization and the improvement of modern ships' equipment, navigational devices, and tools. The digital twin collects all data and streams it to a compliant and secure cloud, where fleet managers get live notifications such as: weather events, maintenance messages, ETA changes, and routes generated by AI algorithms, focusing on optimal fuel consumption and ETA restrictions. Fuel savings and the associated emissions optimization effect range from 5% to 12%, depending on the specific operating characteristics of the individual vessel (Schulz *et al.*, 2021).

This paper has confirmed a profound impact of AI technology on the maritime transportation sector, shaping its development, efficiency, and scalability to meet the high standards and requirements. Also, AI-driven software brings many positive aspects to the overall safety and security of the ship's digital and transport system in general, covering a wide range of functions and areas of vital significance to stakeholders, leading to greater productivity of companies and resilience of critical assets. However, careful attention must be given to data privacy and cybersecurity.

References

AbuHantash, A., Goussous, A., El-Or, A. (2024). Bridging strategy and execution: Unlocking real digital transformation in the maritime sector. PwC Middle East.

Amadasu, O. (2024). The Integration of Artificial Intelligence and Blockchain in Logistics and Facility Management: A Case Study on Sea Cargo Handling Services. *International Journal of Engineering and Advanced Technology Studies* 12 (3), 1-17.

Bosch. (2026). One Platform for Logistics, Solution Providers & System Integrators, available at https://www.l-os.com/eu/en (May 2026).
Calsoft. (2026). Everything You Need to Know About CargoWise, available at https://www.calsoft.com/everything-you-need-to-know-about-cargowise/ (May 2026).
CargoWise Solution. (2026). Elevate your customer's supply chain with greater control, collaboration, and connectivity, available at https://www.cargowise.com/solutions/cargowise-neo/ (accessed on March 2026).
Christopher, M., & Peck, H. (2004). Building the resilient supply chain. *The International Journal of Logistics Management*, 15 (2), 1–14.
Chui, M., Manyika, J. Miremadi, M., Henke, N., Chung, R., Nel, P. (2018). Notes from the AI Frontier – Insights from Hundreds of Use Cases. McKinsey and Company: New York.
Descartes Group. (2026). Improving Logistics and Supply Chain Productivity, Performance, and Security, available at https://www.descartes.com/home (May 2026).
Du Plessis, M., Gerber, R., Goedhals-Gerber, L., Van Eeden, J. (2025). Shaping the future of freight logistics: Use cases of artificial intelligence. *Sustainability*, 17(3), 1355.
Durlik, I., Miller, T., Kostecka, E., Tuński, T. (2024). AI in maritime transportation: A comprehensive review of safety and risk management applications. *Applied Sciences*, 14(18), 8420.
European Commission. (2020). ePIcenter (Enhanced Physical Internet-Compatible Earth-friendly freight Transportation answER) – Project description. The EU Programme Horizon 2020, *Grant Agreement* No. 861584, CORDIS database, European Union.
European Maritime Safety Agency (EMSA). (2022). Modernised ADRIREP system: Technical requirements.
Gómez Ruiz, M. Á., de Almeida, I., Pérez Fernández, R. (2023). Application of machine learning techniques to the maritime industry. *Journal of Marine Science and Engineering*, 11(9), 1820.
Hesse, T., Peylo, C. (2019). Auf dem Weg zu einem intelligenten Mobilitätsraum. Lernende Systeme – Die Plattform für KI (Bericht der AG Mobilität und intelligente Verkehrssysteme).
Höhn, D., Mumm, L., Reitz, B., Tsiroglou, C., Hahn, A. (2025). Enabling future maritime traffic management: A decentralized architecture for sharing data in the maritime domain. *Journal of Marine Science and Engineering*, 13(4), 732.
Koch, P., Stach, T., Constapel, M., Burmeister, H.-Ch. (2022). VerifAI – Studie zur zielbasierten Standardisierung in der Prüfung und Zulassung intelligenter Entscheidungseinrichtungen von teilautonomen Überwasserfahrzeugen. Fraunhofer-CML und Dienstleistungen & Bundesamt für Seeschifffahrt und Hydrographie.
Komadina, P., Brčić, D., Frančić, V. (2013). VTMIS služba u funkciji unaprjeđenja sigurnosti pomorskog prometa i zaštite okoliša na Jadranu. *Pomorski zbornik* 47-48, 27-40.
Kuma, J. A. (2015). Vessel traffic service as a maritime security tool: vessel traffic management information systems (VTMIS) in Ghana. Dissertation, Malmo: World Maritime University.
Kumar, N. (2025). Artificial intelligence in maritime operations: Transforming the future of the global shipping industry. *International Journal of Maritime Technology (IJMTT)*, 6(2), 1–10.
Murrenhoff, A., Bedarf, M., Brügemann, L., van Bree, B., Ambra, T. (2023). AI in logistics (White paper). Alliance for Logistics Innovation through Collaboration in Europe (ALICE).
Noack, T., et.al. (DLR). (2017). Zuverlässige maritime Navigation und Verkehrslageerfassung im Kontext klassischer und neuer Sensoren, Systeme und Verfahren. 38. Duisburger Kolloquium Schiffstechnik/Meerestechnik.
Paladin Z., Scrima P., Vosinakis G., Hajduch G., Moutzouris M., Bolakis, C., et al. (2023). Maritime information sharing environment deployment using the advanced multilayered Data Lake capabilities: EFFECTOR project case study. *Sci. J. of Maritime Research*, 36(2), 291-304.
Paladin, Z., Scrima P., Jacobé de Naurois C., Rizogiannis C., Blum A., et al. (2022). Combined AI Capabilities for Enhancing Maritime Safety in a Common Information Sharing Environment. *Proceedings of 35th Bled e-Conference*, 145-160.
Paladin, Z., Šorović, M., Bauk, S., Knoors, F., Stephens, J., Kapidani, N., Lukšić, Ž., Mujalović, R. (2024). Maritime transport and logistics digital solutions optimization using advanced data

sharing platforms: ePIcenter project's case. *Proceedings of the 37th Bled eConference, University of Maribor Press*, 403–420.

Raza, Z., Woxenius, J., Vural, C., Lind, M. (2023). Digital transformation of maritime logistics: Exploring trends in the liner shipping segment. *Computers in Industry* 145, 103811.

Schulz, S., Krause, M., Stuck, J, et al. (2021). Digitalisierung maritime Wirtschaft: Positionen und Forderungen [Position paper]. Industrie und Handelskammer Schleswig-Holstein (IHK – SH).

Šorović, M., Brčić, D., Frančić, V., Maričević, T., Đurović, Z., Kapidani, N., Lukšić, Ž. (2023). Integrated information solution as a platform for data exchange in modernized ADRIREP system. *Proceedings of the 22nd International Symposium INFOTEH-Jahorina.*

Soumpenioti, V., Panagopoulos, A. (2023). AI technology in the field of logistics. 18th International Workshop on Semantic and Social Media Adaptation & Personalization (SMAP), IEEE.

Stach, T., Koch, P., Constapel, M., Burmeister, H., Gillert, A., Vahl, M. (2024). Studie zur Bewertung von Algorithmen für nautische Anwendungen (BAnA). Fraunhofer CML & Fraunhofer IGD.

Tejwani, D. (2024). Navigating the Future: How AI, big data, and autonomous systems are reshaping maritime transport. Article No. 125 [UNCTAD Transport and Trade Facilitation N°103].

UPS. (2026). UPS & the Power of Route Optimization, available at https://www.joneselitelogistics.com/blog/ups-the-power-of-route-optimization/(May 2026).

Vu, V., Le, P., Do, T., et al. (2024). An insight into the application of AI in maritime and logistics toward sustainable transportation. *Int. Journal on Informatics Visualization*, 8(1), 158–174.

Psychologically Grounded Chatbots in Online Furniture Retail

Claudia Pedron, Esther Federspiel

OST – Eastern Switzerland University of Applied Sciences, Institute for Information and Process Management (IPM), St. Gallen, Switzerland
claudia.pedron@ost.ch, esther.federspiel@ost.ch

High-involvement online purchase decisions are characterized by extensive information processing, perceived risk, and uncertainty. These challenges are particularly visible in furniture retail, where financial investment, long usage cycles, limited physical experience, and symbolic relevance make the purchase decision particularly complex. Although conversational AI is increasingly used in digital commerce, many chatbots remain transaction-oriented and insufficiently address psychological decision dynamics. Using a design-oriented conceptual research approach informed by design science principles, this article develops a phase-sensitive conceptual reference model for psychologically grounded conversational AI in high-involvement online purchasing. Furniture retail serves as an illustrative application context. Chatbots are conceptualized as dialogical decision companions whose roles evolve across the purchasing process, from preference exploration and risk reduction to trade-off moderation and decision stabilization. The model integrates psychological factors such as perceived risk, trust, aesthetic preferences, and sustainability values, as well as chatbot-related design factors such as personalization, usefulness, and interaction quality.

DOI
https://doi.org/10.18690/um.fov.4.2026.16

ISBN
978-961-299-152-4

Keywords:
conversational AI, high-involvement purchase decisions, chatbot-based decision support, consumer purchase decision process, psychologically grounded design, design science research

University of Maribor Press

1 Introduction

Electronic markets have fundamentally changed how consumers search for information, compare alternatives, and make purchase decisions by reducing search costs, increasing transparency, and enabling data-driven personalization (Huang & Rust, 2021; Verhoef et al., 2021; Wedel & Kannan, 2016). These developments are particularly relevant for high-involvement online purchase decisions, which, in contrast to low-involvement decisions, require extensive information processing and are characterized by financial, functional, and social uncertainty. In online settings, such challenges are intensified by the lack of physical product experience and by uncertainty regarding quality, fit, and return processes (Yu et al., 2023; Zhang et al., 2022). Accordingly, digital interaction systems increasingly serve to support consumers, foster trust, and reduce perceived uncertainty.

AI-supported chatbots have become established in service and sales contexts as interactive systems that structure information and guide consumer decision-making. Prior research shows that such systems can improve personalization, responsiveness, and decision confidence (Huang & Rust, 2021; Luo et al., 2019; Verhagen et al., 2014). Yet existing applications largely focus on efficiency, automation, and product-centered recommendation logic, while psychological decision dynamics such as risk perception, preference construction, and affective stabilization remain insufficiently addressed. Similarly, prior research on AI-supported chatbots in e-commerce mainly examines adoption factors, specific design cues, performance outcomes, or operational implementation models (Iniesta & López, 2025; Li et al., 2025; Sidlauskiene et al., 2023; Skuridin & Wynn, 2024). Consequently, limited guidance exists on how chatbots can address consumers' changing psychological needs as they vary across the phases of high-involvement purchase decisions. Thus, there is a need for a theory-based conceptual model that links psychological decision factors to phase-specific chatbot design principles in high-involvement online purchase decisions. This article addresses this gap and examines the following research question:

RQ1: How can a theory-based conceptual reference model be developed for designing psychologically grounded, phase-sensitive chatbots for high-involvement online purchase decisions?

To operationalize this question, two sub-questions are examined:

- RQ1a: Which psychological factors are relevant to high-involvement online purchase decisions?
- RQ1b: How can these factors be systematically integrated into the design of an AI-supported chatbot across the phases of the purchasing process?

Furniture retail is used as an illustrative application context because furniture purchases typically involve high cognitive effort, intensive information processing, long decision times, and considerable financial risk (Pepels, 2013). In addition, furniture often carries symbolic and social meaning by shaping living spaces and expressing lifestyle and identity (Foscht et al., 2017). Thus, the context is suitable for developing a transferable conceptual reference model for high-involvement online purchasing.

The remainder of the article is structured as follows: Section 2 presents the theoretical framework, Section 3 describes the research design, Section 4 introduces the conceptual reference model, and Section 5 discusses the implications and concludes the article.

2 Theoretical Background

The development of the conceptual reference model combines a process perspective on purchasing decisions with a psychological impact perspective. To this end, the Engel-Kollat-Blackwell (EKB) model is used to structure the decision-making process, while the stimulus-organism-response (S-O-R) model explains the internal mechanisms through which external stimuli shape decision-related responses. These perspectives are complemented by psychological factors relevant to high-involvement online purchases and by selected chatbot-related design factors derived from prior research. Together, they provide the conceptual foundation for the model developed in this study.

2.1 The S-O-R model as a psychological impact perspective

The S-O-R model explains purchasing behavior as the result of internal processing triggered by external stimuli (Kotler et al., 2007). Stimuli such as product information, prices, interface cues, or social signals are processed cognitively and affectively and may lead to observable responses such as purchase, continued search, or decision abandonment. A key implication is that consumer behavior is shaped not by objective product characteristics alone, but by their subjective interpretation.

In chatbot-supported purchasing, this means that chatbots do not directly determine decisions, but can influence pre-decisional factors such as trust, perceived usefulness, motivation, and perceived risk.

2.2 The EKB model as a process perspective on purchasing

While the S-O-R model explains how external stimuli are internally processed, the EKB model structures purchasing as a sequence of decision phases, from problem recognition to information search, evaluation of alternatives, purchase, and post-purchase evaluation (Kotler et al., 2007; Pepels, 2013). This process perspective is particularly relevant in high-involvement online purchasing, where decisions develop over time and involve repeated comparison, reassessment, and justification (Huang & Rust, 2021; Luo et al., 2019). The EKB model therefore provides the temporal structure within which chatbot support can be conceptualized.

2.3 Psychological factors influencing high-involvement online purchases

To apply these perspectives to chatbot-supported purchasing, it is necessary to identify the psychological factors that are particularly relevant in high-involvement online decisions. In this study, three interrelated categories are distinguished: cognitive, affective, and value-based factors.

Cognitive factors refer to rational, information-based evaluation and judgment processes through which consumers assess alternatives, uncertainty, costs, and expected benefits (Rani, 2014). In high-involvement online purchases, they are especially relevant during information search and evaluation. Perceived risk is particularly salient when product quality cannot be physically inspected in advance

(Zhang et al., 2022). Performance risk, financial risk, and uncertainty about fit or logistics intensify information search and comparison, while trust helps reduce uncertainty and stabilize decision-making. Transparent product information, warranties, origin disclosures, and clear service information can strengthen perceived trustworthiness (Yu et al., 2023).

Affective factors refer to emotional reactions and feeling-based evaluations that shape how consumers perceive alternatives and form preferences beyond purely functional assessment (Rani, 2014). They are especially relevant in high-involvement purchases that involve identity, aspiration, or symbolic meaning. Emotional resonance, aesthetic liking, and the anticipation of comfort or status influence both the attractiveness of alternatives and the formation of preferences. In furniture retail, as an illustrative context, these mechanisms are particularly visible because design and aesthetic perception are core parts of the decision itself (Balicka & Niedbała, 2022).

Value-based factors refer to relatively stable beliefs, priorities, and normative orientations that guide how consumers frame, justify, and evaluate consumption choices over time (Rani, 2014). They extend beyond immediate product appraisal and influence both problem framing and post-purchase justification. Sustainability considerations, ownership preferences, and materialistic orientations may shape what consumers regard as an acceptable or desirable solution.

2.4 Chatbot-related design factors

In addition to these psychological decision factors, prior chatbot research highlights several design-related perceptions that shape whether and how chatbot support is used. Relevant constructs include information value, performance expectancy, perceived usefulness, effort expectancy, facilitating conditions, perceived risk, anthropomorphism (i.e., the extent to which the chatbot is perceived as human-like), and perceived competence (Bansal et al., 2026; Kumar et al., 2025; Le, 2023). These constructs do not replace cognitive, affective, and value-based factors; rather, they describe design conditions through which chatbot interaction can influence them.

From a cognitive perspective, information quality, usefulness, ease of use, and performance expectancy are particularly relevant for uncertainty reduction and information support (Le, 2023).

From an affective perspective, communication style and anthropomorphic design may shape perceived warmth, authenticity, competence, and engagement (Kumar et al., 2025).

From a value-related perspective, facilitating conditions (e.g., ease of use and system support) and recommendation competence (i.e., the ability to provide relevant and preference-consistent suggestions) can support the alignment between chatbot interaction and consumers' broader priorities and preferences.

Advanced AI capabilities can further shape these effects by influencing perceived interaction quality, personalization, and decision support through S-O-R-related mechanisms (Bansal et al., 2026).

3 Research design

This study adopts a design-oriented conceptual research approach informed by design science research (DSR) principles (Hevner et al., 2004; Peffers et al., 2007). Rather than presenting a fully implemented and empirically validated DSR artifact, the study develops a conceptual reference model that can guide the future design and evaluation of psychologically grounded chatbot systems for high-involvement online purchasing. Furniture retail is used as an illustrative application context to ground the model development.

Following the DSR process model proposed by Peffers et al. (2007), five steps were undertaken:

- 1. Problem identification: The practical problem was specified through an exploratory market analysis of AI-supported applications in online furniture and interior retail. Relevant tools were identified through structured web searches using keywords such as *AI furniture chatbot*, *AI interior design tool*, *AI room planner*, and *AI furniture recommendation*. Provider websites, app store entries, and third-party overview pages were screened iteratively, with

additional cases identified through cross-referencing until no substantially new tool categories emerged. Tools were included if they offered recognizable AI-supported functionality related to furniture or interior decision-making. The final sample comprised around 20 tools and platforms and was assessed using predefined categories covering AI functionality, support across purchase phases, and psychologically relevant support features; the aim was analytical coverage rather than statistical representativeness. The analysis showed that existing applications provide various forms of AI support but offer limited integrated support for psychologically relevant decision processes across purchase phases. A summary of the analyzed tools and assessment criteria is provided in Appendix Table A1.

- 2. Objective definition: Based on the identified discrepancy between theoretically relevant decision dynamics and the functional scope of existing systems, the design objective was defined as the development of a phase-sensitive conceptual reference model for AI-supported chatbot design. The model is intended to support consumers not only at the transactional level, but also in psychologically relevant processes such as uncertainty reduction, preference clarification, confidence building, and post-purchase support.
- 3. Design and development: The conceptual reference model was developed by integrating the theoretical foundations discussed in Section 2, namely the EKB model as a phase structure of the purchase process, the S-O-R framework as an explanatory lens for internal processing, psychological factors relevant to high-involvement online purchase decisions, and chatbot-related design factors. The market analysis served as a practical benchmark for identifying which support functions are already addressed in existing AI-based applications and which remain underrepresented. Based on these theoretical and practical insights, phase-specific chatbot roles and corresponding design requirements were derived.
- 4. Conceptual demonstration: The resulting model was conceptually demonstrated through illustrative decision scenarios in online furniture retail. These scenarios are reflected in Section 4, where the model illustrates how phase-specific chatbot functions address uncertainty about product fit, difficulties in comparing alternatives, low confidence before purchase completion, and the need for reassurance after purchase. The

demonstration shows how the proposed model extends existing product- and transaction-oriented systems by integrating psychologically grounded support across purchase phases.

- 5. Conceptual evaluation: The model was evaluated conceptually in terms of internal consistency, theoretical coherence, and practical plausibility, following Gregor and Hevner (2013). Specifically, the evaluation examined the alignment between decision phases, psychological mechanisms, and chatbot functions, the consistency of the model with its theoretical foundations, and its plausibility in light of the market analysis. Practical plausibility was additionally informed by discussions with representatives of an interior-design software company and interior design consultants. As the study develops a conceptual model, the evaluation remains analytical rather than empirical and provides a basis for future prototypical implementation, empirical validation, and subsequent design science iterations.

4 A phase-sensitive model of psychologically integrated chatbot support

The conceptual reference model integrates the theoretical foundations with insights from the exploratory market analysis. It conceptualizes chatbot interaction as phase-sensitive decision support across the purchase journey. While existing AI-supported applications mainly address inspiration, visualization, product discovery, and transaction-related assistance, the model focuses on psychologically demanding decision processes that remain insufficiently covered in current systems.

The model is structured along the five phases of the purchase process and links each phase to dominant decision requirements. Accordingly, the chatbot's role changes over time: from clarifying preferences and structuring information to moderating trade-offs, stabilizing the purchase decision, and supporting post-purchase evaluation. Figure 1 summarizes this logic by linking the phases of the purchase process to dominant psychological decision requirements and corresponding chatbot roles.

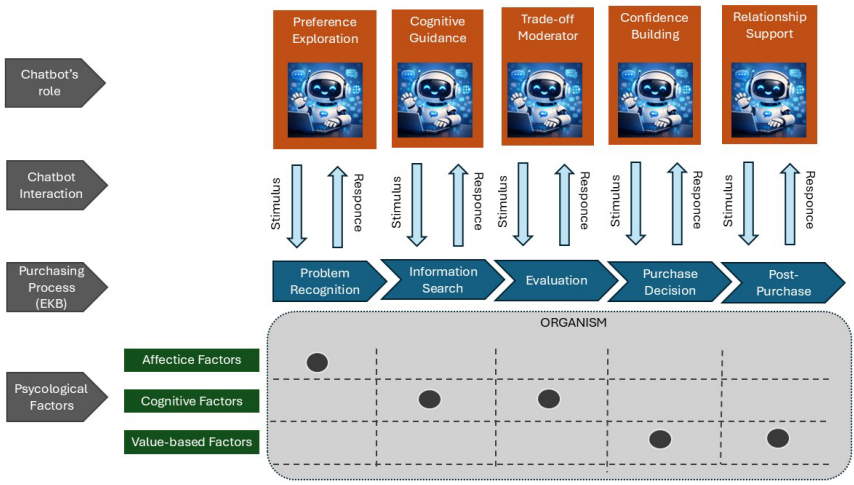

Figure 1: Phase-sensitive conceptual reference model for psychologically integrated chatbot support
Source: Own Illustration

4.1 Problem identification: Elicitation of latent needs and contextual framing

In the problem-identification phase, the chatbot helps consumers articulate diffuse and often affective needs in more explicit preference terms. Rather than recommending products immediately, it supports the clarification of latent preferences by translating moods, atmospheres, and contextual cues into decision-relevant criteria. This role builds on the affective and value-based factors discussed in Section 2.3, as these factors are particularly relevant when consumers have not yet translated subjective needs into explicit decision criteria.

For example, a desire for "coziness" may be specified in terms of preferred materials, colors, or stylistic directions. Style categories and situational cues, such as usage context or life stage, can further support this clarification. The outcome of this phase is therefore not product selection, but the development of a coherent preference structure.

4.2 Information search: Reducing perceived risk and building trust

In the information-search phase, the chatbot acts as a structuring information broker. Its task is to reduce uncertainty by presenting relevant product and service information in a clear, comprehensible, and decision-relevant way. Instead of merely listing product attributes, the chatbot translates them into functional implications and can clarify delivery conditions, return policies, warranties, or sustainability-related aspects. This role reflects the cognitive factors discussed in Section 2.3, particularly perceived risk, information processing, and trust formation, as well as chatbot-related constructs such as information value and perceived usefulness discussed in Section 2.4.

The primary objective in this phase is to make information easier to evaluate and to reduce uncertainty before alternatives are compared more closely.

4.3 Evaluation of alternatives: Moderation of goal conflicts and social dynamics

In the evaluation phase, the chatbot supports consumers in comparing alternatives across criteria such as design, price, quality, sustainability, and usage context. Rather than hierarchically ranking options, it contextualizes alternatives in relation to the consumer's previously articulated preferences. This role reflects the cognitive, affective, and value-based factors discussed in Section 2.3, particularly comparison, preference formation, aesthetic evaluation, and value-related considerations, as well as chatbot-related constructs such as recommendation competence and perceived usefulness discussed in Section 2.4.

For instance, the chatbot may contrast a product with stronger aesthetic appeal with another option that is more robust and easier to maintain, and thus more suitable for households with children. Since furniture purchases are often made jointly, the chatbot can also act as a neutral mediator by highlighting differing preferences and structuring compromise options. Its role in this phase is therefore to support preference-consistent trade-off reasoning rather than to prescribe a single optimal choice.

4.4 Purchase decision: Legitimation and development of decision confidence

Shortly before purchase completion, the chatbot supports decision stabilization by linking the chosen option to previously articulated preferences and by providing transparency about the practical conditions of purchase. For example, it may summarize how the selected product fits the consumer's priorities regarding design, sustainability, and budget, while also clarifying delivery times and return options. This role is grounded in the importance of trust, perceived risk reduction, and decision confidence in high-involvement purchase decisions (Section 2.3), as well as in chatbot-related design factors such as transparency, perceived usefulness, and interaction quality (Section 2.4).

The purpose of chatbot support in this phase is not to reopen comparison, but to reinforce confidence in the selected option. By synthesizing relevant criteria and making residual conditions transparent, the chatbot helps legitimize the decision and reduce lingering uncertainty.

4.5 Transition to the post-purchase phase: Preventive dissonance reduction

At the transition to the post-purchase phase, the chatbot supports consumers by reaffirming the fit between the final choice and the needs and preferences articulated earlier in the process. For instance, it may confirm that the selected product corresponds to the consumer's preference for a warm and calm atmosphere. This role reflects the importance of reassurance and decision confidence in high-involvement purchase decisions, as discussed in Section 2.3, and is further supported by chatbot-related design factors such as perceived competence and consistent communication (Section 2.4).

At this stage, the chatbot should not encourage further comparison, but help consumers feel confident about the decision already made. By summarizing the reasons for the choice and making its consistency with earlier preferences explicit, the chatbot can reduce post-decisional uncertainty and support the transition from purchase completion to post-purchase evaluation.

5 Discussion, implications, and conclusion

The proposed model suggests that AI-supported chatbots in high-involvement online purchasing contexts should not be conceptualized merely as extended search or recommendation systems. In such settings, consumers face elevated perceived risk, longer evaluation cycles, and more demanding cognitive, affective, and value-based decision processes. Chatbots should therefore provide phase-sensitive support that addresses uncertainty reduction, preference clarification, structured comparison, reassurance, and decision confidence.

The study contributes to design-oriented research by developing a theory-based conceptual reference model that combines the EKB model, the S-O-R framework, psychological purchase factors, and chatbot-related design factors into a phase-sensitive design logic for high-involvement decisions. By embedding design features such as trust, personalization, anthropomorphism, and usefulness in this logic, the model extends prior chatbot research and complements more operational chatbot models with a psychologically grounded perspective across the full customer journey. It shows how chatbot roles evolve from preference elicitation to uncertainty reduction, trade-off moderation, decision stabilization, and post-purchase reassurance.

For practice, this implies designing chatbots as strategic interaction systems rather than isolated automation tools. The market analysis indicates that current systems mainly support inspiration, visualization, and transaction efficiency, while deeper psychological support remains limited. Effective chatbot design should therefore support trust building, trade-off moderation, reassurance, and post-purchase stabilization, while ensuring transparency and argumentative clarity to avoid perceptions of manipulation. Implementation also requires enriched data structures that extend beyond product attributes to include semantic and contextual metadata reflecting relevant psychological dimensions.

Future research should empirically validate the model and examine its effects on trust, perceived risk, decision confidence, chatbot usage, and longer-term customer relationships.

Acknowledgment

This work was supported by an Innosuisse InnoCheck project, a Swiss innovation funding instrument that supports the initial assessment of applied research and innovation ideas. The authors gratefully acknowledge the valuable discussions with Bruno Anderegg and Frank Wicht from the software company EasternGraphics (https://www.easterngraphics.com/), which helped refine the practical relevance of the proposed conceptual reference model.

AI-based language tools (ChatGPT, OpenAI) were used to improve the linguistic quality of the manuscript. All intellectual contributions remain the responsibility of the authors.

References

Balicka, A., & Niedbała, M. (2022). Social factors affecting consumer decisions during purchasing furniture in Poland. Annals of Warsaw University of Life Sciences – SGGW. Forestry and Wood Technology, 118, 22–34. https://doi.org/10.5604/01.3001.0015.9775

Bansal, S., Nangia, P., Chanaliya, N., & Thaichon, P. (2026). The role of advanced AI in improving chatbot-driven consumer experiences. *Strategic Business Research, 2*, 100023. https://doi.org/10.1016/j.sbr.2025.100023

Foscht, T., Swoboda, B., & Schramm-Klein, H. (2017). Käuferverhalten: Grundlagen – Perspektiven – Anwendungen (5. Aufl.). Springer Gabler. https://doi.org/10.1007/978-3-658-17465-1

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. MIS Quarterly, 37(2), 337–355. https://doi.org/10.25300/MISQ/2013/37.2.01

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. MIS Quarterly, 28(1), 75–106. https://doi.org/10.2307/25148625

Huang, M.-H., & Rust, R. T. (2021). Engaged to a Robot? The Role of AI in Service. Journal of Service Research, 24(1), 3–19. https://doi.org/10.1177/1094670520902266

Iniesta, M. B., & López, D. (2025). A systematic literature review on conversational AI's influence in consumer behavior. In Decision sciences (Lecture Notes in Computer Science, Vol. 14778, pp. 27–39). Springer. https://doi.org/10.1007/978-3-031-78238-1_3

Kotler, P., Keller, K. L., & Opresnik, M. O. (2007). Marketing-Management (12. Aufl.). Pearson.

Kumar, A., Shankar, A., Behl, A., Chakraborty, D., & Gundala, R. R. (2025). Anthropomorphic generative AI chatbots for enhancing customer engagement, experience and recommendation. *Journal of Consumer Marketing, 42*(4), 472–483. https://doi.org/10.1108/JCM-06-2024-6922

Le, X. C. (2023). Inducing AI-powered chatbot use for customer purchase: The role of information value and innovative technology. *Journal of Systems and Information Technology, 25*(2), 219–241. https://doi.org/10.1108/JSIT-09-2021-0206

Li, Y., Gan, Z., & Zheng, B. (2025). How do artificial intelligence chatbots affect customer purchase? Uncovering the dual pathways of anthropomorphism on service evaluation. Information Systems Frontiers, 27, 283–300. https://doi.org/10.1007/s10796-023-10438-x

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. Humans: The impact of AI Chatbot Disclosure on Customer purchases. Marketing Science, 38(6), 937–947. https://doi.org/10.1287/mksc.2019.1192

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Pepels, W. (2013). Konsumentenverhalten (2. Aufl.). UVK.

Rani, P. (2014). *Factors influencing consumer behaviour. International Journal of Current Research and Academic Review, 2*(9), 52–61. https://www.ijcrar.com/vol-2-9/Pinki%20Rani.pdf

Sidlauskiene, J., Joye, Y., & Auruskeviciene, V. (2023). AI-based chatbots in conversational commerce and their effects on product and price perceptions. Electronic Markets, 33, Article 24. https://doi.org/10.1007/s12525-023-00633-8

Skuridin, A., & Wynn, M. (2024). Chatbot design and implementation: Towards an operational model for chatbots. Information, 15(4), 226. https://doi.org/10.3390/info15040226

Verhagen, T., van Nes, J., Feldberg, F., & van Dolen, W. (2014). Virtual customer service agents: Using social presence and personalization to shape online service encounters. Journal of Computer-Mediated Communication, 19(3), 529–545. https://doi.org/10.1111/jcc4.12066

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. Journal of Business Research, 122, 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. Journal of Marketing, 80(6), 97–121. https://doi.org/10.1509/jm.15.0413

Yu, C., Liu, W., Fei, Y., Chen, J., & Hu, Z. (2023). Influencing factors of online furniture purchase behavior based on Analytic Hierarchy Process. BioResources, 18(2), 2857–2873. https://doi.org/10.15376/biores.18.2.2857-2873

Zhang, S., Zhu, J., Wang, G., Reng, S., & Yan, H. (2022). Furniture online consumer experience: A literature review. BioResources, 17(1), 1627–1642. https://doi.org/10.15376/biores.17.1.1627-1642

About the authors

Claudia Pedron

Claudia Pedron, Prof. Dr., is Professor of Business Information Systems at the Institute for Information and Process Management (IPM), OST – Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland. She has extensive experience in business process management, process automation, IT and management systems, and digitalization. She has acquired and led numerous research and service projects in process management, digital transformation, real estate technology, and decision-support systems. She holds a Ph.D. from ETH Zurich and an Executive Master in Business Engineering from the University of St. Gallen. Her publications address process-oriented decision support, digital services, e-health, and technology-supported management systems.

Esther Federspiel

Esther Federspiel, lic.phil., is a lecturer and project manager at the Institute for Information and Process Management (IPM), OST – Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland. Her professional background includes teaching, project management, marketing management, communication, and public relations. She studied journalism/public communication and psychology at the University of Zurich and holds a MAS in Communication Management from Lucerne University of Applied Sciences and Arts. Her work focuses on innovation culture, customer and employee relationships, customer integration in innovation and value creation processes, online communities, gamification, and decision processes under uncertainty.

Summary

This paper develops a theory-based conceptual reference model for designing AI-supported chatbots in high-involvement purchase contexts, using online furniture retail as an illustrative case. While existing chatbot systems primarily focus on product recommendations and transactional efficiency, they rarely integrate psychological decision dynamics. Building on the Stimulus–Organism–Response framework and the five-stage consumer decision process, the study conceptualizes chatbots as psychologically grounded, phase-sensitive decision companions. The proposed model outlines how chatbots can support preference exploration, risk reduction, trade-off moderation, and decision stabilization. By integrating psychological determinants into conversational design, the paper extends decision support research in electronic markets and provides a foundation for future empirical validation.

Appendix A: Summary of Findings from the Exploratory Market Analysis

Appendix A summarizes the findings of the exploratory market analysis of AI-supported tools and platforms in the furniture and interior context. The table consolidates the observed functionalities of the analyzed tools and assesses how they support the five phases of the purchase decision process as well as selected psychological factors. The overview is intended to document the empirical basis for identifying the practical gap addressed by the conceptual reference model.

Table A1: Summary of Market Analysis Findings on AI-Supported Tools Across Purchase Phases and Psychological Factors

	Product name	Website / source	Brief tool description (AI focus)	Assessment across the five phases and psychological factors
1	Houzz	https://www.houzz.com/	AI-supported inspiration and product platform for home and interior design; personalizes ideas, products, and professionals based on style preferences and usage signals.	Strong in phases 1-3 (problem recognition, information search, evaluation of alternatives), moderate in phase 4, weak in phase 5. Psychological factors: aesthetics/inspiration strong; trust through community and experts moderate; sustainability/environmental orientation only selectively visible.
2	DesignMyRoom	https://apps.apple.com/us/app/design-my-room-official/id1664182248	AI app for rapid room redesigns based on a photo; generates several style variants for the visual exploration of interior design ideas.	Strong in phases 1-2, moderate in phase 3, weak in phases 4-5. Psychological factors: visual need exploration strong; risk, trust, social dynamics, and environmental aspects largely not integrated.
3	Homestyler	https://www.homestyler.com/	Cloud-based 3D planning platform with AI for floor-plan conversion, furnishing, and photorealistic visualization; supports style selection and scenarios.	Strong in phases 1-3, moderate in phase 4, weak in phase 5. Psychological factors: aesthetics, preference exploration, and comparison strong; perceived risk can be partially reduced through 3D visualization; sustainability largely not addressed systematically.
4	RoomsGPT	https://www.roomsgpt.io/	Generative AI for room redesign based on a photo; creates	Strong in phases 1-2, moderate in phase 3, weak in phases 4-5.

	Product name	Website / source	Brief tool description (AI focus)	Assessment across the five phases and psychological factors
			stylized alternatives for different room types within seconds.	Psychological factors: emotional resonance and style discovery strong; risk, trust, and sustainability dimensions only very limited.
5	Planner 5D	https://planner5d.com/	2D/3D home and room planner with AI for floor-plan recognition, layout suggestions, and furnishing; combines visualization with planning functions.	Strong in phases 1-3, moderate to strong in phase 4, weak in phase 5. Psychological factors: comparison, spatial confidence, and functional fit good; affective and social factors partially addressed; sustainability not central.
6	Havenly	https://havenly.com/	Digital interior design platform with personalized suggestions, 3D visualizations, and purchasing options; AI supports matching and personalization.	Strong in phases 1-4, moderate in phase 5. Psychological factors: style, budget, and advisory needs are well addressed; trust is stronger through designer involvement than in purely AI-based tools; environmental aspects only limited.
7	Modsy (historical / integrated into Lennar)	https://www.modsy.com/	Historical 3D interior design solution with AI-supported room layouts and shoppable design suggestions; according to the website, integrated into Lennar since 2023.	Historically strong in phases 1-4, weak in phase 5. Psychological factors: visual confidence, comparison, and decision confidence relatively good; sustainability and deeper psychological/social modelling limited.
8	AiHouse	https://my.aihouse.com/	AI-supported 3D interior and manufacturing platform with room planning, renderings, and automatic structure recognition; more professionally oriented.	Strong in phases 2-4, moderate in phase 1, weak in phase 5. Psychological factors: risk and functional assessment good; emotional need exploration and environmental values less strong; more planning-oriented than dialogue-oriented.
9	Design Flex	https://www.cyncly.com/products/design-flex/	Described in the market document as an interactive 3D environment with real-time	According to the document description, strong in phases 2-3, moderate in phases 1 and 4, weak in phase 5.

	Product name	Website / source	Brief tool description (AI focus)	Assessment across the five phases and psychological factors
			visualization, automatic furniture arrangement, and AI suggestions.	Psychological factors: functional visualization good, but no clear evidence of trust, social moderation, or environmental aspects.
10	Home Design AI	https://home-design.ai/	AI tool for redesigning interiors and floor plans based on photos and prompts; quickly and accessibly generates design ideas and layout variants.	Strong in phases 1-2, moderate in phase 3, weak in phases 4-5. Psychological factors: need exploration and aesthetics strong; trust, risk reduction, sustainability, and post-purchase support hardly developed.
11	Wayfair	https://www.wayfair.com/	Large online furniture retailer using AI for search, recommendations, assortment personalization, and marketing optimization across the digital purchase process.	Strong in phases 2-4, moderate in phases 1 and 5. Psychological factors: personalization, convenience, and choice reduction strong; deeper psychological dialogue guidance, social negotiation, and sustainability integration only partial.
12	Ashley Furniture	https://www.ashleyfurniture.com/	Furniture retailer with widely used AI/digital applications for commerce, assortment, marketing, and customer experience; focus is more operational-commercial.	Moderate to strong in phases 2-4, weak to moderate in phases 1 and 5. Psychological factors: convenience and product search supported; phase-sensitive decision guidance and environmental values not prominent.
13	Coleman Furniture	https://colemanfurniture.com/	Online furniture retailer with visual product search and "shop the look" approaches; AI primarily supports discovery and inspirational product exploration.	Strong in phases 1-3, moderate in phase 4, weak in phase 5. Psychological factors: visual inspiration and comparison of alternatives good; trust, risk reduction, and sustainability only selectively addressed.
14	Freedom Furniture	https://www.freedom.com.au/	Omnichannel furniture retailer with AI-supported search and personalization to improve discovery, relevance, and cart value.	Moderate to strong in phases 2-4, moderate in phase 1, weak in phase 5. Psychological factors: personalization good, but psychological depth models, social

	Product name	Website / source	Brief tool description (AI focus)	Assessment across the five phases and psychological factors
				moderation, and environmental factors are not comprehensively visible.
15	Temple & Webster	https://www.templeandwebster.com.au/	Online retailer using AI in customer service, logistics, product creation, and personalization; focused on scaled commerce support.	Moderate to strong in phases 2-4, weak to moderate in phases 1 and 5. Psychological factors: search and service efficiency present; deeper preference exploration, risk dialogue, and sustainability advice rather limited.
16	Sofology	https://www.sofology.co.uk/	Sofa retailer with visual image search or AI-supported product search to find suitable models based on inspirational references.	Strong in phases 1-3, moderate in phase 4, weak in phase 5. Psychological factors: style and image preferences good, but little evidence of comprehensive trust, environmental, or social decision support.
17	IKEA (including IKEA Kreativ)	https://www.ikea.com/c/en/home-design/	Global furniture provider with virtual room planning, AR/scan functions, and IKEA Kreativ as an AI-supported room designer for budget- and style-related product selection.	Strong in phases 1-4, moderate in phase 5. Psychological factors: good combination of inspiration, risk reduction, and budget orientation; sustainability more visible at brand level than in many other cases, but not deeply integrated in a dialogic manner.
18	Blueport	https://www.blueport.com/about-us	Omnichannel e-commerce platform for big-ticket retail that provides furniture retailers with personalized digital shopping experiences and commerce functions.	Indirect fit: strong for phases 2-4 on the retailer side, but not an end-customer chatbot. Psychological factors are technologically supported, but not implemented as an independent dialogical decision companion with an environmental/value layer.
19	Vaimo	https://www.vaimo.com/	Digital commerce and experience provider; illustrates how AI search, recommendation, and advisory logic can be	Indirect fit: strong for phases 2-4 as enablement, weaker for phases 1 and 5. Psychological and environmental factors are more implementation

	Product name	Website / source	Brief tool description (AI focus)	Assessment across the five phases and psychological factors
			integrated into retailer ecosystems.	options than a visible core function of the product.
20	Quin AI	https://www.quinengine.com/	Generative AI platform for real-time behavioral predictions and personalized website/app experiences; optimizes engagement and conversion in-session.	Strong in phases 2-4, moderate in phase 1, weak in phase 5. Psychological factors: behavior and intent are anticipated in a data-driven way; explicit sustainability or socio-cultural reflection remains limited.

Source: Own elaboration

Exploring AI Adoption for Manufacturing SMEs in Industry 5.0

Lars-Olof Johansson
Halmstad University, School of IT, Halmstad, Sweden
lars-olof.johansson@hh.se

This article explores AI adoption in manufacturing SMEs in Industry 5.0. The authors present an interdisciplinary approach to AI adoption grounded in literature from information systems, organization research, and industry research. The aim of the article is to identify characteristics of AI adoption in manufacturing SMEs in Industry 5.0. In information systems, researchers claim that the rise of AI technologies is prompting growing calls to rethink the theoretical foundations of adoption in IS research. The research approach was inspired by interpretive field studies and was conducted within the ongoing project, HUMAI. The article's contribution is a set of characteristics of AI adoption in manufacturing SMEs in Industry 5.0. Three characteristics were identified: 1) The AI adoption should support a manufacturing SME's sustainability work, 2) A Socio-technical and human-centric perspective should be applied throughout the AI adoption process, and 3) Competence development and learning should be addressed throughout the AI adoption process.

DOI
https://doi.org/10.18690/um.fov.4.2026.17

ISBN
978-961-299-152-4

Keywords:
AI adoption,
SME,
industry 5.0,
socio-technical,
Bled eConference

University of Maribor Press

1 Introduction

In the contemporary business landscape, the strategic integration of Artificial Intelligence (AI) has become a critical determinant of competitive advantage and long-term sustainability. For Small and Medium-sized Enterprises (SMEs), which form the backbone of most global economies. Within the EU (European Union), only 12.6% of SMEs are using AI (European Commission 2025). European SMEs, which represent over 90% of European enterprises, are facing challenges in adopting AI. One of the critical sectors highlighted by the EU is the European manufacturing sector, which comprises 2.2 million enterprises, mostly SMEs. For manufacturing industries in Europe, the EU policy Industry 5.0 (European Commission 2022) is of great importance. The vision of Industry 5.0 rests on three pillars: resilience, sustainability, and human-centricity. Resilience addresses agility and flexibility with adaptable technologies (Such as AI). Sustainability - leads action on sustainability and respects planetary boundaries, and human-centric - promotes talent, diversity, and empowerment. It is through human-centric approaches combined with digital devices that the goals of Industry 5.0 can be achieved.

Research in organizational science has shown that SMEs differ. SMEs are typically characterized by flat and informal structures and weakly formalized procedures (Torrès and Julien 2005). Roles often overlap, and coordination relies heavily on direct communication rather than standardized processes. In other words, a small-sized firm does not necessarily have to adhere to the classical management method. The management process in small firms is unique and cannot be considered to be the same as professional management in larger organizations, practiced on a reduced scale (Jennings and Beaver 1997). SME strategy is often emergent rather than deliberate, embedded in day-to-day operations rather than articulated in formal plans (Mintzberg and Waters 1985).

Information systems (IS) research has long examined how IT/IS innovations are adopted (Akhlaghpour, Karahanna et al. 2025). The rise of AI technologies introduces new adoption logics and patterns of use. These developments have prompted growing calls to rethink the theoretical foundations of adoption in IS research. Moreover, research indicates that a significant majority of SMEs are still in the early stages of their AI journey, underscoring a widespread need for foundational guidance and support (Khan, Kauttonen et al. 2025). Organizational AI research

does not yet reflect organizational adoption and the impact that current-generation AI technologies will have (Nguyen, Sidorova et al. 2022). Hence, the article aims to identify characteristics of AI adoption in manufacturing SMEs in Industry 5.0. The research question is: What are the characteristics of AI adoption in manufacturing SMEs in Industry 5.0?

2 AI adoption, Industry 5.0, and SMEs

The adoption of Artificial Intelligence (AI) has become a critical determinant of competitive advantage for SMEs in the business landscape. For Small and Medium-sized Enterprises (SMEs), AI represents a transformative opportunity to enhance efficiency and unlock new value. However, the path to AI adoption is fraught with unique challenges that often hinder SMEs from realizing their full potential. By examining current research on AI adoption, studies on SMEs, and EU policy documents, an AI adoption framework for SMEs in Industry 5.0 will be developed.

2.1 SMEs and AI adoption

Research on SMEs clearly shows that they are different. SMEs are directly mentioned in (European Commission, 2025) Apply AI strategy (Fig. 1). Torrès and Julien (2005) state that SMEs need not adhere to the classical management method. SMEs are typically characterized by flat and informal structures, limited specialization, and weakly formalized procedures (Torrès and Julien 2005). And as such, the management process in small firms is unique (Jennings and Beaver 1997). Moreover, roles in the SME often overlap, and coordination relies heavily on direct communication rather than standardized processes. SMEs also face resource constraints in terms of capital, competence (Fig. 1), and time, which shape their organizational behavior (Barney 1991). A limitation on resources is that SMEs lack the competence, time, and finances to undertake complex, often costly AI projects (Knoblach, Vogl et al. 2025). A frequently cited obstacle is the shortage of technical knowledge (Fig. 1) and the absence of in-house expertise required to develop, implement, and maintain AI systems (Khan, Kauttonen et al. 2025, Knoblach, Vogl et al. 2025). SMEs are also structurally more fragile, with high dependence on key individuals, customers, or technologies. Learning in SMEs is predominantly informal, experiential, and embedded in work practices rather than organized through formal training systems (Billett 2001).

One of the findings from Khan, Kauttonen et al. (2025) is that a significant majority of SMEs are still in the preliminary stages of their AI journey, indicating a widespread need for foundational guidance and support. From a technical perspective, there are two challenges for SMEs: data management and incompatibility. Many SMEs struggle with data availability, poor data quality, and the lack of a coherent strategy for data collection and management (Fig. 1), which are foundational prerequisites for most AI applications (Khan, Kauttonen et al. 2025, Knoblach, Vogl et al. 2025). AI solutions are often incompatible with the existing devices, software, and legacy systems prevalent in many SMEs, creating a significant integration barrier (Knoblach, Vogl et al. 2025). Further on, Nguyen, Sidorova et al. (2022) claim that organizational AI research does not yet reflect organizational adoption and the impact that current-generation AI technologies will have (Fig. 1).

Information systems (IS) research has long examined how IT/IS innovations are adopted (Akhlaghpour, Karahanna et al. 2025). TAM (Technology Acceptance Model (Davis 1989)), TOE (Technology-Organization-Environment (Tornatzky, Fleischer et al. 1990)), and UTAUT (Unified Theory of Acceptance and Use of Technology (Venkatesh, Morris et al. 2003)) are all examples of frameworks used to understand IT adoption. The rise of AI technologies introduces new adoption logics and patterns of use. Developments in AI have prompted growing calls to rethink the theoretical foundations of adoption in IS research (Fig. 1). Moreover, AI adoption has been framed as a fundamentally socio-technical challenge that demands attention to emotions, values, and lived experience that complement human cognition and emotion, rather than inadvertently undermining them. A well-known socio-technical framework (Task, Technology, People, and Structure) for organizational change is Leavitt (1965). This framework has been modified by Grant and Mergen (1996) to include communication as an important aspect. In Yu, Xu et al. (2023), a socio-technical perspective is presented to understand AI adoption. In the article, three research agendas are proposed: 1) AI and training, performance, feedback, and employee experience, 2)AI and organizational design, and 3) AI and employment. Moreover, a socio-technical system is presented that comprises three subsystems: the personnel subsystem, the technical subsystem, and the organizational structure/work/job design subsystem.

2.2 Industry 5.0

The vision of Industry 5.0 rests on three pillars (European Commission 2022): resilience, sustainability, and human-centricity. Resilience addresses agility and flexibility with adaptable technologies (Such as AI). Sustainability (Fig. 1) - leads action on sustainability in line with the Agenda 2030 (United Nations 2015) and respects planetary boundaries; human-centric - promotes talent, diversity, and empowerment. It is through human-centric approaches combined with digital devices that the goals of Industry 5.0 (Fig. 1) can be achieved. In Agenda 2030, one of the goals is to ensure sustainable consumption and production patterns (goal 12). The 17 goals of Agenda 2030 are also related to the three overlapping categories of sustainability proposed by Brundtland (1987): ecological, economic, and social.

In the vision of Industry 5.0, humans are positioned at the center of production and service systems (European Commission 2022). European industry is expected to thrive through the integration of human-centric approaches and digital technologies. This paradigm emphasizes the empowerment of individuals and highlights the importance of research to understand the complex interplay between humans and technology—particularly Artificial Intelligence (AI).

Even though human-centered design has been a subject of research for over five decades, the transition to Industry 5.0 presents new challenges. These include the need for a highly skilled workforce capable of engaging with advanced technologies (George and George 2023) and the imperative to ensure positive and meaningful human–technology interactions (Ferreira and Serpa 2018). In recent years, there has been a growing interest in Human-Centered Artificial Intelligence (HCAI). One researcher argues that "the success of Artificial Intelligence (AI)-powered services relies on human-centered design" (Wärnestål 2022).

The vision of human-centricity within Industry 5.0 can be examined on two interrelated levels: the technical and the organizational (European Commission 2024). On a technical level, human-centricity implies that the implementation of digital technologies—such as AI—should not only fulfill functional and operational requirements but also enhance human well-being, capabilities, skills, and working conditions. This perspective seeks to optimize the interaction between humans and machines by reinforcing the human role through technological integration. It

underscores the importance of maintaining individual autonomy and control while upholding core human values. At the organizational level, the human-centric approach prioritizes worker well-being and satisfaction by considering employees' unique needs, competencies, and viewpoints. This perspective fosters a culture of collaboration and inclusivity, promoting continuous learning and skill development. Such efforts are crucial to ensure that employees remain adaptable and equipped to navigate rapidly evolving technological landscapes.

2.3 Summary

The literature review has been summarized in Figure 1. In the model, three tentative characteristics have been identified: 1) the importance of government policies for AI adoption in manufacturing SMEs; 2) the differences and importance of SMEs; and 3) the need to rethink the socio-technical perspective on AI adoption for SMEs in Industry 5.0. In the intersection between SMEs and government policies, Industry 5.0 and Agenda 2030 have been identified.

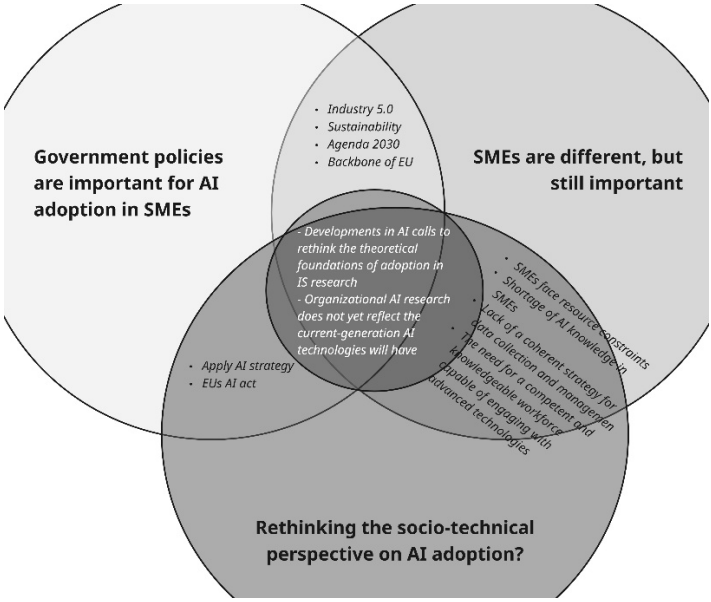

Figure 1: Summary of literature

3 Research approach

The research approach was inspired by interpretive field studies (Walsham 2006) and was conducted within the ongoing project, HUMAI. HUMAI is an acronym for "test and demonstration of HUMan-centered AI". The project involves five partners: Halmstad University, RISE (Research Institutes of Sweden), National Gummi (NG), HMS, and Connectitude. The manufacturing partner, NG, mostly produces rubber products primarily based on fossil raw materials, with a focus on rubber sealing strips used in automobiles and other applications. The goal of the project is to develop an AI demonstrator that supports process operators during production and supports production managers in their decision-making.

Figure 2 illustrates the layout of the production line, which spans approximately 60 meters from the extruder to the packing station. The goal of HUMAI is to develop a human-centered technical solution, grounded in artificial intelligence, to support and enhance production at NG. The four non-manufacturing partners collaborate on developing this solution.

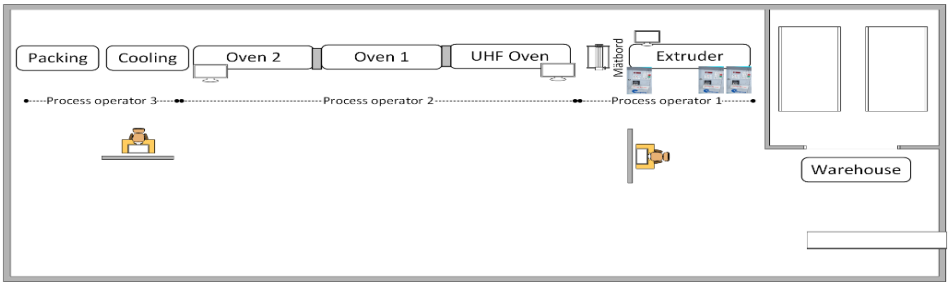

Figure 2: Production Line layout

3.1 Data Collection

In HUMAI, three different data collection techniques were used: field visits, stimulated recall interviews (SRI), and workshops. The research team has conducted five field visits to NG. After each field visit, the researchers have gained a broader understanding of the production. Each visit included meetings and walkthroughs of the production line, during which researchers documented the environment through photography (see Figure 3).

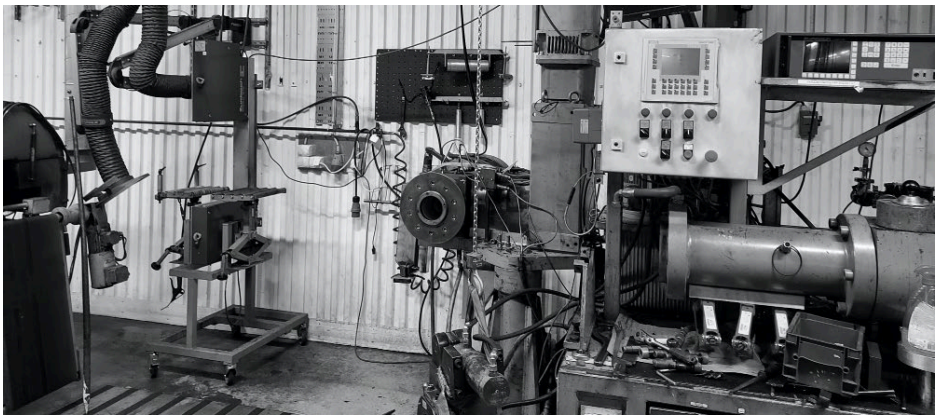

Figure 3: Photo of the extruder

Stimulated Recall Interviews (SRI) are a qualitative data collection method in which the interviewer uses visual, audio, or video stimuli to prompt participants to recall and reflect upon social interactions (Dempsey 2010). In HUMAI, SRI is employed to understand operators' actions and thought processes in situ. The method has been described as both a focused approach (Messmer 2015) and a technique for investigating how individuals navigate interactions. It is an introspective procedure that enables researchers to explore human reasoning and behavior by inviting participants to revisit moments captured via images (Lyle 2003). Although widely used in ethnography, pedagogy, and didactics, SRI remains relatively uncommon within information systems and engineering research. Approximately thirty photos were taken to capture the entire production process. In collaboration with NG's production managers, these images were sequenced to reflect the actual manufacturing flow. This sequence was later developed into a presentation, which also incorporated the production line map (Figure 1). A total of five SRI sessions were conducted with operators. Each interview lasted between 90 and 120 minutes. To create a comfortable interview environment, sessions began with an informal conversation about names, roles, family, and background. Four of the participants had more than three years of experience at NG, while the fifth had worked at the company for six months. The third data collection technique used in HUMAI is workshops. On four separate occasions, the researchers have had workshops with the production managers at NG. Each workshop had a specific task, such as describing the production process or managing data and parameters that affect the production.

3.2 Analysis

The analysis has also been inspired by Walsham (2006). The article recommends developing a coherent, interesting story for the reader. Which is not the easiest task! In HUMAI we have used two analysis techniques: thematic analysis (Braun and Clarke 2006) and process analysis (Langley 1999, Langley, Smallman et al. 2013). During the literature review, five themes were identified using the Braun and Clarke (2006) framework. In analyzing the empirical data, the researchers used a mix of thematic and process analysis to develop episodes grounded in everyday practice at NG. Two scenarios will be presented in the results section.

4 Results

Episode 1 – Searching for collaborators

Early on in HUMAI, the researchers visited NG and the production managers. At the meeting, there was a discussion about why and how the idea of adopting AI started at NG. According to one of the managers, it all started with the idea of digitalizing support processes for process operators during production. The background story at NG was that the management group was inspired by EU Industry 5.0 documents, especially the idea of being human-centric. The next step taken by NG was to contact organizations with which they had been collaborating before: TEK (https://www.tek.se/), HMS (https://www.hms-networks.com), and High5 (https://highfive.se/). After some time, HMS responded, one of the first things we need to do is to find ways to extract data from the extruder (Figure 2) and from the ovens (Figure 1). In any digital solution, data is the raw material. NG also contacted organizations knowledgeable about AI and the manufacturing industry: RI.SE (https://www.ri.se/en), Connectitude (https://connectitude.com/), and AI Sweden (https://www.ai.se/).

In parallel with the work done by HMS, the managers at NG started a project, project process management. In that project, the production process was described from the start of production through delivery to customers. The organization wanted to know more about how the operators worked, their competence, and to identify other sources of data that could affect the quality of the rubber. One major finding was that data from intermediate storage were not used to adjust the extrusion

parameters. Extrusion is one of the key stages in production, where the effects of material quality, mixing, and environmental conditions during intermediate storage (e.g., humidity, temperature, and storage duration) become apparent. In the worst cases, entire batches must be discarded due to quality issues. Throughout production, operators must make numerous individual adjustments, each of which can significantly impact the final product. Inefficient or incorrect parameter settings can lead to significant waste.

Episode 2 - Why did the production go wrong?

The manufacturing partner, NG, mostly produces rubber products primarily based on fossil raw materials, with a focus on rubber sealing strips used in automobiles and other applications. The rubber manufacturing process at NG is complex, as the quality and properties of the end products are influenced by numerous parameters that are difficult to control. One of the production managers said, When things go *wrong, we usually have no idea why the production went wrong.* He continued to describe various parameters that affect the production, including temperature, humidity, time spent in the warehouse, pressure, and belt speed. After a while, he also said, then we also have the human factors, the teams of operators. It is crucial to involve the process operators from the start, so that they know about the adoption process and that we are interested in their opinion.

A team of three operators runs the production line (Figure 1). They are responsible for managing a 60-meter stretch of the production process. To ensure efficiency and product quality, it is crucial that at least two of the three operators have more than two years of experience. Sometimes the team includes one or more operators contracted through staffing agencies. These individuals may have no prior experience at NG. As a result, permanent staff must continuously monitor and support the temporary workers. Moreover, staff absences (e.g., due to illness) are common and can lead to ad hoc team formations, exacerbating production challenges. One of the operators is responsible for setting the correct parameters on the extruder and the ovens. The parameter settings are described in the manufacturing specification ("tissen") for a specific order.

Based on the stimulated recall interviews with the operators, they did, to an extent, agree with the production manager's statement, "*We usually have no idea why the production went wrong*«. The operators mentioned one example: if the production goes wrong in the ovens, we usually fix the problem by adjusting the temperature or the belt speed. The operators also mentioned that fixing problems in the ovens is time-critical. Another common problem is the vulcanization of the rubber sealing strips. Vulcanization is measured at the end of the production line, meaning that if the rubber does not meet the vulcanization quality criteria, at least 60 meters of it must be discarded. This not only results in additional costs for the organization but also has environmental consequences. The underlying causes of problems with vulcanization are the settings of the extruder and oven parameters. The more experienced the operators are, the easier it is to adjust the parameters to achieve greater accuracy.

5 Deriving characteristics of AI adoption in manufacturing SMEs

Based on the episodes presented in the previous section, we have identified challenges faced by manufacturing SMEs that hinder AI adoption. In this section, we will discuss the challenges presented in relation to the tentative themes of characteristics outlined in the literature section: *government policies are important for AI adoption in manufacturing SMEs; SMEs are different but still important; and Rethinking the socio-technical perspective on AI adoption for SMEs in Industry 5.0.* The reported episodes demonstrate the complexity of the rubber production process, which is crucial to understand if the goal is to adopt an AI solution to support operators. It also demonstrates the search for collaboration to address challenges during the AI adoption process.

5.1 The AI adoption should support the sustainability work of manufacturing SMEs

In the described episodes, rubber waste is a problem for NG. It not only results in additional costs for the organization but also has environmental consequences. Every year, NG publishes a sustainability report in accordance with the United Nations Agenda 2030 (United Nations 2015) and Swedish regulations. Government policies are important for manufacturing SMEs in Europe. The sustainability awareness among operators and managers is high. In Industry 5.0, sustainability is

linked to planetary boundaries (European Commission 2022), as outlined in Agenda 2030. NG shares characteristics with other SMEs, such as flat and informal structures (Torrès and Julien 2005), and strategy is often emergent rather than deliberate and articulated in formal plans (Mintzberg and Waters 1985).

5.2 A Socio-technical and human-centric perspective should be applied throughout the AI adoption process

NG is a manufacturing SME industry, and as such, it relates to human-centricity in Industry 5.0 (European Commission 2022). In the vision of Industry 5.0, humans are positioned at the center of production. Furthermore, it also emphasizes the empowerment of individuals and highlights the importance of research to understand the complex interplay between humans and technology. NG wanted the process operators to be involved throughout the HUMAI project. The developments in AI have prompted growing calls to rethink the theoretical foundations of adoption in IS research (Akhlaghpour, Karahanna et al. 2025), and organizational AI research does not yet reflect organizational adoption and the impact that current-generation AI technologies will have (Nguyen, Sidorova et al. 2022). Parameters and parameter settings are frequently mentioned in the episodes. Parameters are examples of data that the organization must manage and that managers and operators use. NG had problems with data management; some data (humidity, temperature, and storage duration) were not digitized. SMEs are known to lack a coherent strategy for data collection and management (Khan, Kauttonen et al. 2025, Knoblach, Vogl et al. 2025)

5.3 Competence development and learning should be addressed throughout the AI adoption process

In the first episode, "Searching for collaborators," it is clear that NG cannot run the AI adoption process on its own. Generally speaking, SMEs face resource constraints in terms of competence (Barney 1991). At NG, there is a shortage of AI knowledge, a common issue among SMEs (Khan, Kauttonen et al. 2025; Knoblach, Vogl et al. 2025). Competence among process operators is important; sometimes the team includes one or more operators contracted through staffing agencies. These individuals may have no prior experience at NG. As a result, permanent staff must continuously monitor and support the temporary workers. In a sense, the more

experienced staff conduct informal learning with the rental staff to develop their competence. There is a need for a competent and knowledgeable workforce capable of engaging with advanced technologies (George and George 2023) for example, in SMEs.

6 Conclusion

As outlined in the introduction, the aim was to identify characteristics of AI adoption in manufacturing SMEs in Industry 5.0. Hence, the research question is: What are the characteristics of AI adoption in manufacturing SMEs in Industry 5.0?

In this study, we began exploring empirical findings from HUMAI to gain a deeper understanding of the characteristics of AI adoption in manufacturing SMEs in Industry 5.0. In conclusion, we have derived three characteristics of AI adoption in SMEs: 1) The AI adoption should support a manufacturing SMEs sustainability work, 2) A Socio-technical and human-centric perspective should be applied throughout the AI adoption process, and 3) Competence development and learning should be addressed throughout the AI adoption process

The article proposes three characteristics of AI adoption in manufacturing SMEs in Industry 5.0 for information systems researchers interested in AI adoption. This list is not exhaustive; instead, the characteristics have emerged through the HUMAI project. The proposed characteristics have not yet been tested or evaluated in any AI adoption project, underscoring the need to assess them in practice during adoption research to generate insights for the AI adoption researcher.

References

Akhlaghpour, S., et al. (2025). "Beyond IT Adoption: The Evolving Boundaries of Diffusion Research in the AI Era." *Communications of The Association for Information Systems* **57**(1): 42

Barney, J. (1991). "Firm resources and sustained competitive advantage." *Journal of Management* **17**(1): 99-120

Billett, S. (2001). "Learning through work: workplace affordances and individual engagement." *Journal of Workplace Learning* **13**(5): 209-214

Braun, V. and V. Clarke (2006). "Using thematic analysis in psychology." *Qualitative research in psychology* **3**(2): 77-101

Brundtland, G. H. (1987). "Our common future—Call for action." *Environmental conservation* **14**(4): 291-294

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of use, and User Acceptance of Information Technology." *MIS Quaterly*(september)

Dempsey, N. P. (2010). "Stimulated recall interviews in ethnography." *Qualitative sociology* **33**: 349-367

European Commission (2022). Industry 5.0: A Transformative Vision for Europe-ESIR Policy Brief No. 3. D.-G. f. R. a. Innovation. Publications Office of the European Union, European Commission

European Commission (2024). ERA - Industrial Technologies Roadmap on Human-Centric Research and Innovation for the manufacturing sector. D.-G. f. R. a. Innovation. Luxembourg, Publications Office of the European Union

European Commission (2025). Apply AI Strategy. E. A. Office, European Union

Ferreira, C. and S. Serpa (2018). "Society 5.0 and social development: Contributions to a discussion. Management and Organizational Studies, 5 (4), 26–31."

George, A. S. and A. H. George (2023). "Revolutionizing Manufacturing: Exploring the Promises and Challenges of Industry 5.0." *Partners Universal International Innovation Journal* **1**(2): 22-38

Grant, D. and E. Mergen (1996). "Applying quality to Leavitt's framework to solve information technology problems: A case study." *Information Technology & People* **9**(2): 43-60

Jennings, P. and G. Beaver (1997). "The performance and competitive advantage of small firms: a management perspective." *International small business journal* **15**(2): 63-75

Khan, U. A., et al. (2025). *AI Adoption in Finnish SMEs: Key Findings from AI Consultancy at a European Digital Innovation Hub*. 2025 IEEE 23rd World Symposium on Applied Machine Intelligence and Informatics (SAMI), IEEE

Knoblach, J., et al. (2025). "The SME AI Experience Center: Overcoming AI Adoption Barriers for SMEs in Industrial Production." *KI-Künstliche Intelligenz*: 1-6

Langley, A. (1999). "Strategies for theorizing from process data." *Academy of Management Review* **24**(4): 691-710

Langley, A., et al. (2013). "Process studies of change in organization and management: Unveiling temporality, activity, and flow." *Academy of Management Journal* **56**(1): 1-13

Leavitt, H. (1965). Applied organizational change in industry: Structural, technological,and humanistic approaches. *Handbook of Organizations*. (ed). NewYork, Rand McNally and Co. **5**

Lyle, J. (2003). "Stimulated recall: A report on its use in naturalistic research." *British educational research journal* **29**(6): 861-878

Messmer, R. (2015). *Stimulated recall as a focused approach to action and thought processes of teachers*. Forum: Qualitative Social Research, Freie Universität Berlin

Mintzberg, H. and J. A. Waters (1985). "Of strategies, deliberate and emergent." *Strategic Management Journal* **6**(3): 257-272

Nguyen, Q. N., et al. (2022). "Artificial intelligence in business: A literature review and research agenda." *Communications of The Association for Information Systems* **50**(1): 7

Tornatzky, L. G., et al. (1990). *The processes of technological innovation*. Lanham, MD, USA, Lexington Books

Torrès, O. and P.-A. Julien (2005). "Specificity and denaturing of small business." *International small business journal* **23**(4): 355-377

United Nations (2015). Agenda 2030 - Sustainable development, United Nations

Venkatesh, V., et al. (2003). "User acceptance of information technology: Toward a unified view." *MIS Quarterly*: 425-478

Walsham, G. (2006). "Doing interpretive research." *European Journal of Information Systems* **15**(3): 320-330

Wärnestål, P. (2022). *Designing AI-Powered Services*. Lund, Sweden, Studentlitteratur AB

Yu, X., et al. (2023). "Antecedents and outcomes of artificial intelligence adoption and application in the workplace: the socio-technical system theory perspective." *Information Technology & People* **36**(1): 454-474

About the author

My main research interest is digital service innovation from a learning perspective. Research has identified collaboration between heterogeneous actors with different knowledge and backgrounds as problematic in digital service innovation. There is a consensus in the literature that we must find ways to relate, share, translate, and exchange knowledge in order to succeed with innovation activities. In other words, it is important to understand and describe the learning processes that take place during and between innovation activities. Teaching takes place in various courses within Informatics. In recent years, I have been involved in the research program LeaDS (Learning in a digitalized Society). From autumn 2024, my teaching will be in the Digital Business Development and Digital Learning programs.

I am also appointed an excellent teacher in Information Systems.

Organisational Readiness for Open Government Data-Driven Co-Creation

Petra Vujković,[1, 2] Tina Jukić,[1] Sanja Vrbek[1]

[1] University of Ljubljana, Faculty of Public Administration, Ljubljana, Slovenia
petra.vujkovic@fu.uni-lj.si, tina.jukic@fu.un-lj.si, sanja.vrbek@fu.uni-lj.si
[2] University of Rijeka, Faculty of Economics & Business, Rijeka, Croatia
petra.vujkovic@fu.uni-lj.si

Open government data (OGD) is widely promoted for transparency and innovation, yet reuse remains uneven. Co-creation offers a route to public value by enabling public organisations and external stakeholders to jointly design or improve services and policies using open data. This paper synthesises what the OGD-driven co-creation literature says about internal organisational conditions that support effective participation. We conducted a systematic literature review of English-language studies indexed in the Web of Science database. The analysis focuses on organisational factors, understood as both drivers and barriers, within public organisations rather than ecosystem-level conditions. The findings indicate that readiness depends on more than data release: it also requires strategic commitment to openness, clear procedures and incentives, a learning-oriented culture, leadership that enables cross-boundary collaboration, staff competences in data and participatory methods, and the ability to assure data quality. We discuss implications for strengthening organisational capacities for sustained, data-enabled co-creation and identify priorities for further empirical research.

DOI
https://doi.org/10.18690/um.fov.4.2026.18

ISBN
978-961-299-152-4

Keywords:
barriers,
co-creation,
drivers,
open government data,
public organisations

University of Maribor Press

1 Introduction

Over the past decade, open government data (OGD) has become an important policy instrument within broader open government and digital governance reforms. Governments increasingly publish administrative datasets to improve transparency and accountability and to enable the reuse of public sector data for innovation and new services (Janssen et al., 2012; Janssen & van den Hoven, 2015). In the literature, these developments are commonly discussed as part of broader debates on open government, public governance, and the role of data disclosure in supporting citizen participation (Janssen et al., 2012; Ruijer et al., 2018).

Empirical research shows, however, that increasing data availability does not automatically lead to effective reuse or policy impact. Janssen et al. (2012) identify technical, organisational, and institutional barriers that constrain data utilisation. Ruijer et al. (2018) demonstrate that transparency initiatives do not necessarily result in democratic engagement. Together, these studies indicate that the outcomes of OGD initiatives depend on the conditions under which data are accessed and used in practice (Janssen et al., 2012; Ruijer et al., 2018). As open data does not create value simply by being accessible, its contribution depends on how it is interpreted, combined, and applied in decision-making or service development processes (Charalabidis et al., 2018; Hitz-Gamper et al., 2019).

One such mechanism is co-creation. In the public sector context, co-creation refers to collaborative processes that involve multiple stakeholders, including public authorities, citizens, and other actors, in the development and improvement of public services (Vrbek & Jukić, 2024). By enabling citizens, civil society actors, and private organisations to collaborate with public authorities, OGD can serve as a foundational resource for jointly developing public services and solutions (McBride et al., 2019). In this context, OGD-driven co-creation refers to collaborative processes in which publicly available datasets are utilised by diverse actors to design, implement, or improve services that generate public value (Callinan et al., 2018). Such initiatives typically involve multiple stakeholders and rely on open data resources as a key input (McBride et al., 2019).

Despite the growing body of research on both OGD and co-creation, these strands of literature have largely developed separately (Callinan et al., 2018). While OGD research has often focused on the provision of datasets, governance arrangements, and technological infrastructures, co-creation scholarship has primarily examined participatory processes and stakeholder interaction. The organisational conditions required to integrate these domains in practice remain underexplored, and in particular, limited attention has been devoted to understanding how public organisations can develop the internal organisational conditions that influence their readiness (i.e., maturity) to support and sustain OGD-driven co-creation initiatives.

Accordingly, this paper seeks to identify and systematise factors (understood as drivers and barriers) influencing the readiness of public organisations to engage in OGD-driven co-creation. The study addresses the following research question: *What makes public organisations ready to co-create successfully based on OGD?*

The next section presents the methodological approach, followed by a structured analysis of the factors affecting organisational readiness for OGD-driven co-creation. The paper concludes with a discussion of the findings, their implications, and the study's limitations.

2 Methodology

This study is based on a systematic literature review of OGD-driven co-creation, aimed at identifying key organisational factors (i.e., drivers and barriers) influencing the organisational readiness of public organisations to engage in OGD-driven co-creation.

The search was conducted in the Web of Science (WoS) database using the query: (All fields = "open data" OR "open government data") AND (All fields = co-creation) AND LANGUAGE: English. The search yielded 51 publications. A screening of titles and abstracts excluded four articles that were either not related to OGD-driven co-creation or focused primarily on the private sector. A subsequent full-text assessment led to the exclusion of two further articles. The final sample consisted of 45 studies, supplemented by one further article considered relevant but not retrieved through the initial search.

The selected publications were systematically coded to identify factors (i.e., drivers and barriers), which were classified according to whether they relate to: (i) features of public organisations, (ii) external co-creators, (iii) the co-creation process, or (iv) the broader environment. For the purposes of this article, the analysis concentrates primarily on factors from the first group shaping the organisational readiness of public organisations to engage in OGD-driven co-creation, understood here as their ability to co-create successfully based on OGD. Selected contextual factors were also considered where relevant for understanding organisational readiness.

3 Factors affecting the organisational readiness of public organisations to engage in OGD-driven co-creation

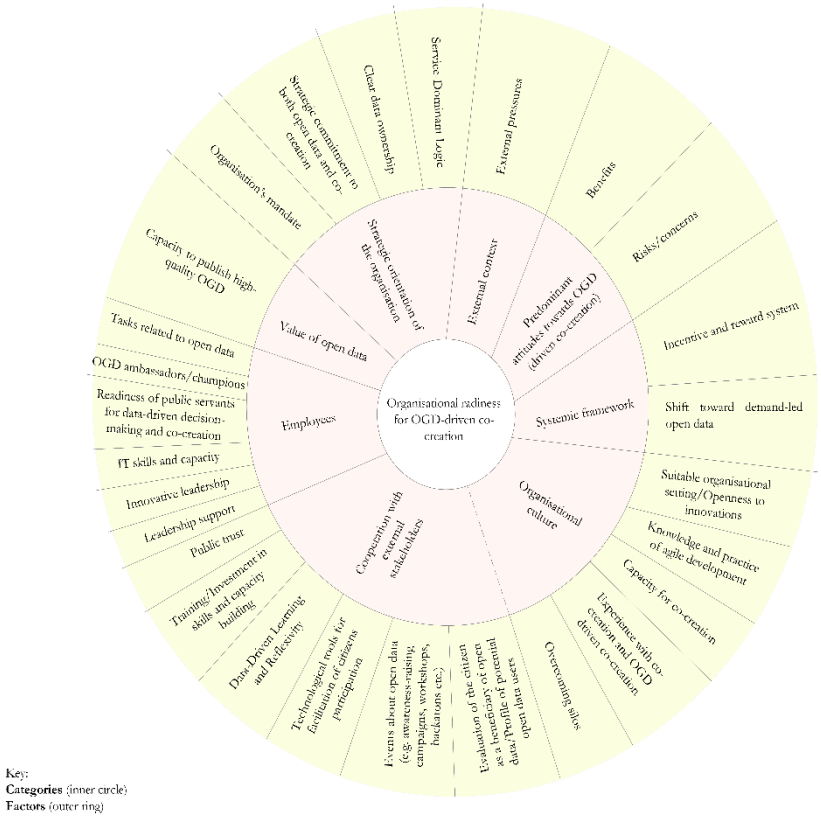

Figure 1: Conceptual overview of categories and factors shaping the organisational readiness for OGD-driven co-creation.

Source: Own

Building on a systematic review of the literature on OGD-driven co-creation, this section synthesises the key factors shaping the organisational readiness of public organisations to engage in OGD-driven co-creation. These factors are organised into analytically distinct categories of organisational readiness (see Figure 1, which provides a conceptual overview of the identified categories and associated factors). The categories were developed inductively through the reviewed studies. Each category and its associated factors are discussed in the subsections that follow.

3.1 Strategic orientation of the organisation

Within this category, four key factors emerge from the literature (see Figure 1 for an overview). A clear strategic commitment to both OGD and co-creation is identified as a key enabling factor for OGD-driven innovation and the creation of public value (Scheibein et al., 2022; Wiedenhöft et al., 2023). However, this commitment cannot remain limited to formal declarations and strategic documents. It entails reconsidering how public organisations balance efficiency concerns with democratic engagement, with greater emphasis placed on the latter (Cubuk et al., 2019). Moreover, the institutionalisation of open data depends on binding legal and political mandates. Public bodies operating under explicit and enforceable requirements to publish and reuse data tend to implement OGD in a more systematic and consistent manner, including in co-creation contexts (McBride et al., 2020).

Data ownership represents an additional critical issue, particularly in large-scale collaborative initiatives that result in the joint generation of data (Toli et al., 2022). For this reason, establishing transparent ownership and licensing arrangements at the beginning of co-creation processes is essential. Such arrangements support responsible data management, facilitate open reuse, and contribute to the continuity of data-driven public service innovation (Toli et al., 2022). In these situations, questions of ownership go beyond internal uncertainties within public organisations, such as determining which entity retains rights over a specific dataset.

An additional factor shaping organisational readiness for OGD-driven co-creation is the adoption of a service-dominant logic (SDL). In contrast to state-centric understandings of public services as outputs delivered unilaterally by government institutions, SDL conceptualises service provision as a participatory and interaction-based process. Within this perspective, citizens and other non-state stakeholders are

regarded as active co-creators of public value rather than passive recipients (McBride et al., 2018; Potra et al., 2015). Adopting this logic requires a corresponding shift in administrative mindsets, recognising citizens as legitimate partners and prosumers within service ecosystems (Nikiforova & McBride, 2021; Toots et al., 2017a).

3.2 External context

External pressures, particularly in the form of citizen demand, have been identified as a primary motivating factor encouraging public organisations to increase data availability and pursue more participatory and collaborative governance approaches (Toots et al., 2017a). More broadly, drivers for opening government data and engaging in co-creation processes tend to stem from external forces (such as citizen demand, technological developments, and policy pressures) rather than from purely internal administrative initiatives (McBride et al., 2019; Toots et al., 2017a).

3.3 Predominant attitudes towards OGD-driven co-creation

Within this category, two key factors emerge from the literature (see Figure 1 for an overview). The first factor relates to the degree of awareness of the potential benefits of OGD (McBride et al., 2019; Toots et al., 2017; Wiedenhöft et al., 2023). When these benefits (particularly those extending beyond economic value) are not clearly understood or sufficiently acknowledged, the incentive to participate in OGD initiatives remains weak (Scarano et al., 2019). In the specific setting of OGD-driven co-creation, Khayyat and Bannister (2017) point to the perception among certain public organisations that such initiatives are warranted primarily when the main advantages are realised by governmental actors. For this reason, enhancing the recognition of the wider societal value that may emerge from OGD-driven co-creation becomes important.

The second factor concerns perceived risks linked to OGD, especially those associated with confidentiality and privacy. The possibility of re-identification through the combination of datasets can limit decisions to make data openly available, reusable, or linkable (Dall-Orsoletta et al., 2022; Scheibein et al., 2022; Toots et al., 2017). Furthermore, compliance with data protection frameworks, including the GDPR, may generate administrative burdens and constrain flexibility in the design of co-creation processes (Kavouras et al., 2023). Additionally,

unwillingness to share data - particularly among certain stakeholders (e.g. private sector actors) due to concerns related to confidentiality, competitiveness, and misuse - may hinder open collaboration and limit the innovative capacity of OGD initiatives (Taylor, 2014; Toli et al., 2022).

3.4 Systemic framework

Within this category, several systemic factors emerge from the literature (see Figure 1 for an overview). A structured system of incentives plays a central role in motivating public organisations to prioritise open data activities and engage in co-creation. Where incentive structures are weak or absent, open data initiatives risk being sidelined within routine administrative practices (McBride et al., 2020). Equally important is the availability of organisational resources. Research highlights that human competencies, financial capacity, technical expertise, time allocation, and IT infrastructure substantially shape the capacity of public organisations to support Open Government Data Value Co-Creation (OGD VCC) (Wiedenhöft et al., 2023). Adequate resourcing strengthens the management of data infrastructures and facilitates stakeholder involvement in co-creative service development.

The existence of clearly defined internal procedures for data publication represents an additional systemic factor. Transparent and consistently applied organisational processes clarify responsibilities, regulate data-related rights, and uphold data quality standards. In contrast, unclear procedures for data preparation and dissemination may discourage both internal data providers and external users, thereby limiting the effectiveness of open data initiatives (McBride et al., 2019; Wiedenhöft et al., 2023). Institutional rigidity represents an additional structural constraint affecting the systemic environment. Bureaucratic structures that restrict flexibility may hinder cross-departmental coordination and cross-sector collaboration, which are necessary for sustainable open data governance (Toots et al., 2017a).

A further systemic factor relates to predictable publication practices. Regular and transparent data release schedules enhance usability and enable the development of data-driven services (Taylor, 2014). Finally, literature emphasises the importance of shifting from supply-oriented data publication towards demand-responsive approaches. Rather than prioritising internal administrative convenience, strategic data release focuses on datasets that are most relevant to stakeholders (Köster &

Suarez, 2016). This transition strengthens usability and public value creation through open data (Taylor, 2014). Operationally, such a shift requires feedback mechanisms that allow administrations to identify user needs and adapt data provision accordingly (Khayyat & Bannister, 2017; Toots et al., 2017a).

3.5 Organisational culture

Within this category, five key factors related to organisational culture emerge from the literature (see Figure 1 for an overview). Organisational openness to new ideas and practices has been identified as a key enabler of value creation based on OGD (Mergel et al., 2018; Toots et al., 2017a). However, innovation in the public sector frequently encounters barriers, particularly institutional resistance to change and rigid administrative structures (McBride et al., 2018; Wiedenhöft et al., 2023). Such constraints undermine organisational adaptability and weaken the development of mechanisms necessary for value co-creation, including cross-sector collaboration and structured stakeholder feedback (Wiedenhöft et al., 2023). Political resistance, risk-averse leadership, and conservative governance practices may further limit openness and experimentation (Seo & Myeong, 2021; Toots et al., 2017a). In addition, regulatory rigidity and systemic inertia characteristic of traditional bureaucratic settings can limit the adoption and scaling of OGD-related strategies (Seo & Myeong, 2021). Within many public organisations, fear of failure discourages engagement with innovative approaches and collaboration with external actors (Seo & Myeong, 2021), partly linked to efforts to preserve institutional authority (Potra et al., 2015). As a result, OGD initiatives may be regarded as marginal or even threatening (McBride et al., 2020). Addressing these challenges is associated with shifts towards more open and collaborative governance arrangements (Callinan et al., 2018).

Beyond these constraints, organisational culture also manifests through capabilities and practices that actively enable co-creation. In this respect, organisational capabilities related to agile development and co-creation practices play a critical role in enabling OGD-driven co-creation. Agile approaches support iterative learning, early testing, and adaptive development processes by replacing linear project models with more flexible cycles and instruments such as Minimum Viable Products (MVPs) (McBride et al., 2019; Toots et al., 2017a). Capacity for co-creation is equally critical, involving the application of participatory methods such as data walks and policy

canvases to engage diverse stakeholders and design socially responsive services (Cubuk et al., 2019; Jarke, 2019). The use of open-access tools and international knowledge exchange further supports inclusive innovation processes (Atenas et al., 2019; Khayyat & Bannister, 2017). Experience with co-creation, particularly in OGD-driven initiatives, enhances organisational readiness and strengthens stakeholder relationships by fostering trust and demonstrating concrete outcomes (McBride et al., 2020; Mergel et al., 2018). Nevertheless, issues related to scalability and the evaluation of social impact remain challenging (Jussila et al., 2019; Scarano et al., 2019). Finally, overcoming internal and cross-organisational silos remains essential, as fragmented structures hinder collaboration and data integration, thereby limiting coordinated and innovative public service development (Jussila et al., 2019; Potra et al., 2015).

3.5.1 Cooperation with external stakeholders

Within this sub-category, six interrelated factors emerge from the literature (see Figure 1 for an overview). Several factors relate to organisations' engagement with external stakeholders, including efforts to assess user profiles, organise open data-related events, promote data-driven learning through feedback, and investing in the development of skills and capacities.

In this context, digital technologies play a central role in facilitating stakeholder participation. The effective deployment of digital technologies supports meaningful engagement in OGD-driven co-creation processes. Digital tools such as social media, collaborative platforms, and open-source applications provide opportunities for stakeholders to contribute to the design, testing, and implementation of public services (Cubuk et al., 2019; Emaldi et al., 2017; Gebetsroither-Geringer et al., 2018; Huang et al., 2021). These platforms also enable participatory innovation by offering access to data repositories, structured toolkits, and feedback mechanisms (Jussila et al., 2019; Morelli et al., 2017). More advanced technologies, including 3D visualisations, may further enhance engagement, although their effectiveness depends on usability, accessibility, and available infrastructure (Atenas et al., 2019; Dambruch, 2018; Gebetsroither-Geringer et al., 2018).

Public trust represents an additional critical factor influencing stakeholder participation in OGD-driven co-creation. The credibility and reputation of public organisations influence citizens' willingness to participate in co-creation activities. Higher levels of trust support sustained engagement and contribute to the perceived legitimacy of collaborative initiatives (Wiedenhöft et al., 2023).

3.6 Employees

Within this category, a set of key employee-related factors emerge from the literature (see Figure 1 for an overview). These factors encompass the formal incorporation of open data-related tasks into employees' regular job responsibilities; the presence of open data champions who actively promote and advocate the OGD agenda; employees' IT competencies and their capacity to manage, analyse, and interpret data; and leadership support expressed through priority-setting, allocation of resources, and sustained organisational commitment to advancing open data practices.

An additional factor concerns the presence of innovative leadership. Leaders who demonstrate openness to experimentation and new approaches, particularly in uncertain contexts, play an important role in fostering organisational learning, adaptability, and co-creation practices. Such leadership contributes to embedding open data reforms and aligning co-creation initiatives with broader strategic objectives (McBride et al., 2018b; McBride et al., 2019; Mergel et al., 2018).

A further factor relates to the readiness of public servants to engage in data-driven decision-making and co-creation processes. Beyond fulfilling open data obligations, employees need to recognise citizens as legitimate partners in the development of public services and policies. This requires an attitudinal shift towards data-informed approaches and acceptance of co-creation as a valid governance practice (McBride et al., 2018b).

3.7 Value of Open data

A fundamental driver of OGD VCC lies in the ability of public organisations to ensure the availability of high-quality and sustainable datasets (Wiedenhöft et al., 2023). Effective value co-creation requires a comprehensive understanding of the

data lifecycle, encompassing stages from data collection and processing to publication and subsequent reuse (Wiedenhöft et al., 2023). This capability involves not only technical proficiency but also familiarity with the data value chain, including activities such as data collation, cleaning, transformation, and standardisation (McBride et al., 2018). Such competencies contribute to improving data accuracy, interoperability, and reusability, thereby enhancing their applicability in co-creation contexts. Crucially, the capacity to integrate heterogeneous datasets and connect them to societal or user needs constitutes a central mechanism through which meaningful public value can be generated (Wiedenhöft et al., 2023).

4 Discussion and conclusion

The findings of the systematic literature review indicate that organisational readiness for OGD-driven co-creation cannot be reduced to technical data publication capacity alone. Although the ability to generate, curate, and maintain high-quality datasets constitutes an important precondition, it does not in itself ensure meaningful collaborative engagement. This observation aligns with broader research on co-creation and organisational maturity, which emphasises that structural, organisational, and behavioural factors shape the capacity of public organisations to engage in collaborative processes (Baptista et al., 2019; Jukić et al., 2022).

OGD-driven co-creation requires broader organisational adaptation, including adjustments to strategic priorities, internal procedures, and institutional arrangements. Rather than prioritising internal control and procedural compliance, public organisations need to adopt more externally oriented and flexible approaches to governance. In this respect, moving from supply-driven data release towards demand-responsive practices becomes particularly relevant, as co-creation depends on the alignment of datasets with societal needs and stakeholder expectations (Voorberg et al., 2015).

Such transformation also entails a reconsideration of how public value is understood and generated. Drawing on public service logic, value emerges through interaction and resource integration among multiple actors rather than through unilateral service delivery (Rösler et al., 2021; Jukić et al., 2022). OGD-driven co-creation is therefore characterised by iterative learning, experimentation, and adaptive development processes. These dynamics may challenge traditional conceptions of efficiency and

performance that favour predictability and short-term accountability. Consequently, organisational readiness encompasses not only structural reforms but also deeper cultural and cognitive shifts within public administration.

The review further underscores the significance of leadership, organisational culture, and employee capacities. Research consistently demonstrates that leadership plays a central role in enabling collaborative innovation by legitimising experimentation, mobilising resources, and fostering organisational learning (Baptista et al., 2019; Rösler et al., 2021). At the same time, public servants must develop both the competences and willingness to integrate data and external input into decision-making processes, recognising citizens and other stakeholders as legitimate partners in public value creation (Voorberg et al., 2015). Skills related to data management, participatory methods, digital tools, and cross-sector collaboration therefore represent core components of organisational readiness. Moreover, prior experience with co-creation and the strategic institutionalisation of collaborative practices strengthen the conditions for sustained engagement (Jukić et al., 2022).

This study contributes to the literature by integrating insights from OGD and co-creation research through the lens of organisational readiness. While previous studies have examined these domains separately, this review provides a structured synthesis of organisational factors that jointly shape readiness for OGD-driven co-creation. Nevertheless, several limitations must be acknowledged. The review relied exclusively on publications indexed in a single database. Furthermore, the identified factors are derived from literature spanning diverse administrative and national contexts, implying that their practical relevance may vary across institutional environments. Future research should therefore examine and validate these factors in specific empirical settings in order to assess their contextual applicability and relative importance.

Acknowledgment

The research was funded by the Slovenian Agency for Research and Innovation (ARIS) under grant n. J5-50166.

References

Atenas, J., Havemann, L., Nascimbeni, F., Villar-Onrubia, D., & Orlic, D. (2019). Fostering Openness in Education: Considerations for Sustainable Policy-Making. *Open Praxis*, 11(2), 167–183. https://doi.org/10.5944/openpraxis.11.2.947

Baptista, N., Alves, H., & Matos, N. (2019). Public Sector Organizations and Cocreation With Citizens: A Literature Review on Benefits, Drivers, and Barriers. *Journal of Nonprofit & Public Sector Marketing*, 32(3), 217–241. https://doi.org/10.1080/10495142.2019.1589623

Callinan, C., Scott, M., Ojo, A., & Whelan, E. (2018). How to Create Public Value Through Open Data Driven Co-Creation: A Survey of the Literature. *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*, 363–370. https://doi.org/10.1145/3209415.3209457

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., & Ferro, E. (2018). The Multiple Life Cycles of Open Data Creation and Use. In Y. Charalabidis, A. Zuiderwijk, C. Alexopoulos, M. Janssen, T. Lampoltshammer, & E. Ferro, *The World of Open Data* (Vol. 28, pp. 11–31). Springer International Publishing. https://doi.org/10.1007/978-3-319-90850-2_2

Cubuk, E. B. S., Karkin, N., & Yavuz, N. (2019). Public sector innovativeness and public values through information and communication technologies. 353–361. https://doi.org/10.1145/3325112.3325215

Dall-Orsoletta, A., Romero, F., & Ferreira, P. (2022). Open and collaborative innovation for the energy transition: An exploratory study. *Technology in Society*, 69, 101955. https://doi.org/10.1016/j.techsoc.2022.101955

Dambruch, J. (2018). COMPUTING FEEDBACK FOR CITIZENS' PROPOSALS IN PARTICIPATIVE URBAN PLANNING. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-4/W7, 35–42. https://doi.org/10.5194/isprs-annals-IV-4-W7-35-2018

Emaldi, M., Aguilera, U., López-de-Ipiña, D., & Pérez-Velasco, J. (2017). Towards Citizen Co-Created Public Service Apps. *Sensors*, 17(6), 33. https://doi.org/10.3390/s17061265

Gebetsroither-Geringer, E., Stollnberger, R., & Peters-Anders, J. (2018). INTERACTIVE SPATIAL WEB-APPLICATIONS AS NEW MEANS OF SUPPORT FOR URBAN DECISION-MAKING PROCESSES. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-4/W7, 59–66. https://doi.org/10.5194/isprs-annals-IV-4-W7-59-2018

Hitz-Gamper, B. S., Neumann, O., & Stürmer, M. (2019). Balancing control, usability and visibility of linked open government data to create public value. *International Journal of Public Sector Management*, 32(5), 457–472. https://doi.org/10.1108/ijpsm-02-2018-0062

Huang, H.-Y., Kovacs, M., Kryssanov, V., & Serdult, U. (2021). Towards a Model of Online Petition Signing Dynamics on the Join Platform in Taiwan. *2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG)*, 199–204. https://doi.org/10.1109/ICEDEG52154.2021.9530852

Janssen, M., & van den Hoven, J. (2015). Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy? *Government Information Quarterly*, 32(4), 363-368. https://doi.org/10.1016/j.giq.2015.11.007.

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012) Benefits, Adoption Barriers and Myths of Open Data and Open Government, *Information Systems Management*, 29(4), 258-268. https://doi.org/10.1080/10580530.2012.716740

Jarke, J. (2019). Open government for all? Co-creating digital public services for older adults through data walks. *Online Information Review*, 43(6), 1003–1020. https://doi.org/10.1108/OIR-02-2018-0059

Jukić, T., Pluchinotta, I., Hržica, R., & Vrbek, S. (2022). Organizational maturity for co-creation: Towards a multi-attribute decision support model for public organizations. *Government Information Quarterly*, 39(1), 101623. https://doi.org/10.1016/j.giq.2021.101623

Jussila, J., Kukkamäki, J., Mäntyneva, M., & Heinisuo, J. (2019). Open Data and Open Source Enabling Smart City Development: A Case Study in Hame Region. *Technology Innovation Management Review*, 9(9), 25–34. https://doi.org/10.22215/timreview/1266

Kavouras, I., Sardis, E., Protopapadakis, E., Rallis, I., Doulamis, A., & Doulamis, N. (2023). A Low-Cost Gamified Urban Planning Methodology Enhanced with Co-Creation and Participatory Approaches. *Sustainability*, 15(3), 2297. https://doi.org/10.3390/su15032297

Khayyat, M., & Bannister, F. (2017). Towards a model for facilitating and enabling co-creation using open government data. *Information Polity*, 22(4), 211–231. https://doi.org/10.3233/IP-170406

Köster, V., & Suarez, G. (2016). Open Data for Development: Experience of Uruguay. 207–210. https://doi.org/10.1145/2910019.2910060

McBride, K., Olesk, M., Kütt, A., & Shysh, D. (2020). Systemic change, open data ecosystem performance improvements, and empirical insights from Estonia: A country-level action research study. *Information Polity*, 25(3), 377–402. https://doi.org/10.3233/IP-190195

McBride, K., Toots, M., Kalvet, T., & Krimmer, R. (2019). Turning Open Government Data into Public Value: Testing the COPS Framework for the Co-creation of OGD-Driven Public Services. In M. P. Rodríguez Bolívar, K. J. Bwalya, & C. G. Reddick (Eds.), *Governance Models for Creating Public Value in Open Data Initiatives* (Vol. 31, pp. 3–31). Springer International Publishing. https://doi.org/10.1007/978-3-030-14446-3_1

McBride, K., Aavik, G., Kalvet, T., & Krimmer, R. (2018b). Co-creating an Open Government Data Driven Public Service: The Case of Chicago's Food Inspection Forecasting Model. *Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2018.309

McBride, K., Toots, M., Kalvet, T., & Krimmer, R. (2018). Open Government Data Driven Co-creation: Moving Towards Citizen-Government Collaboration. In P. Parycek, O. Glassey, M. Janssen, H. J. Scholl, E. Tambouris, E. Kalampokis, & S. Virkar (Eds.), *Electronic Government* (Vol. 11020, pp. 184–195). Springer International Publishing. https://doi.org/10.1007/978-3-319-98690-6_16

Mergel, I., Kattel, R., Lember, V., McBride, K., & Acm. (2018). Citizen-Oriented Digital Transformation in the Public Sector. 844–846. https://doi.org/10.1145/3209281.3209294

Morelli, N., Aguilar, M., Concilio, G., De Götzen, A., Mulder, I., Pedersen, J., & Torntoft, L. K. (2017). Framing Design to support Social Innovation: The Open4Citizens Project. *Design Journal*, 20, S3171–S3184. https://doi.org/10.1080/14606925.2017.1352823

Nikiforova, A., & McBride, K. (2021). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, 58, 101539. https://doi.org/10.1016/j.tele.2020.101539

Potra, S., Branea, A. M., & Izvercian, M. (2015). How to Foster Prosumption for Value Co-Creation? *The Open Government Development Plan*. 239–245.

Rösler, J., Söll, T., Hancock, L., & Friedli, T. (2021). Value Co-Creation between Public Service Organizations and the Private Sector: An Organizational Capabilities Perspective. *Administrative Sciences*, 11(2), 55. https://doi.org/10.3390/admsci11020055

Ruijer, E., Grimmelikhuijsen, S., van den Berg, J., & Meijer, A. (2018). Open data work: understanding open data usage from a practice lens. *International Review of Administrative Sciences*, 86(1), 3-19. https://doi.org/10.1177/0020852317753068

Scarano, V., Andreoli, R., Monaco, D., Negro, A., Santangelo, G., & Vicidomini, L. (2019). Re-using Open Data by Automatically Building Immersive Virtual Reality Worlds as Personal Museums. *Proceedings of the 20th Annual International Conference on Digital Government Research*, 297–305. https://doi.org/10.1145/3325112.3325258

Scheibein, F., Donnelly, W., & Wells, J. S. (2022). Assessing open science and citizen science in addictions and substance use research: A scoping review. *International Journal of Drug Policy*, 100, 12. https://doi.org/10.1016/j.drugpo.2021.103505

Seo, H., & Myeong, S. (2021). Determinant Factors for Adoption of Government as a Platform in South Korea: Mediating Effects on the Perception of Intelligent Information Technology. *Sustainability*, 13(18), 20. https://doi.org/10.3390/su131810464

Taylor, J. (2014). Citizen Enablement from Open Data to Open Policy: A Personal View [Article; Book Chapter]

Toli, A. M., Murtagh, N., & Smyth, H. (2022). Co-owned resources: IP and data in smart cities. *Journal of Service Theory and Practice*, 32(2), 156–178. https://doi.org/10.1108/jstp-11-2019-0243

Toots, M., McBride, K., Kalvet, T., Krimmer, R., Tambouris, E., Panopoulou, E., Kalampokis, E., & Tarabanis, K. (2017). A Framework for Data-Driven Public Service Co-production. 10428, 264–275. https://doi.org/10.1007/978-3-319-64677-0_22

Toots, M., McBride, K., Kalvet, T., & Krimmer, R. (2017a). Open Data as Enabler of Public Service Co-creation: Exploring the Drivers and Barriers. *2017 Conference for E-Democracy and Open Government (CeDEM)*, 102–112. https://doi.org/10.1109/CeDEM.2017.12

Voorberg, W. H., Bekkers, V. J. J. M., & Tummers, L. G. (2015). A Systematic Review of Co-Creation and Co-Production: Embarking on the social innovation journey. *Public Management Review*, 17(9), 1333–1357. https://doi.org/10.1080/14719037.2014.930505

Vrbek, S & Jukić, T. (2024). Co-creation service readiness model: a decision support for the selection of public services suitable for improvement through co-creation. *Transforming Government: People, Process and Policy*, 18(1), 13-32. https://doi.org/10.1108/TG-03-2023-0031

Wiedenhöft, G. C., Matheus, R., Saxena, S., & Alexopoulos, C. (2023). Barriers towards Open Government Data Value Co-Creation: An empirical investigation. *THE ELECTRONIC JOURNAL OF INFORMATION SYSTEMS IN DEVELOPING COUNTRIES*, 89(5), e12270. https://doi.org/10.1002/isd2.12270

HYPERTENSION EHEALTH INTERVENTION NATIONAL UPSCALING

LUUK P.A. SIMONS

Delft University of Technology, Faculty of Computer Science, Delft, the Netherlands
l.p.a.simons@tudelft.nl

Hypertension is a main risk factor worldwide for premature death. After having conducted several hypertension eHealth interventions locally, we now study national upscaling. Innovation diffusion is generally challenging; there are additional challenges in eHealth and even more so for prevention interventions. Still, the Dutch GLI-system (Combined Lifestyle Intervention) appears promising for enabling national innovation diffusion. The local 'Blood Pressure Challenge' intervention has been redesigned to a 'Blood Pressure GLI'-format. This new service concept is evaluated using Rogers' framework of innovation diffusion in a user study with n=7 potential adoption partners (physiotherapists, dietitians and lifestyle coaches) having GLI execution experience. Main findings are that the Dutch GLI-system appears supportive for innovation diffusion and that the respondents see clear benefits in the Blood Pressure GLI service concept and its tangible blood pressure improvements. Finally, the user feedback suggested several service redesign improvements.

DOI
https://doi.org/10.18690/um.fov.4.2026.19

ISBN
978-961-299-152-4

Keywords:
hypertension,
ehealth,
innovation adoption,
lifestyle,
intervention
user test
upscaling

University of Maribor Press

1 Introduction to the research question

Hypertension is the main risk factor for death worldwide, according to Lancet publications from the largest health study ever (Lozano, 2012, Lim, 2012). Still, most people lack the health self-management competences to improve their blood pressure. Which is why most people get hypertension before they retire (Ostchega, 2020, Zhou, 2021, Carey & Whelton, 2020). Hence, we previously developed and tested a 'Blood Pressure Challenge,' an eHealth intervention using various tools and daily microlearning strategies to strongly reduce blood pressure (avg. -20/-10 systolic/diastolic) within several weeks (Simons et al., 2022, 2024, 2025, 2026).

After more than 10 employer/employee groups and four community pilots in Leiden and Rotterdam, the next question came up: how can we upscale to a more national and structural level, as a follow up to this series of one-off local innovation projects? Also, several local health care partners (physiotherapists, dietitians and lifestyle coaches) started asking for options for their adoption, training and reimbursement for the Blood Pressure Challenge. Next, the idea came up to transform the challenge into a GLI format. The GLI (Combined Lifestyle Intervention) is a nationally regulated, evidence-based healthcare pathway in the Netherlands that is reimbursed through basic health insurance, incorporating 2-year-long lifestyle interventions for obesity and lifestyle-related diseases. There is a review board and procedure to assess if a lifestyle intervention meets requirements and becomes part of the GLI pathway. When an intervention is 'GLI-accredited' local practitioners can be trained and licensed, if they follow the standard and report their results in a national database.

Since the GLI admission procedure requires quite some investments in time, money and data collection, an evaluation study was conducted with potential adoption partners having GLI experience, on the Blood Pressure GLI design concept.

Our ***Research Question:***

To which extent does the Dutch GLI-format help in national upscaling and what is the feasibility and attractiveness of the Blood Pressure GLI for potential adoption partners?

2 Theory and concepts

In this section we take two steps. First, we briefly review some of the challenges in innovation diffusion and of upscaling in health care, eHealth and prevention. Second, using Rogers' framework of innovation diffusion, we review some of the benefits of the Dutch GLI system for upscaling the Blood Pressure Challenge.

2.1 Upscaling challenges, in eHealth and prevention

Upscaling innovations toward market-wide adoption requires navigating the "chasm" between early adopters and the early majority (Moore, 2014). A primary challenge is managing innovation diffusion, where relative advantage and compatibility with existing user values determine the rate of spread (Rogers, 2003). Organizations often struggle with the Complexity-Consistency Paradox: maintaining core intervention fidelity while allowing for local adaptation and customization (Fixsen et al., 2009). Furthermore, scaling is frequently hindered by the lack of a "business model fit" in adopters and the inability to mobilize resources across varied contexts (Christensen et al., 2006; Teece, 2010). Success requires moving beyond "island solutions" toward systemic integration within the broader market infrastructure (West & Bogers, 2014).

In the ***healthcare and eHealth*** sectors, upscaling is uniquely constrained by institutional complexity and fragmented reimbursement systems (Greenhalgh et al., 2017). Innovations often fail to move beyond the pilot phase—a phenomenon known as "pilotitis"—due to a lack of alignment with clinical workflows and professional identities (May et al., 2009, Egermark et al., 2022). The NASSS framework (Nonadoption, Abandonment, Scale-up, Spread, and Sustainability) highlights that challenges arise from the interplay of multiple domains: the health condition, technology, value proposition, and organizational capacity (Greenhalgh et al., 2017). Additionally, interoperability issues and stringent regulatory requirements (e.g. data privacy, execution protocols or reimbursement rules) create significant barriers to national expansion (Adler-Milstein et al., 2021; Cresswell & Sheikh, 2013). Consequently, sustainable scaling necessitates a co-evolution of technology, health practices, and policy (Lennon et al., 2017).

Upscaling ***prevention and healthy lifestyle interventions*** faces a "prevention paradox," where long-term health gains often conflict with short-term clinical funding models. A primary challenge is the low intensity of lifestyle support in traditional medical workflows, often marginalized in favor of acute care (Fineberg, 2013). Structural success is often hindered by the lack of intersectoral collaboration between healthcare and community-based services (Mair et al., 2012, Ockene et al., 2011). Consequently, scaling requires moving beyond individual behavior change toward "whole-system" approaches that address environmental determinants (Skivington et al., 2021; Giunca et al., 2025).

2.2 Rogers' framework of innovation diffusion & Blood Pressure GLI

Table 1: Summary of GLI system benefits for practitioners of the Blood Pressure GLI service concept, based on Rogers' innovation diffusion factors

Diffusion Factor	Blood Pressure GLI
Attractiveness	Professional legitimacy and structural health care system payments for practitioners
Alignment	Systemic compatibility: integrated into existing routines
Simplicity	Low "setup" effort: using existing GLI ways of working incl. billing and referral relations
Observability	Clear intervention results. Blood pressure as motivating biofeedback. National monitoring systems.
Trialability	Blood Pressure GLI processes easy to imagine & experience by shadowing in ongoing groups

The transition of the Blood Pressure Challenge from a pilot initiative to an accredited GLI represents a strategic shift toward systemic scaling, which can be elucidated through the lens of Rogers' Diffusion of Innovations theory. While NASSS (Greenhalgh et al., 2017) offers a deeper look at systemic complexity, Rogers' framework was selected for its robust ability to evaluate the perceived attributes of the intervention among existing GLI practitioners, which is a primary driver of initial national uptake. Rogers (2003) posits that the adoption rate of an innovation is predominantly governed by five attributes: attractiveness (or relative advantage), alignment (or compatibility), simplicity (or low complexity), observability (or clarity of benefits), and trialability (including previews or shadowing in existing implementations). Table 1 summarizes the GLI system's benefits according to these factors, illustrating how its formalized healthcare infrastructure mitigates some of

the typical barriers associated with the "pilot" phase of medical interventions. Further explanation of Table 1 follows below.

Central to this transformation is the enhancement of ***attractiveness*** through professional and financial legitimacy. While the earlier Blood Pressure Challenge pilots relied primarily on clinical health outcomes to attract interest, the GLI framework provides structural reimbursement via basic health insurance, transforming the intervention from a entrepreneurial project into a standardized revenue stream for practitioners (Greenhalgh et al., 2017, Egermark et al., 2022). Furthermore, the system ensures ***alignment*** through systemic compatibility. Unlike pilot programs that often operate as administrative "add-ons," the GLI aligns with the existing working-, reporting- and billing routines of Dutch primary care. Since a nationwide network of over 1,200 providers is already proficient in similar GLI protocols, the Blood Pressure GLI becomes a more "plug-and-play" solution (Lennon et al., 2017).

Simplicity is enhanced (complexity is reduced) by utilizing pre-established referral and billing pathways, effectively removing the need for practitioners to construct their own operational infrastructure. This ease of implementation is coupled with ***observability***, where the integration into national monitoring systems allows for the validation of results at a population health level, which enhances the existing blood pressure benefits (of using rapid results for coaching purposes) in comparison to most other GLI formats focusing on the more cumbersome metric of weight loss. Finally, regarding ***trialability*** it is easy to imagine for aspirant Blood Pressure GLI adopters to imagine how things will go, thanks to adoption of GLI protocols. Besides, it is relatively normal in the GLI system for aspirant GLI adopters to shadow (= watch along in one or more intervention sessions) with existing GLI groups, which enables a 'try before you buy' experience.

While Rogers identifies five drivers, scholars suggest adding customization to account for the necessary "reinvention" in complex health settings. Effective diffusion requires balancing intervention fidelity with local adaptation to satisfy diverse patient needs and organizational- or community contexts (Fixsen et al., 2009). This flexibility, often termed "reinvention" in implementation science, enhances long-term sustainability and ownership among practitioners (Greenhalgh et al., 2017). Furthermore, customization allows interventions to address socio-

cultural and health nuances of specific populations, thereby also reducing barriers for preventive health (Lorenc et al., 2013). Ultimately, providing practitioners with the agency to tailor non-core components accelerates adoption across heterogeneous healthcare landscapes (May et al., 2009). ***In summary, this provides an evaluation framework with six factors for our user evaluation*** with potential adoption partners, as further described in the next section.

3 Methods and materials

In a design research approach, we follow the methodology of Vaishnavi & Kuechler, (2004) which goes from problem awareness and solution suggestion to development, evaluation and conclusion. Previously, a tentative Blood Pressure GLI service concept was formulated, based on practitioner interactions. In this paper we focus on the first evaluation- and concept redesign phase. After reporting the user study evaluation results in section 4, we discuss redesign options in section 5.

The ***Blood Pressure Challenge phase with twice-daily biofeedback*** consists of[1]:

- Telephone intake & instructions for BP (Blood Pressure) home measurements
- Start- and repeat surveys (BRAVO[2] behaviors and RAND-SF8)
- Group sessions (weeks 1, 2, 3, 6, 10, face-to-face)
- Daily eCoaching in week 1
- (at least three days, via MS Teams or face-to-face)
- Twice-daily BP measurements and logging email
- Feedback on group progress at repeat workshops
- Healthy recipe suggestions
- Content (portal and/or email) on health, BP, and behavior strategies

The updated ***2-year-long Blood Pressure GLI concept*** that we ***user-tested*** was:

- It scales the successful challenge phase, which achieved significant clinical and behavioral improvements:
  - Clinical Impact: Pilots yielded a 10% blood pressure reduction within weeks, effectively halving cardiovascular risks.

[1] This is a 10-week version for municipality-based challenges, providing an extension to employer-based versions.
[2] Bewegen, Roken, Alcohol, Voeding, Ontspanning = Exercise, Smoking, Alcohol, Foods, Relaxation.

  - Health Behaviors (BRAVO survey): Participants demonstrated improved scores, specifically in physical activity, nutrition, and stress management.
  - Well-being (RAND-SF8 survey): Statistically significant increases were recorded in both physical and mental health health-related quality of life.
  - Satisfaction: High participant engagement resulted in group satisfaction scores ranging from 8.5 to 9 out of 10.
- Intake: Baseline survey and twice-daily BP self-monitoring.
- Year 1, Challenge phase: Group sessions (weeks 1, 2, 3, 6, 10, face-to-face)
- Daily eCoaching in week 1 (at least three days, via MS Teams or face-to-face)
- Year 1, Treatment phase: 5 workshops (months 3, 5, 7, 9, 11).
- Year 2, Maintenance phase: 6 Bi-monthly maintenance workshops.
- Year 1 & 2: four Biofeedback measurement weeks (BP logging twice-daily)
- Training for practitioners: 30 hours self-study, half-day workshop, test. Plus 1 or 2 half-day workshops per year as follow up training.
- Operations: Full ePortal support (content, group workflow management, surveys and BP measurements, eTools for health coaching)
- Practitioner support: communication-, education- and coach materials for recruitment and workshops, with room for practitioner inputs.
- Scaling and Efficiency: 4–6 groups of 12–18 participants annually, optimizing 65–80 hours per group.

Participants: From Jan 2026 to Feb 2026, the **user evaluation** was conducted with ***n=7 local health care partners (2 physiotherapists, 2 dietitians and 3 lifestyle coaches) with previous GLI experience***. Since our purpose was a first design evaluation iteration and to 'harvest' design suggestions as soon as possible, we choose a sample size of seven. For this purpose, sometimes five participants can be enough (Virzi, 1992, Nielsen & Landauer 1993), but given some diversity in participant backgrounds, we chose seven. Next, three factors were important in selecting these participants. First, previous GLI experience was very useful for real-world assessments and for extracting design improvement suggestions from them. Second, current GLI execution rests on three professions (physiotherapists, dietitians and lifestyle coaches, with the latter being the majority), hence we chose a representation of 2, 2 and 3 respectively. Finally, we favored practitioners with Blood Pressure challenge experience, which counted for three of our participants.

In terms of evaluation process, first the Blood Pressure GLI concept was explained to them. Next, they were asked for feedback on each of the six innovation diffusion factors from theory: first a grade on a scale from 1 to 10, and then an explanation. Since this was not an evaluation with patients, but with health practitioners on a voluntary basis, no explicit ethics approval procedure was used. Still, all participants did provide explicit consent. The results of the user evaluation are described in the next section.

4 Results, evaluation from GLI practitioners

The results of the user evaluation, conducted among seven implementation partners including physiotherapists, dietitians, and lifestyle coaches with prior experience in GLI execution, provide critical insights into the innovation adoption and diffusion potential of the Blood Pressure GLI.

Table 2: Evaluation of Blood Pressure GLI concept by GLI practitioners (n=7)

Diffusion Factor	Blood Pressure GLI
Attractiveness	Scores: 8, 8, 1, 8.5, --[3] , 8, 1 (Avg. 5.8) Fast, measurable results and daily feedback. Quote: "Blood pressure is a much better motivator than the scale."
Alignment	Scores: 8, 9, 3, 8, -- , 5, 6 (Avg. 6.5) The proposed GLI formats integrate well into existing routines
Simplicity	Scores: 6, 8, 4, 8, -- , 7, 7.5 (Avg. 6.8) The concept and training are seen as realistic. Clear structure and support from the licensor are crucial. Recruitment efficiency is crucial to avoid waiting list drop-outs.
Observability	Scores: 7, 7, 3, 7, -- , 7.5, 8 (Avg. 6.6) Clear results. Blood pressure focus and biofeedback provide unique GLI value. And: “A more extensive product flyer would help me for clarity.”
Trialability	Scores: 8, 9, 3, 9, -- , 10, 8.5 (Avg. 7.9) Strong desire to ‘shadow’ groups & experience a Blood Pressure GLI in practice.
Customization	Scores: 8, 8, 3, 8.5, -- , 7.5, 9 (Avg. 7.3) While valuing support with default materials, all partners express a wish to add their "own flavor" to a lifestyle intervention.

As summarized in user evaluation Table 2 below, the intervention was assessed across six diffusion factors, yielding average scores ranging from 5.8 to 7.9. Across respondents there is large variation in scores. The third respondent had outsourced

[3] As explained, one respondent gave only qualitative feedback and no scores.

all GLI activity to an external partner, given the "incompatibility with our large-scale physiotherapy processes" and financial unattractiveness for them; resulting in low scores on all factors. The fifth respondent gave no scores, because she had ceased doing GLI's, although she still provided this user evaluation with useful inputs on GLI operations and context alignment. The seventh respondent recently shifted to coaching on 'who you are' instead of 'what to do'. Hence the attractiveness of the Blood Pressure GLI for her practice was seen as low (score 1). In contrast with these three respondents, evaluation scores of the other respondents were relatively high.

The qualitative feedback indicates that the intervention's ***attractiveness*** is significantly bolstered by its reliance on concrete, biometric results. Respondents noted that the "powerful feedback loop" provided by daily blood pressure monitoring serves as a superior motivator compared to traditional weight-centric programs, which participants often find emotionally taxing. One respondent familiar with the BP Challenge notes that Blood Pressure coaching provides more 'substance' than an average GLI, more health education and more measurement feedback (RAND SF-8 and BRAVO behaviors, besides blood pressure). The **observability** of benefits (= rapid blood pressure improvements in daily measurements) is perceived as a unique selling point that facilitates both client engagement and professional satisfaction. Other observable benefits mentioned are the clear business case and service concept, although some indicated that they would value a more extensive product description or -flyer.

Regarding ***alignment*** and ***simplicity***, practitioners identified that the Blood Pressure GLI integrates effectively into their existing clinical "GLI factories," provided that systemic barriers such as administrative complexity and mandatory fees to intermediate organizations are minimized. The proposed training and structural framework were deemed realistic, yet the participants emphasized that the simplicity of execution is dependent on a user-friendly digital portal with BP monitoring and the ability to fill groups quickly to prevent dropout during waiting periods.

The highest evaluation scores were observed in ***trialability*** (7.9) and ***customization*** (7.3). There is a significant desire for "trialability" through shadowing existing groups and sessions to build implementation confidence. Furthermore, the results highlight a crucial tension in implementation science: while practitioners highly value

standardized foundational materials, they require the agency for customization. They expressed a consistent need to add a "personal flavor" to the intervention, allowing flexibility to address emergent group needs such as emotional eating or supplemental exercise, provided this customization aids rather than compromises overall quality.

Additional user feedback on Blood Pressure GLI concept and upscaling

Beyond the primary diffusion factors, the user evaluation yielded several critical insights regarding the long-term sustainability and operational viability of the Blood Pressure GLI. A primary concern identified by participating practitioners involves longitudinal ***client retention***, particularly during the maintenance phase of the second year. The initially proposed structure of six sessions in year two is perceived as insufficient to prevent attrition. Based on their experience, they recommend enhancing continuous engagement through digital communication platforms, such as moderated WhatsApp groups, or providing supplemental group-based physical activities to maintain a consistent connection with clients.

Regional collaboration also emerged as a fundamental trait in upscaling. Practitioners emphasized that physical ***outreach to general practitioners*** and practice nurses, supported by communication materials and partnerships with national organizations like the Heart Foundation, is essential for ***generating awareness and high-quality client referrals***. They note that low-quality client referrals create extra work and large drop-out rates, essentially sabotaging feasibility in terms of business case and quality of care.

In summary, the respondents recognized a clear opportunity for the Blood Pressure GLI concept. They also indicated that upscaling success depends on optimizing recruitment efficiency and implementing robust retention strategies throughout the full two-year duration.

5 Discussion and Conclusion

In answer to our Research Question (*To which extent does the Dutch GLI-format help in national upscaling and what is the feasibility and attractiveness of the Blood Pressure GLI for potential adoption partners?*) and based on Rogers' framework and the user study, we find: in addition to the GLI system benefits identified in the Theory and Concepts section (see Table 1), potential adoption practitioners with GLI experience indicate

that the Blood Pressure GLI service concept is attractive and feasible, with some suggestions for improvement (see section 5.2).

There are three limitations to this study we want to discuss. First, this is only a pre-test of a BP GLI service concept. Although it helps that several respondents have previous BP *Challenge* experience and that all respondents have GLI experience, we are still in the design phase. Hence, this is not a true test of market adoption. Second, given the exploratory nature of our study, we only user-tested with n=7 respondents (see Method section). We foresee that a second design and user-test iteration will be needed to further refine the service concept and improve adoption. Still, in the current design exploration, we benefitted a lot from the diversity of their backgrounds (physiotherapists, dietitians and lifestyle coaches) and GLI experiences. Third, regarding adoption, other factors are important too (like market dynamics, regulatory changes, other stakeholder interests etc). So, this study is limited in scope since it focusses mostly on attractiveness and feasibility for potential adopters.

5.1 Implications for theory

Firstly, given the health care context and the fact that funding tariffs are fixed, as well as many of the rules, protocols and obligations, there is a ***special kind of cost-benefit assessment*** for health practice owners. Questions like 'Can we be efficient?' and 'How much work are we required to do?' are ***closely linked to 'What is the quality of the referrals*** of clients from Physician Assistants?' If this referral quality is low (e.g. the wrong target group; too complex patients with too many other problems; not the right motivation or expectations) then the drop-out ratio becomes too high which also stops the funding for these clients and ruins the business case for the practice owner. Besides, quality of care is compromised, which goes against professional pride and standards of practice (hampering diffusion factor alignment/compatibility).

Secondly, intervention ***attractiveness*** has two faces. Besides the business case attractiveness for the ***practice owner*** previously mentioned, there is a second question of attractiveness: for the ***lifestyle coach*** that will execute the intervention. For the coach the following factors and questions are generally important: Will my clients become happy (with this type of biofeedback and coaching, with the results, with the lessons learned)? Do I learn new things, and can I grow my competence

level? Do I get a nice balance between supporting materials/best practices versus options for customization (factor 6 that was added to Rogers' factors)? Again, ***alignment*** with professional norms and ways of working is important.

5.2 Implications for practice & service redesign lessons

To increase attractiveness and feasibility, the following service redesign suggestions followed from the user study:

- Communication materials to help raise awareness and attract high-quality client referrals from Family Physician practices.
- Explicit cooperation with heart foundation (Hartstichting) to aid awareness and connect to existing local 'heart check' initiatives and local care networks.
- Support for the first start-up group for a new BP GLI adopter ("there are always start-up issues that need solving") & for complex BP questions that a new BP GLI adopter runs into.
- Yearly repeat trainings only 1 half-day mandatory.
- More frequent workshops (e.g. monthly) in maintenance year 2.
- WhatsApp group (or other social media) for 2-year client support and retention.
- Workshop schedule (2-year) for adopters: which workshops/exercise are mandatory and where is room for customization?

Next steps we are taking towards national upscaling are: implementing longer term support (+1,5 years) and evaluating its effectiveness, testing with different target-/risk-groups like perimenopausal women, extending needs assessments of health practice owners, and training other health practice owners for BP GLI execution and evaluating health effectiveness as well as BP GLI attractiveness.

5.3 Conclusion

The Dutch GLI systems appears supportive to 'crossing the upscaling chasm' for health prevention interventions. And on a user validation level of the Blood Pressure GLI redesign, based on feedback from the potential adoption practitioners, they see a clear potential for a blood pressure-oriented GLI. In their assessments, the greatest

opportunities lie in the tangibility of the results; the greatest threats lie in recruitment (efficiency and effectiveness) and maintaining participant retention over the full two-year period.

References

Adler-Milstein, J., Garg, A., Zhao, W., & Patel, V. (2021). A survey of health information exchange organizations in advance of a nationwide connectivity framework. *Health Affairs*, *40*(5), 736-744.

Carey, R. M., & Whelton, P. K. (2020). Evidence for the universal blood pressure goal of< 130/80 mm Hg is strong: controversies in hypertension-pro side of the argument. *Hypertension*, *76*(5), 1384-1390.

Christensen, C. M., Baumann, H., Ruggles, R., & Sadtler, T. M. (2006). Disruptive innovation for social change. *Harvard Business Review*, 84(12), 94.

Cresswell, K., & Sheikh, A. (2013). Organizational issues in the implementation and adoption of health information technology innovations: An interpretative synthesis of the literature. *Journal of the American Medical Informatics Association*, 20(e1), e13-e21.

Egermark, M., Blasiak, A., Remus, A., Sapanel, Y., & Ho, D. (2022). Overcoming pilotitis in digital medicine at the intersection of data, clinical evidence, and adoption. *Advanced Intelligent Systems*, *4*(9), 2200056.

Fineberg, H. V. (2013). The paradox of disease prevention: celebrated in principle, resisted in practice. *Jama*, *310*(1).

Fixsen, D. L., Blase, K. A., Naoom, S. F., Van Dyke, M., & Wallace, F. (2009). Implementation: The missing link between research and practice. *NIRN implementation brief*, *1*(1), 218-227.

Giurca, B. C., Lau, M., & Fallows, E. (2025). Lifestyle medicine: Bridging behavioural science and clinical care. *Clinics in Integrated Care*, *33*, 100288.

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., Hinder, S., ... & Shaw, S. (2017). Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research*, 19(11), e367.

Lennon, M. R., Bouamrane, M. M., Devlin, A. M., O'Connor, S., O'Donnell, C., Westerlund, H., ... & Mair, F. S. (2017). Readiness for delivering digital health at scale: Lessons from a longitudinal indirect-comparison study of health technology deployment. *Journal of Medical Internet Research*, 19(2), e42.

Lorenc, T., Petticrew, M., Welch, V., & Tugwell, P. (2013). What types of interventions generate inequalities? Evidence from systematic reviews. *Journal of Epidemiology and Community Health*, 67(2), 190-193.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., ... & Remuzzi, G. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet*, *380*(9859), 2095-2128.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., ... & Pelizzari, P. M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet*, *380*(9859), 2224-2260.

Mair, F. S., May, C., O'Donnell, C., Finch, T., Sullivan, F., & Murray, E. (2012). Factors that promote or inhibit the implementation of e-health systems: An explanatory systematic review of synthesized evidence. *BMJ Open*, 2(3), e000094.

May, C. R., Mair, F., Finch, T., MacFarlane, A., Dowrick, C., Treweek, S., ... & Montori, V. M. (2009). Development of a theory of implementation and integration: Normalization Process Theory. *Implementation Science*, 4(1), 1-18.

Moore, G. A. (2014). Crossing the Chasm: Marketing and Selling Disruptive Products to Mainstream Customers. HarperBusiness, New York.

Nielsen, J., & Landauer, T. K. (1993, May). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213).

Ockene, J. K., Schneider, K. L., Lemon, S. C., & Ockene, I. S. (2011). Recent Advances in Preventive Cardiology and Lifestyle Medicine. *Circulation*, *124*, 1276-1282.

Ostchega, Y., Fryar, C. D., Nwankwo, T., & Nguyen, D. T. (2020). Hypertension prevalence among adults aged 18 and over: United States, 2017–2018, NCHS, National Health and Nutrition Examination Survey, 2017–2018: https://stacks.cdc.gov/view/cdc/87559.

Rogers, E. M. (2003). Diffusion of Innovations (5th ed.). Free Press, New York.

Simons, LPA, Gerritsen, B, Wielaard, B, Neerincx MA (2022). Health Self-Management Support with Microlearning to Improve Hypertension, pp. 511-524, 35th Bled eConference. June 26-29, Bled, Slovenia, Proceedings retrieval from www.bledconference.org. ISBN-13: 978-961-286-616-7, DOI: 10.18690/um/fov.4.2022

Simons, LPA, Gerritsen, B, Wielaard, B, Neerincx, MA (2024). Employee Hypertension Self-Management Support with Microlearning and Social Learning. International Journal of Networking and Virtual Organisations, 30(4), 350-365. https://doi.org/10.1504/IJNVO.2024.140218

Simons, LPA, Wielaard, B, & Neerincx, MA (2025). Beyond Average Results in Hypertension E-Support and Self-Management: Three Pilot Studies With Social Learning. In N. Wickramasinghe (Ed.), Impact of Digital Solutions for Improved Healthcare Delivery (pp. 257-282). IGI Global. https://doi.org/10.4018/979-8-3693-5237-3.ch009

Simons, LPA, Murukannaiah, PK, Neerincx, MA (2026). Dementia Empowerment with Heart Health Intervention and LLM-based Health AI Research Assistant. Chapter in N. Wickramasinghe (ed.), *Digital Convergence: Better Solutions for People with Dementia* (pp.323-345), CRC Routledge. DOI: 10.1201/9781003485681-16

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., ... & Moore, L. (2021). A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. *BMJ*, 374, n1861.

Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, 43(2-3), 172-194.

Virzi, R. A. (1992). Refining the test phase of usability evaluation: how many subjects is enough?. *Human factors*, *34*(4), 457-468.

West, J., & Bogers, M. (2014). Leveraging external sources of innovation: A review of research on open innovation. *Journal of Product Innovation Management*, 31(4), 814-831.

Zhou, B., Perel, P., Mensah, G. A., & Ezzati, M. (2021). Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. *Nature Reviews Cardiology*, *18*(11), 785-802.

Strategic Embedding of Responsible AI: Governance Architecture and Hybrid Human–AI Value Creation

Valtteri Tomminen,[1,2] Matti Leppäniemi,[1]
Tiina Kemppainen,[1] Lauri Frank[2]

[1] University of Jyväskylä, School of Business and Economics, Jyväskylä, Finland
valtteri.j.tomminen@jyu.fi, matti.leppaniemi@jyu.fi, tiina.j.kemppainen@jyu.fi
[2] University of Jyväskylä, Faculty of Information Technology, Jyväskylä, Finland
valtteri.j.tomminen@jyu.fi, lauri.frank@jyu.fi

Artificial intelligence (AI), particularly generative and agentic systems, is increasingly embedded in organizational processes and customer-facing interactions. While responsible AI research emphasizes governance frameworks and ethical considerations, strategic and digital transformation research focuses on performance and innovation outcomes. These perspectives remain insufficiently integrated, and we lack empirical insight into how responsibility becomes structurally embedded within AI-enabled value creation. Drawing on 44 semi-structured interviews with senior professionals, this study develops a conceptual framework explaining how organizations embed responsible AI through (1) strategic AI governance architecture and (2) hybrid human–AI value creation design. The findings show that responsibility is not treated as a compliance layer but enacted as a structural and operational design feature. Governance architectures align AI with strategy and accountability, while hybrid configurations operationalize responsibility through human oversight, controlled autonomy, transparency, and workflow integration. The study contributes by conceptualizing responsible AI as a strategic embedding process that integrates governance and hybrid human–AI design in organizational value creation.

DOI
https://doi.org/10.18690/um.fov.4.2026.20

ISBN
978-961-299-152-4

Keywords:
responsible AI, AI governance, strategic management, ethical AI, hybrid intelligence, human-AI interaction, value creation, generative AI, agentic AI

University of Maribor Press

1 Introduction

Artificial intelligence (AI) is increasingly embedded in the core processes of contemporary organizations. In particular, generative AI (GenAI) and emerging agentic AI systems are transforming how organizations design interactions, automate workflows, and support strategic decision-making (Dwivedi et al., 2023; Grewal et al., 2024; Hosseini & Seilani, 2025; Hughes et al., 2025). AI systems are becoming integral components of organizational infrastructures, shaping operational processes and organizational interactions with stakeholders (Davenport et al., 2020; Schneider et al., 2023). At the same time, the rapid expansion of AI has intensified ethical and governance concerns raising questions related to transparency, bias, fairness, accountability, and data protection (Balasubramaniam et al., 2023; Meske et al., 2022; Mittelstadt et al., 2016). In response, a growing body of research has developed principles and frameworks for responsible AI governance (Mäntymäki et al., 2022; Papagiannidis et al., 2025; Rakova et al., 2021).

However, the existing research streams remain largely fragmented. Responsible AI research tends to frame governance primarily as a compliance, risk-management, or ethical oversight issue (Ashok et al., 2022; Papagiannidis et al., 2025). On the other hand, strategic management and digital transformation research highlight AI's potential for competitive advantage, personalization, and performance improvement (Alsheibani et al., 2020; Keding, 2021; Ruokonen & Ritala, 2024), often without fully integrating responsible governance into core strategy formulation and value creation processes. This separation becomes increasingly problematic as AI systems gain autonomy and directly shape customer-facing interactions and organizational decisions.

Despite the growing importance of responsible AI governance, we lack empirically grounded insights into how responsibility becomes strategically embedded in organizations and how governance mechanisms interact with hybrid human–AI desing in organizational value creation. To address this gap, this study is guided by the following research question: *How do organizations strategically embed responsible AI governance and hybrid human–AI design in organizational value creation?*

Based on 44 semi-structured interviews with senior professionals across industries, this study develops a strategic embedding framework that explains how responsible AI governance becomes integrated into organizational value creation. The findings show that responsibility is embedded through two interrelated dimensions: strategic AI governance architecture and hybrid human–AI value creation design. Together, these dimensions explain how organizations translate responsible AI principles into organizational practices and value creation outcomes.

2 Theoretical Background

Artificial intelligence, especially technologies like Generative and Agentic AI, has evolved from a decision-support technology to an increasingly autonomous actor in organizational processes (Chan & Choi, 2025; Grewal et al., 2024;Baird & Maruping, 2021; Brohi et al., 2025). As these technologies become integrated into customer interactions and internal workflows, they reshape organizational decision-making structures and value creation procesess. This development has amplified ethical and governance challenges, as AI systems often operate as complex and partially opaque socio-technical systems, raising concerns about explainability, bias, discrimination, and accountability (Balasubramaniam et al., 2023; Meske et al., 2022; Mittelstadt et al., 2016). In response, scholars have proposed frameworks for responsible AI governance that emphasize transparency, accountability, documentation, and stakeholder-oriented oversight (Akbarighatar, 2025; Papagiannidis et al., 2025; Rakova et al., 2021).

Responsible AI refers to the design, development, and deployment of AI systems in ways that ensure transparency, accountability, fairness, and alignment with organizational and societal values (Papagiannidis et al., 2025; Zimmer et al., 2022). It emphasizes the governance structures, organizational practices, and oversight mechanisms that guide how AI systems are implemented and used in practice, extending beyond ethical principles toward organizational capabilities that enable the safe and accountable use of AI technologies. AI governance refers to the organizational structures, processes, and practices that translate responsible AI principles into concrete organizational arrangements and guide the responsible use of AI systems in organizations (Madanchian & Taherdoost, 2025; Papagiannidis et al., 2025). Governance is not limited to technical safeguards; it also concerns how decision-making authority and accountability are distributed within organizations

(Gasser & Almeida, 2017; Mäntymäki et al., 2022), including the definition of roles and responsibilities, data governance mechanisms, and lifecycle traceability (Meske et al., 2022). Furthermore, as AI systems are increasingly developed and deployed within broader ecosystems involving technology providers, data suppliers, and partners, responsibility cannot be confined to internal processes but must also encompass ecosystem-level coordination and stakeholder considerations (Mäntymäki et al., 2022; Minkkinen et al., 2023).

This broadens strategic governance from internal compliance to inter-organizational accountability. As AI systems become embedded in organizational workflows, value creation increasingly occurs through hybrid configurations that combine human and artificial intelligence capabilities (Heyder et al., 2023; Mortati & Freitas, 2025). In this study, value creation refers to organizational outcomes that emerge when AI capabilities are effectively integrated into workflows, decision-making, and stakeholder interactions. Therefore, AI-enabled value creation depends on how AI systems are embedded in organizational processes and coordinated with human expertise.

This leads to the concept of hybrid intelligence that emphasizes the complementary integration of human judgment and machine capabilities to enhance performance (Dellermann et al., 2019). Rather than replacing humans, AI augments expertise, automates routine tasks, and enables new forms of coordination (Alsheibani et al., 2020; Heyder et al., 2023; Keding, 2021). However, this hybridization also introduces new ethical and organizational complexities. AI systems may generate plausible yet inaccurate or biased outputs, requiring human oversight and validation (Meske et al., 2022). As Heyder et al. (2023) argue, ethical management of human–AI interaction must consider both structural governance mechanisms and everyday practices through which AI is designed and used. Responsibility thus resides not only in formal policies but also in how human actors interpret, supervise, and correct AI outputs.

Therefore, transparency and disclosure practices therefore become central to maintaining trust in AI-mediated interactions (Balasubramaniam et al., 2023). From a value creation design perspective, hybrid human–AI systems must balance autonomy and control. Strategic management of hybrid intelligence therefore requires appropriate design choices regarding human oversight, workflow

integration, and decision authority leading to value creation outcomes (Dellermann et al., 2019; Heyder et al., 2023). Table 1 summarizes the key theoretical concepts guiding this study.

Table 1: Key theoretical concepts

Theoretical concept	Definition	Key studies
Responsible AI	The design, development, deployment, and use of AI systems in ways that support transparency, accountability, fairness, and alignment with organizational and societal values.	Papagiannidis et al. (2025); Zimmer et al. (2022); Balasubramaniam et al. (2023); Mittelstadt et al. (2016)
Responsible AI Governance	The organizational structures, processes, and practices through which organizations guide, monitor, and control the responsible use of AI systems.	Mäntymäki et al. (2022); Papagiannidis et al. (2025); Rakova et al. (2021)
Hybrid Intelligence	The complementary integration of human judgment and AI capabilities to enhance organizational performance and decision-making.	Dellermann et al. (2019); Heyder et al. (2023); Mortati & Freitas (2025)
Human–AI Value Creation	Value creation that emerges when AI capabilities are integrated into workflows, decision-making, and stakeholder interactions together with human expertise and oversight.	Dellermann et al. (2019); Heyder et al. (2023); Davenport et al. (2020); Blümel et al. (2024)

3 Methodology

This study adopts a qualitative, inductive research design to examine how organizations strategically embed responsible AI governance and hybrid human–AI structures in practice. Accordingly, the design emphasizes rich, context-sensitive accounts of managerial interpretations, organizational practices, and governance mechanisms (Langley, 1999), enabling an in-depth exploration of how governance structures and hybrid configurations are constructed and enacted within organizations.

Empirical data were collected through 44 semi-structured interviews with senior professionals in AI-intensive and digitally transforming organizations. Participants were recruited through purposive sampling to ensure representation from AI-intensive and digitally transforming organizations across multiple industries. Selection criteria emphasized senior professionals with direct strategic, governance, or implementation responsibilities related to AI adoption. Informants represented diverse roles, including CEOs, CIOs, AI directors, marketing directors, strategy

managers, and transformation leaders across industries. Interviews were conducted between February and December 2025 and ranged from 34 to 87 minutes, resulting in approximately 2,400 minutes of recorded material. All interviews were recorded and transcribed verbatim. The interview guide followed a semi-structured format that ensured consistency while allowing flexibility to explore emerging themes (Brinkmann & Kvale, 2015). The guide covered themes such as AI adoption, governance structures, human oversight, workflow integration, and value creation. Interviews continued until thematic saturation was reached, meaning that later interviews generated limited new conceptual insights and primarily reinforced existing themes.

Data analysis followed an inductive approach. Rather than applying a predefined theoretical framework, the analysis sought to identify recurring patterns in how organizations described embedding responsibility into AI-enabled value creation. Interview transcripts were first reviewed and coded to capture themes related to governance structures, data management, accountability mechanisms, hybrid work practices, and customer-facing value creation processes. To enhance credibility and consistency, coding was done by the author team using Nvivo, and coding decision were discussed among the authors. The coding process remained open to unexpected insights and cross-cutting themes. Emerging codes were then compared and grouped into higher-level categories reflecting broader organizational patterns. Particular attention was paid to how governance mechanisms connected to strategic objectives and how human oversight and workflow design shaped hybrid value creation. Through constant comparison across cases, conceptual distinctions were refined and patterns clarified (Corbin & Strauss, 2008). In the final stage, two overarching dimensions were identified: (1) strategic AI governance architecture and (2) hybrid human–AI value creation design. These dimensions synthesize how responsibility becomes structurally embedded in governance arrangements and operationalized within value creation processes.

4 Findings

The findings show that responsible AI is not treated as a separate compliance initiative but becomes strategically embedded through two interrelated dimensions: strategic AI governance architecture and hybrid human–AI value creation design.

Together, these dimensions explain how organizations integrate responsibility into both structural governance arrangements and value creation processes.

The first dimension captures how organizations construct strategic AI governance architectures that embed AI into business strategy, formal accountability mechanisms, and cross-functional coordination. Rather than viewing governance as a restrictive layer, informants described it as an enabling structure that supports AI-driven innovation while managing risk and accountability. Interviewees repeatedly emphasized that AI must be integrated into business strategy and core processes rather than treated as a technological experiment: *"You don't gain a competitive advantage from just adopting AI. It's about how you implement them and how deeply you can integrate them into your processes, that's where you can gain a competitive advantage."* (CEO & Co-Founder). Beyond strategic alignment, organizations described the implementation of formal governance mechanisms, including clear rules, documentation practices, and defined roles. Governance was characterized as systematic and disciplined rather than bureaucratic obstruction: *"You need to have some kind of understanding of AI governance and you need to have your own AI strategy for what you do, what customer data you use, what employee data you use"* (Development Director).

Data governance and information control were central components of this architecture. Informants emphasized restrictions regarding sensitive data and explicit accountability structures: *"So we have ethical restrictions for AI. For example, we don't feed personal data into AI at all, which is a no-brainer for us. All business-critical personal data is excluded."* (Growth Marketing Director). Importantly, even as AI systems become more autonomous, ultimate responsibility was consistently anchored in human actors: *"We always have human oversight into these matters. Ultimately, we ourselves are responsible for what we do"* (VP, Digital Customer Experience). Strategic governance was also described as extending beyond internal structures toward ecosystem-level responsibility. Collaboration with external partners, agencies, and technology providers complicated accountability arrangements: *"In our case, this is a bit problematic because we have our own organization, but we also work with external stakeholders and partners. This perhaps adds an extra challenge to AI management, as not all parts of the process take place in-house."* (Director, Digital).

The second dimension explains how responsibility is enacted within everyday value creation processes through hybrid human–AI configurations. Informants consistently described AI not as a fully autonomous substitute for human judgment, but as part of a deliberately designed hybrid system. AI systems were perceived as powerful yet imperfect, requiring validation and supervision: *"How to make AI sufficiently reliable and high-quality is still partly unresolved, we want to maintain human oversight, because ultimately, whether what we do is done by humans or AI, we are still responsible for the end result and we need to know these things."* (VP, Digital Customer Experience). Some interviewees described layered control mechanisms embedded within AI systems: *"AI errors and bias can be reduced by having, for example, an agent generating an email, and then having another agent check the email before it is sent."* (AI Lead & Senior Consultant).

Human oversight was therefore not a temporary safeguard but a structural feature of hybrid system design. Hybrid human-AI value creation was particularly visible in customer-facing processes, where AI enables personalization and automation. However, informants also recognized ethical tensions associated with hyper-personalization and predictive reasoning: *"Where customers and their expectations are going right now is that they expect personalization. We are already talking about hyper-personalization, and when AI is included in that, we can build an even more personalized shopping experience."* (Vice President, Marketing). However, excessive personalization was also perceived as potentially intrusive: *"Yes, AI personalizes, but that can be both good and bad. On the one hand, the experiences are so seamless and personalized, but on the other, there are privacy issues, which can lead to manipulation and a loss of consumer trust."* (Creative Strategist). Transparency emerged as a key mechanism for maintaining legitimacy and trust in AI-mediated interactions: *"From my point of view, transparency is vital in the utilization of AI in a company, for its customers, employees, and all stakeholders."* (Chief Marketing Officer).

Organizations also emphasized that automation must be introduced in a controlled and context-sensitive manner. Rather than maximizing autonomy, firms described deliberate decisions about where and how AI should operate: *"It's not necessarily smart to automate everything with AI, and then there's the fact that someone still has to check that everything is right, that it's not removing people from the equation."* (Team Manager, Customer Service). Hybrid systems were designed to integrate AI into existing workflows instead of introducing detached tools: *"Then you have to think about different strategies, it*

has to be connected to the workflow, so that it's not something new that you have to start using a new tool, but rather something that you experience as part of your work." (GenAI Lead).

Overall, the findings are summarized in table 2 indicating that responsible AI governance is enacted as a strategic architecture combining leadership commitment, formalized rules, data control, human accountability, and ecosystem awareness. Governance architecture defines the structural conditions of AI deployment, while hybrid design operationalizes responsibility through human oversight, controlled autonomy, personalization boundaries, and workflow integration. Together, these dimensions illustrate how responsibility becomes embedded in both organizational structures and AI-driven practices.

Table 2: Overview of Findings

Dimension	Core Subthemes	Illustrative Evidence
Strategic AI governance architecture	Leadership commitment; AI-strategy alignment; accountability roles; data governance; ecosystem coordination	*"You need to have AI governance and your own AI strategy" (Development Director).* *"We have our own organization, but we also work with external stakeholders and partners. This perhaps adds an extra challenge to AI management" (Director, Digital)*
Hybrid human-AI value creation design	Human oversight; autonomy boundaries; layered validation; workflow integration; transparency	*"We want to maintain human oversight, because ultimately, whether what we do is done by humans or AI, we are still responsible for the end result and we need to know these things." (VP, Digital Customer Experience).* *"Then you have to think about different strategies, it has to be connected to the workflow" (GenAI lead).*

5 Discussion

This study examined how organizations strategically embed responsible AI governance and hybrid human–AI structures to enable value creation. The findings indicate that responsibility does not emerge from abstract ethical principles alone, nor from isolated governance mechanisms. Instead, responsible AI becomes embedded through the interplay of strategic AI governance architecture and deliberately designed hybrid human–AI value creation. The findings and main contribution of the study culminate in the *Strategic Embedding Framework of Responsible AI* (Figure 1). The framework integrates strategic AI governance architecture and hybrid human–AI value creation design into a unified explanatory framework. It conceptualizes responsible AI as a dynamic organizational capability that emerges from the reciprocal interaction between structural governance mechanisms and

operational hybrid configurations, thereby generating distinct value creation outcomes that arise when responsible AI governance and hybrid human–AI value creation design operate together. This perspective shifts responsible AI from a compliance-oriented concern to a strategic design challenge embedded in governance structures and hybrid human–AI value creation.

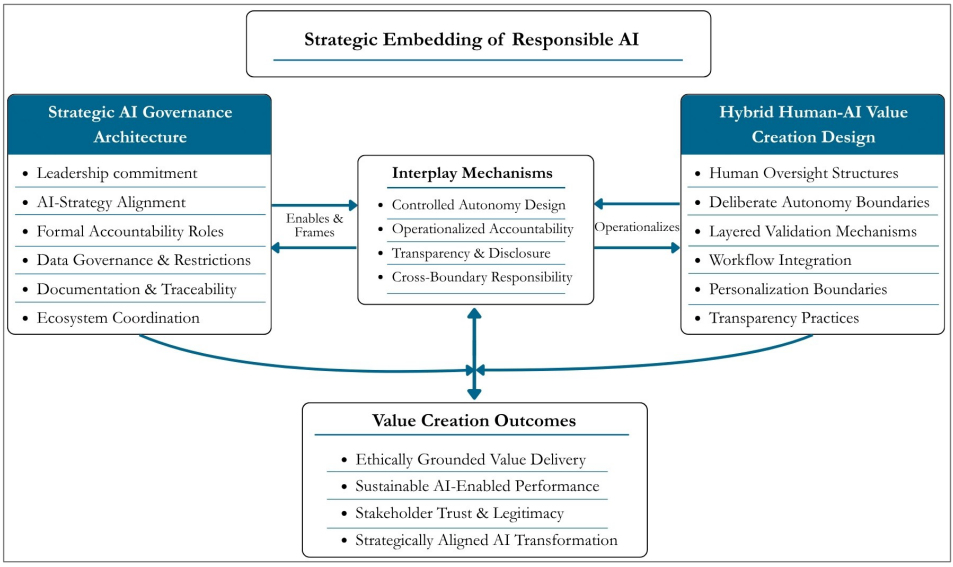

Figure 1: Strategic Embedding Framework of Responsible AI

The first core dimension of the framework, strategic AI governance architecture, captures how organizations anchor responsibility at the structural level. Governance is embedded in leadership commitment, AI–strategy alignment, formal accountability roles, data governance infrastructures, documentation practices, and ecosystem coordination mechanisms. Rather than functioning as an external compliance layer, governance operates as an enabling architecture that shapes the scope, conditions, and boundaries of AI deployment. In this respect, governance constitutes a strategic capability that supports innovation, performance, and ethical alignment simultaneously.

The second core dimension, hybrid human–AI value creation design, explains how responsibility is operationalized in practice. Hybrid configurations translate governance principles into workflows, customer-facing processes, and decision-

making routines. The findings show that organizations deliberately design human oversight structures, layered validation mechanisms, autonomy boundaries, workflow integration practices, and transparency mechanisms to ensure that AI outputs remain accountable, contextually appropriate, and aligned with stakeholder expectations. Hybrid design thus functions not only as a performance-enhancing arrangement but also as an ethical infrastructure that distributes authority and responsibility across human and artificial actors.

The framework further explicates the mechanisms linking governance architecture and hybrid system design. Governance enables controlled autonomy by defining the strategic and ethical parameters within which AI systems operate, while hybrid configurations operationalize accountability by embedding oversight and validation into daily work processes. Transparency mechanisms connect internal governance arrangements to external legitimacy by making AI-mediated activities visible and interpretable to stakeholders. As AI systems operate within broader ecosystems of technology providers and partners, cross-boundary coordination becomes central to responsible AI embedding.

Together, these dimensions generate value creation outcomes when strategic AI governance architecture and hybrid human–AI value creation design operate together. Organizations create value not merely by adopting AI technologies but by embedding responsibility into how AI is governed and applied in practice. These outcomes include ethically grounded value delivery, sustainable AI-enabled performance, stakeholder trust and legitimacy, and strategically aligned AI transformation. Controlled autonomy and human oversight support durable performance, while transparency and accountability strengthen legitimacy in AI-mediated interactions. Together, these outcomes suggest that responsible AI creates value by enabling AI deployment that is reliable, accountable, and strategically integrated across organizational processes.

This study contributes to the responsible AI governance literature in several ways. Prior research has emphasized ethical principles, transparency, documentation, and accountability mechanisms (Mäntymäki et al., 2022; Papagiannidis et al., 2025; Rakova et al., 2021), often framing governance as a compliance-oriented or risk-mitigation activity (Ashok et al., 2022). Our findings extend this perspective by showing that governance becomes strategically embedded within organizational

architecture. Leadership commitment, AI–strategy alignment, formal accountability structures, and disciplined data governance are not peripheral safeguards but core components of AI strategy. Consistent with calls to conceptualize AI governance as a strategic element (Papagiannidis et al., 2025), the framework suggests how governance defines the parameters of controlled autonomy and shapes AI deployment from the outset. Governance thus operates as an enabling infrastructure that supports innovation while constraining risk, bridging strategic management perspectives on AI capability development (Alsheibani et al., 2020; Keding, 2021; Schneider et al., 2023) with responsible AI scholarship.

Furthermore, the study contributes to research on hybrid intelligence and human–AI interaction. Prior work conceptualizes hybrid intelligence as the complementary integration of human and machine capabilities to enhance performance (Dellermann et al., 2019). Our findings extend this perspective by suggesting that hybrid design functions as an ethical infrastructure. Human oversight structures, layered validation mechanisms, autonomy boundaries, workflow integration, and transparency practices institutionalize responsibility within value creation. Rather than treating human involvement as a temporary corrective to imperfect AI systems, organizations embed it as a structural design principle. Hybrid intelligence is thus framed not merely as a logic of augmentation, but as a mechanism through which accountability and explainability are enacted in practice.

6 Conclusion

This study examined how organizations strategically embed responsible AI governance and hybrid human–AI structures in value creation processes. Based on 44 interviews, we developed a conceptual framework demonstrating that responsible AI is enacted through (1) strategic AI governance architecture and (2) hybrid human–AI value creation design. Together, these dimensions explain how organizations balance autonomy and accountability, personalization and trust, and innovation and control. The study advances responsible AI governance research by reframing governance as strategic embedding rather than regulatory compliance, extends hybrid intelligence research by conceptualizing hybrid configurations as ethical infrastructures, and shows that responsible AI constitutes an organizational capability reflected in value creation outcomes. As AI systems become increasingly autonomous and organizationally embedded, responsibility depends on

organizations' capacity to design governance architectures and hybrid configurations that continuously align autonomy with accountability. Future research may build on this framework by examining longitudinal developments, sectoral variations, and regulatory influences on governance–hybrid interplay.

Acknowledgment

This study was funded by Business Finland (grant number 6659/31/2023).

References

Akbarighatar, P. (2025). Operationalizing responsible AI principles through responsible AI capabilities. *AI and Ethics (Online)*, *5*(2), 1787–1801. https://doi.org/10.1007/s43681-024-00524-4

Alsheibani, S., Cheung, Y., Messom, C., & Alhosni, M. (2020). Winning AI strategy: Six-steps to create value from artificial intelligence. *AMCIS 2020 Proceedings. 1.*

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management*, *62*, 102433. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, *45*(1), 315–341. https://doi.org/10.25300/MISQ/2021/15882

Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, *159*, 107197. https://doi.org/10.1016/j.infsof.2023.107197

Blümel, J. H., Zaki, M., & Bohné, T. (2024). Personal touch in digital customer service: A conceptual framework of relational personalization for conversational AI. *Journal of Service Theory and Practice*, *34*(1), 33–65. https://doi.org/10.1108/JSTP-03-2023-0098

Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (Third edition). Sage Publications.

Brohi, S., Mastoi, Q., Jhanjhi, N. Z., & Pillai, T. R. (2025). A Research Landscape of Agentic AI and Large Language Models: Applications, Challenges and Future Directions. *Algorithms*, *18*(8), 499. https://doi.org/10.3390/a18080499

Chan, H.-L., & Choi, T.-M. (2025). Using generative artificial intelligence (GenAI) in marketing: Development and practices. *Journal of Business Research*, *191*, 115276. https://doi.org/10.1016/j.jbusres.2025.115276

Corbin, J., & Strauss, A. L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed). Sage.

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid Intelligence. *Business & Information Systems Engineering*, *61*(5), 637–643. https://doi.org/10.1007/s12599-019-00595-2

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research,

practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Gasser, U., & Almeida, V. A. F. (2017). A Layered Model for AI Governance. *IEEE Internet Computing*, *21*(6), 58–62. https://doi.org/10.1109/MIC.2017.4180835

Grewal, D., Satornino, C. B., Davenport, T., & Guha, A. (2024). How generative AI Is shaping the future of marketing. *Journal of the Academy of Marketing Science*. https://doi.org/10.1007/s11747-024-01064-3

Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. *The Journal of Strategic Information Systems*, *32*(3), 101772. https://doi.org/10.1016/j.jsis.2023.101772

Hosseini, S., & Seilani, H. (2025). The role of agentic AI in shaping a smart future: A systematic review. *Array (New York)*, *26*, 100399. https://doi.org/10.1016/j.array.2025.100399

Hughes, L., Dwivedi, Y. K., Li, K., Appanderanda, M., Al-Bashrawi, M. A., & Chae, I. (2025). AI Agents and Agentic Systems: Redefining Global it Management. *Journal of Global Information Technology Management : JGITM*, *28*(3), 175–185. https://doi.org/10.1080/1097198X.2025.2524286

Keding, C. (2021). Understanding the interplay of artificial intelligence and strategic management: Four decades of research in review. *Management Review Quarterly*, *71*(1), 91–134. https://doi.org/10.1007/s11301-020-00181-x

Langley, A. (1999). Strategies for Theorizing from Process Data. *The Academy of Management Review*, *24*(4), 691–710. https://doi.org/10.2307/259349

Madanchian, M., & Taherdoost, H. (2025). Ethical theories, governance models, and strategic frameworks for responsible AI adoption and organizational success. *Frontiers in Artificial Intelligence*, *8*, 1619029. https://doi.org/10.3389/frai.2025.1619029

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *Ai and Ethics (Online)*, *2*(4), 603–609. https://doi.org/10.1007/s43681-022-00143-x

Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. *Information Systems Management*, *39*(1), 53–63. https://doi.org/10.1080/10580530.2020.1849465

Minkkinen, M., Zimmer, M. P., & Mäntymäki, M. (2023). Co-Shaping an Ecosystem for Responsible AI: Five Types of Expectation Work in Response to a Technological Frame. *Information Systems Frontiers*, *25*(1), 103–121. https://doi.org/10.1007/s10796-022-10269-2

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679

Mortati, M., & Freitas, G. (2025). AI in Service Design: A New Framework for Hybrid Human–AI Service Encounters. *Journal of service research, 29*(1), 81-96. https://doi.org/10.1177/10946705251344387

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2), 101885. https://doi.org/10.1016/j.jsis.2024.101885

Rakova, B., Yang, J., Cramer, H., & Chowdhury, R. (2021). Where Responsible AI meets Reality: Practitioner Perspectives on Enablers for Shifting Organizational Practices. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 1–23. https://doi.org/10.1145/3449081

Ruokonen, M., & Ritala, P. (2024). How to succeed with an AI-first strategy? *The Journal of Business Strategy*, *45*(6), 396–404. https://doi.org/10.1108/JBS-08-2023-0178

Schneider, J., Abraham, R., Meske, C., & Vom Brocke, J. (2023). Artificial Intelligence Governance For Businesses. *Information Systems Management*, *40*(3), 229–249. https://doi.org/10.1080/10580530.2022.2085825

Zimmer, M., Minkkinen, M., & Mäntymäki, M. (2022). Responsible Artificial Intelligence Systems Critical considerations for business model design. *Scandinavian journal of information systems, 40*(2), 4.

From Algorithm to Doorstep: Consumer Acceptance of Chatbots in Last-mile Logistics

Alinda Kokkinou,[1] Hans Quak,[1] Ondrej Mitas[2]

[1] Breda University of Applied Sciences, Academy of Built Environment and Logistics, Breda, the Netherlands
kokkinou.a@buas.nl, quak.h@buas.nl
[2] Breda University of Applied Sciences, Academy for Tourism, Breda, the Netherlands
mitas.o@buas.nl

Failed deliveries are a major source of inefficiency in last-mile logistics that could be addressed by AI-enabled chatbots that facilitate real-time interaction between parcel recipients and Logistics Service Providers. This study examines the factors driving parcel recipient intention to use AI-enabled chatbots in this context, drawing on an extended Unified Theory of Acceptance and Use of Technology (UTAUT) model and a scenario-based experiment with 214 university students. Three chatbot conditions were compared: a well-functioning chatbot, an error-prone chatbot, and a privacy-invasive chatbot. Results show that chatbot errors indirectly reduce future usage intentions through lower user satisfaction. Furthermore, expected performance and social influence are the primary predictors of intention to use. These findings extend UTAUT to a utilitarian, post-purchase last-mile logistics context and offer practical guidance for Logistics Service Providers seeking to improve last-mile performance through AI-enabled chatbots.

DOI
https://doi.org/10.18690/um.fov.4.2026.21

ISBN
978-961-299-152-4

Keywords:
artificial intelligence,
UTAUT,
last-mile logistics,
satisfaction,
consumer behavior

University of Maribor Press

1 Introduction

Convenience, variety, and competitive pricing have led to a sustained surge in Business-to-Consumer (B2C) e-commerce (Amaya et al., 2025; Nogueira et al., 2024). Consumers are ordering more frequently online and expecting swift delivery, preferably at home, and for free (Buldeo Rai et al., 2019). As these consumers increasingly settle in cities, this is compounding existing congestion problems, negatively impacting the social and ecological performance of supply chains. Given that the last-mile is also known for being the most expensive part of the supply chain, Logistics Service Providers (LSPs) are facing increasing pressure to improve the triple bottom line performance of their operations (Mangiaracina et al., 2015) while maintaining customer satisfaction (Kokkinou et al., 2026). A source of poor last-mile performance is the high rate of unsuccessful deliveries, which has estimates ranging from 1 to 60%, with 20% frequently mentioned as the upper limit (Niemeijer & Buijs, 2023). The primary reason for failed delivery is typically that the customer is not home (Norman, 2025). These failed deliveries lead to high costs for e-commerce players, negatively affect customer satisfaction, multiply emissions and congestion through repeated delivery attempts, and add pressure on drivers (Seghezzi & Mangiaracina, 2023).

Several solutions have been proposed to reduce failed deliveries and their negative consequences on last-mile performance. Unattended delivery options, such as parcel lockers, are effective but not accessible to all customers (Kokkinou & Quak, 2026). More recently, the use of Internet of Things (IoT) in combination with data analytics have been used to predict whether recipients will be home and (re)arrange routes accordingly (Seghezzi & Mangiaracina, 2023). Nevertheless, this only reduces the probability of failed delivery. Increasingly, customers and parcel recipients demand more control over when, where, and how their parcels are delivered (Norman, 2025). The development of Generative Artificial Intelligence (GenAI), and in particular AI-enabled chatbots, is a promising approach to aligning customer delivery preferences with the options offered by the LSP. AI-enabled chatbots (hereafter referred to as chatbots) are a class of intelligent software agents that facilitate interactive information sharing (Arce-Urriza et al., 2025). Driven by advances in GenAI, they are increasingly adopted across sales, marketing, and customer service for their capacity to deliver engaging, human-like conversations (Pillai & Sivathanu, 2020). Chatbots can facilitate real-time interaction between parcel recipients and LSPs,

enabling them to (re)arrange delivery (time and place), thereby reducing missed deliveries and potentially increasing customer satisfaction (Richey et al., 2023). By offering conversational rescheduling options, chatbots can reduce friction and meet recipients' need for flexibility and autonomy (Alsarayreh et al., 2025; Seegers, 2025). Nevertheless, these outcomes can only be achieved if customers and parcel recipients are willing to use chatbots, a topic that has received limited attention (Richey et al., 2023).

The purpose of the present study is therefore to examine consumers' intention to use chatbots in the context of last-mile deliveries. Chatbot features have been examined across various contexts, including e-commerce, finance, hospitality, tourism, and retail (Alharbi et al., 2025; Arce-Urriza et al., 2025; Liu et al., 2024; Pillai & Sivathanu, 2020). Factors that have been examined include technology aspects, namely ease of use, usefulness, and trust, in combination with chatbot-specific features such as anthropomorphism and chatbot intelligence (Alharbi et al., 2025; Liu et al., 2024; Pillai & Sivathanu, 2020) Yet, last-mile logistics as a study context differs from the hospitality, tourism, and retail contexts due to the utilitarian nature of the interaction and the fact that it occurs post-purchase (Vakulenko et al., 2019). Our study thereby contributes to a better understanding of chatbot usage in a utilitarian, post-purchase context, where the primary goal is the efficient execution of a time-sensitive task, a context where empirical evidence on consumer acceptance remains limited (Alsarayreh et al., 2025; Richey et al., 2023), despite growing interest from the industry (King, 2024; Seegers, 2025).

2 Review of the Literature

In supply chains and logistics, the last mile is known as the most expensive (Gevaers et al., 2014) and polluting (Mangiaracina et al., 2019). This negative impact of last-mile logistics has been addressed through a mix of supply-side and demand-side innovations and interventions. Specifically, LSPs have made efforts to reduce the number and frequency of trips, drawing on efficient routing and the principle of consolidation (Savelsbergh & Van Woensel, 2016). Nevertheless, customers and parcel recipients play an important role therein, as a substantial number of trips are caused by recipients not being home to receive their parcels. LSPs have invested significant resources to improve communication with parcel recipients. For example, most large LSPs offer customers and parcel recipients the possibility to adjust

delivery preferences through an online application (Norman, 2025). For example, in the Netherlands, DHL and PostNL offer parcel recipients the option to leave a package at their neighbours', in a safe location around the house, or have it rerouted to a nearby parcel locker (DHL, n.d.; PostNL, n.d.). Furthermore, LSPs increasingly partner with retailers to offer customers the option to select a delivery time and location at checkout. While these interventions are successful, they remain quite unilateral, leaving the initiative with the parcel recipient. Parcel recipients increasingly look for flexibility and convenience, with for example up to 35% of deliveries in the UK being re-directed while already in transit (King, 2024). There is thus much to be gained from offering real-time interaction between parcel recipients and LSPs that reduces friction in two-way communication (Alsarayreh et al., 2025; Seegers, 2025). Recent advances in GenAI have transformed chatbots from rule-based systems into flexible agents capable of natural, context-aware dialogue (Alharbi et al., 2025). Their use is widespread and expanding rapidly. This is unsurprising, as chatbots allow organizations to reduce operational costs while concurrently improving customer experience, user engagement, and satisfaction (Alsarayreh et al., 2025). Despite extensive research across numerous contexts (Alharbi et al., 2025), the adoption of chatbots in utilitarian, post-purchase settings such as last-mile logistics remains under-researched (Alsarayreh et al., 2025; Arce-Urriza et al., 2025).

Several research models have been used to explain users' intention to adopt a particular focal technology. The two broadest models are the Technology Acceptance Model (TAM) (Davis et al., 1989) and the Unified Theory of Acceptance and Usage of Technology (UTAUT) (Venkatesh et al., 2012). Both theories focus on the functional aspects of technology, such as its perceived usefulness and perceived ease of use (Kasilingam, 2020). UTAUT extends TAM by incorporating social factors, namely facilitating conditions and social influence (Arce-Urriza et al., 2025; Venkatesh et al., 2012). TAM and UTAUT are operationalized into consistent and concise tools of measurement (Kasilingam, 2020) that have demonstrated empirical soundness across a broad range of contexts (Alharbi et al., 2025; Arce-Urriza et al., 2025). Two other models are also relevant for the present study. The Expectation-Confirmation Model (ECM) similarly focuses on users' cognitive processes but also incorporates satisfaction with prior use of the focal technology to explain continued use (Bhattacherjee, 2001). Yet another model, the Technology Readiness Index (TRI) is a more general measure of individuals' technology

readiness, captured by four components: optimism, innovativeness, discomfort, and insecurity (Parasuraman & Colby, 2015).

UTAUT and TAM were selected as more appropriate, as they are parsimonious yet comprehensive models, and have been applied in contexts relevant to this study (Alsarayreh et al., 2025), such as smartphone chatbots (Kasilingam, 2020), chatbots in retail communication (Rese et al., 2020), and chatbots in service settings (Liu et al., 2024; Pillai & Sivathanu, 2020). The TAM and UTAUT models have been extended to account for the specific characteristics of chatbots, and in particular relational elements. These can generally be grouped into two categories. The first category comprises variables describing how human-like the chatbot is, including anthropomorphism (Pillai & Sivathanu, 2020), perceived humanness (Arce-Urriza et al., 2025), and the authenticity of the conversation (Rese et al., 2020). The second category revolves around the concept of trust. Given that chatbots facilitate two-way communication, they may facilitate the exchange of sensitive data (Arce-Urriza et al., 2025). Several variables are linked to trust, namely, perceived risk (Kasilingam, 2020) and privacy concerns (Alharbi et al., 2025). While these models have been tested across diverse contexts, little research has examined their application to last-mile logistics. This context is characterized by utilitarianism and its post-purchase setting (Arce-Urriza et al., 2025). Therefore, the present study seeks to answer the following research question: *What factors drive consumer usage of chatbots in the context of last-mile logistics?*

3 Conceptual Framework

To understand parcel recipients' intentions to use chatbots, three models were tested. Consistent with previous research, the first model examined an augmented UTAUT model. The second model compared three different chatbots developed based on two salient aspects identified in the literature review, namely quality (translated to error proneness) and privacy. The third model examined whether the expanded UTAUT model could explain the effects of the different chatbots on future intention to use them.

The *first model, based on an augmented UTAUT model,* was used to identify the factors explaining parcel recipients' intention to use Chatbots for last-mile logistics. This model included the broadly accepted UTAUT variables (Alsarayreh et al., 2025;

Kasilingam, 2020; Venkatesh et al., 2012): expected performance (*HA1*), ease of use (*HA2*), social influence (*HA3*), and facilitating conditions (*HA4*). Additionally, perceived risk (*HA5*) was also included, given its importance in both chatbot and last-mile research (Zhou et al., 2020). Finally, given that satisfaction with past chatbot interactions has been shown to predict future intentions (Ashfaq et al., 2020), and consistent with ECM (Bhattacherjee, 2001), satisfaction was also included as a predictor (*HA5*).

The *second model was an experimental model based on two salient factors* identified in the literature review. These factors were found to play a role in consumers' satisfaction, attitude, and intention to use chatbots, namely the quality (Toader et al., 2019) and privacy aspects of the chatbot (Arce-Urriza et al., 2025). Errors by chatbots can frustrate users and lead them to perceive chatbots as less reliable and competent (Alharbi et al., 2025). In the related context of human-robot interaction, errors by agents similarly led to them being perceived as less reliable and competent, and having weaker reasoning abilities (Ragni et al., 2016). In an experiment comparing four chatbots in a hedonic pre-purchase context, De Sá Siqueira et al. (2024) found that attitude towards the chatbot was negatively impacted by errors, and this effect was not mitigated by human-like characteristics. We therefore hypothesize that errors by the chatbot decrease parcel recipients' intention to use chatbots for last-mile logistics (*HB1*). Interaction with chatbots typically involves two-way communication, with users sharing information so the chatbot can better understand their queries and provide more accurate answers. Nevertheless, this raises concerns about privacy and risk. Previous research showed that even as customers became more accustomed to AI technologies, their concerns about privacy remained unchanged (Arce-Urriza et al., 2025). Therefore, parcel recipients, confronted with an inquisitive chatbot that asks personal questions, may feel at risk, leading to the following hypothesis: private questions by the chatbot decrease parcel recipients' intention to use chatbots for last-mile logistics (*HB2*).

The *third model tested was a mediation model*, combining the previous two, and illustrated in Figure 1. Using TAM as a research model, De Sá Siqueira et al. (2024) found that errors by chatbots negatively affected perceived usefulness, the TAM counterpart of the UTAUT construct expected performance. We therefore hypothesize that: expected performance mediates the effect of errors by the chatbot on future intention to use (*HM1*). Similarly, chatbot errors are likely to decrease satisfaction,

and we therefore hypothesize that satisfaction mediates the effect of chatbot errors on future intention to use (*HMS*). Finally, as an inquisitive chatbot that asks personal questions may trigger feelings of risk, we hypothesize that perceived risk mediates the effect of private questions by the chatbot on future intention to use (*HM2*).

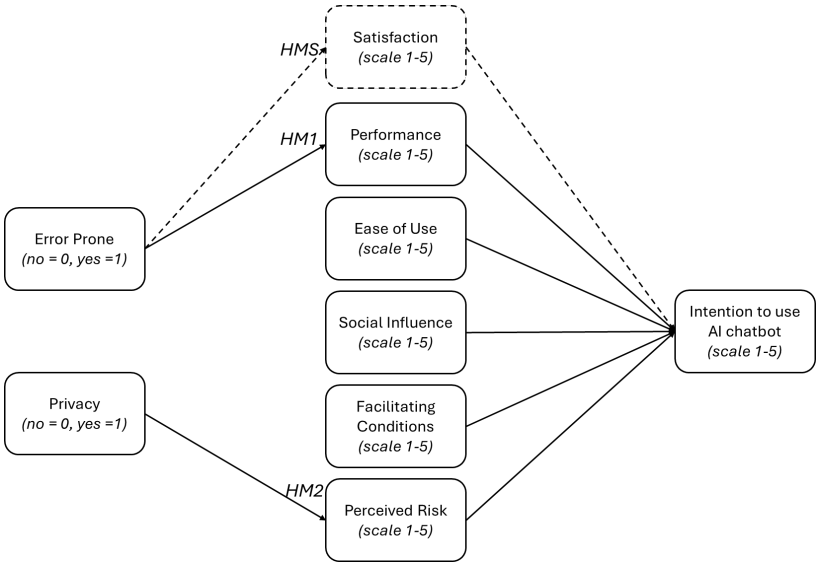

Figure 1: Mediation Model
Source: Own

4 Methods

An online scenario-based experiment was used to test the study hypotheses. Such experiments, also known as vignette-based experiments, are common in both the context of last-mile logistics consumer behavior research (Buldeo Rai et al., 2019, 2021; Kokkinou et al., 2024) and chatbot research (De Sá Siqueira et al., 2024) as they allow researchers to manipulate conditions without interfering with sensitive business processes (Kokkinou & Cranage, 2011).

An online survey was developed using the software Qualtrics. The survey consisted of four parts. After receiving information about the study procedures and providing informed consent, participants were asked to imagine themselves in the situation where they received the following message: "*Dear customer, you recently ordered a pair of shoes to be delivered to your door. However, we believe we may be able to deliver your order sooner*

and would like to ask you to work with our Chatbot, Miles, to arrange delivery at a mutually suitable time. Miles the Chatbot can explore various options with you to deliver your package at a time and place that suits you. Miles is enabled by AI technology and has been trained by our customer service," and asked if they were willing to interact with Miles the Chatbot. If participants refused, they were redirected to the third section of the survey. The second section of the survey consisted of an interaction with one of three versions of the chatbot (De Sá Siqueira et al., 2024). The chatbots were programmed to exhibit human-likeness, following the characteristics as developed by De Sá Siqueira et al. (2024).

Participants were randomly assigned to one of three conditions. Condition 1 consisted of a "Good" version of the chatbot that allowed the participant to reschedule and/or redirect delivery. Possible options included delivery at home, at work, or at a parcel locker located close to either. Condition 2 represented the Privacy concerns version. This version of the chatbot only differed from Condition 1 in that the Chatbot asked personal information from the participant (e.g., to confirm their identify, confirm delivery location, etc.). Condition 3 represented the Error Prone Chatbot. This version of the chatbot was programmed to make illogical mistakes. For example, if a respondent indicated not being at home for delivery, the chatbot would offer to arrange delivery by electric van. The three versions of the chatbot were pre-tested extensively by researchers, lecturers, and (Master) students to ensure that they exhibited the desired behavior. This section of the survey also included control measures. Participants were asked to indicate on a scale from 1 to 5 whether the chatbot that they interacted with (1) asked for personal information, (2) responded illogically, and (3) helped them achieve the objective to reschedule the delivery. Participants were also asked to rate their satisfaction with their interaction with the chatbot. Section 3 of the online survey consisted of measurement scales based on the UTAUT framework (Venkatesh et al., 2012) and adapted from previous studies (Kokkinou et al., 2025; Venkatesh et al., 2012; Zhou et al., 2020).

Participants were recruited from the student population of a Dutch mid-sized university using a census approach. Of the 4000 students approached by e-mail, 233 participated in the research, a response rate of 5.8%. 19 surveys were removed due to excessive incomplete data and/or straightlining. The final sample consisted of 214 students, of whom 37% indicated being male, 54% female, and 1% non-binary. On average, participants reported 2.047 (sd = 1.969) online monthly purchases.

5 Results

Prior to testing the study hypotheses, the reliability of the measurement scales was examined using Cronbach Alpha (Hair, 1998). For all measurement scales except Facilitating Conditions and Perceived Risk, the Cronbach Alpha exceeded 0.7. Given that Facilitating Conditions is a well-established scale in the literature, it was kept as is for further research. Perceived Risk is also a widely used scale, however for this scale only two items were retained for further research. For all measurement scales, the values of the items were averaged to compute a single score. One-way ANOVAs were used to test differences among the three chatbot conditions. Participants assigned to condition 1 reported fewer personal questions ($p<.000$) and illogical interactions ($p<.000$) than participants in the other conditions. Participants in condition 3 reported being less successful in accomplishing the objective of rescheduling delivery than participants in other conditions ($p<.045$).

Table 1: Reliability and Descriptive Statistics

	Source	Nr. Items	Alpha	Mean	St. Dev.
Intention to use Chatbot	Zhou et al. (2020)	5	0.914	2.858	1.034
Performance	Zhou et al. (2020)	3	0.840	3.382	1.051
Ease of Use	Zhou et al. (2020)	4	0.768	4.265	0.678
Social Influence	Zhou et al. (2020)	3	0.893	2.636	0.977
Facilitating Conditions	Zhou et al. (2020)	4	0.613	3.933	0.653
Perceived Risk	Zhou et al. (2020)	2 (from 4)	0.693	3.341	1.039

Source: Authors' Own Work

Table 2: Control Measures: One-Way ANOVAs

	Condition 1 (Good)	Condition 2 (Inquisitive)	Condition 3 (Error Prone)	p-value	No Chatbot
Personal	1.344	2.300	2.017	0.000	
Illogical	1.571	2.116	2.052	0.045	
Objective	4.138	4.200	3.333	0.000	
	n = 59	n = 60	n = 58		n = 36

Source: Authors' Own Work

Of the 214 participants, 36 chose not to interact with the chatbot. For these participants, there is no score for satisfaction with the chatbot. This explains the difference in sample size for models including vs. not including satisfaction as a dependent variable.

Linear regression was used to test the hypotheses of the *UTAUT model* in two steps. First, only the UTAUT variables were included. The overall regression model was statistically significant, $F(5, 208) = 22.95$, $p < .001$, and explained approximately 35.6% of the variance in the outcome variable (adjusted $R^2 = .340$). The examination of individual predictors showed that two variables made significant contributions, namely Performance ($\beta = 0.151$, $p < .000$) and Social Influence ($\beta = 0.229$, $p < .05$), supporting HA1 and HA3. Subsequently, a second model that also included satisfaction was examined. This second overall regression model was also statistically significant, $F(6, 170) = 12.07$, $p < .001$, and explained approximately 29.9% of the variance in the outcome variable (adjusted $R^2 = .274$). Satisfaction was also significant ($\beta = 0.229$, $p < .05$), providing support for HAS.

Table 3: Linear Regression Results

	Model 1		Model 2		Hyp.
	Estimate	p-value	Estimate	p-value	
Intercept	-0.110	0.825	-0.040	0.947	
Performance	0.294	**0.000**	0.200	**0.017**	HA1 Sup.
Ease of Use	0.151	0.133	0.071	0.506	
Social Influence	0.229	**0.002**	0.248	**0.002**	HA3 Sup.
Facilitating Conditions	0.207	0.084	0.199	0.109	
Perceived Risk	-0.026	0.664	-0.041	0.528	
Satisfaction			0.194	**0.004**	HAS Sup.

Notes: Model 1: R2 0.356, Adj. R2 0.340; Model 2: R2 0.299 Adj. R2 0.274

Source: Authors' Own Work

Linear regression was also used to test the *effect of the interventions* on intention to use the chatbot. The overall regression model was not statistically significant, $F(2, 174) = 2.231$, $p = .110$, only explaining 2.5% of the intention to use the chatbot.

A *mediation analysis* was conducted to examine whether satisfaction mediates the relationship between the Error Prone intervention and future intention to use Chatbots for last-mile delivery. The analysis was performed using Hayes' process function in R. A bootstrapping procedure with bias-corrected confidence intervals

based on 5,000 bootstrap samples was employed. The total effect of the Error Prone intervention on satisfaction was statistically significant (b = -0.41, SE = 0.15, t = -2.71, p < .01, 95% CI [-0.70, -0.11]), indicating that this intervention negatively impacted satisfaction. This intervention accounted for 4.0% of the variance in satisfaction, F(1, 175) = 7.34, p <.01. In the model predicting future intention to use AI, both the Error Prone intervention and satisfaction were included as predictors. This model explained 7.5% of the variance in future intention to use AI, F(2, 174) = 7.019, p < .001. Within this model, satisfaction was a significant positive predictor of AI usage (b = 0.220, SE = 0.072, t = 3.059, p < .005, 95% CI [0.078, 0.362]). However, the direct effect of the intervention on AI usage was not statistically significant (b = -0.219, SE = 0.146, t = -1.506, p = .134, 95% CI [-0.507, 0.068]).

The indirect effect of the intervention on intention to use the chatbot through satisfaction was statistically significant (b = -0.089, SE = 0.049, 95% CI [-0.200, -0.011]), indicating that satisfaction mediated the relationship between this intervention and future intention to use a chatbot. This means that when the chatbot made more mistakes, users were less satisfied, which in turn was associated with lower intentions to use chatbots in the future. These findings support a mediation model in which satisfaction serves as a mechanism through which an Error Prone chatbot negatively influences future intention to use AI, supporting H_{MS}.

6 Discussion

The purpose of this study was to examine the factors that drive consumer use of Chatbots in the context of last-mile logistics. The results of the extended UTAUT model showed that expected performance and social influence were positively associated with parcel recipients' future intention to use chatbots for rescheduling deliveries. Satisfaction with a past chatbot interaction was also a significant predictor, supporting previous research that found that positive experiences carry forward into future behavioral intentions (Ashfaq et al., 2020). The experimental manipulations did not produce a statistically significant direct effect on future intention to use. However, the mediation analysis revealed that the error-prone chatbot negatively affected satisfaction, which in turn reduced future intention. This suggests that, rather than directly impacting intention, chatbot errors influence intention through satisfaction (Bhattacherjee, 2001), providing nuance to our understanding of how chatbot quality influences long-term usage.

In terms of *implications for theory*, the findings confirm the partial applicability of the UTAUT framework to this novel context (Alsarayreh et al., 2025). Consistent with prior research in retail and service settings (Arce-Urriza et al., 2025; Liu et al., 2024), expected performance emerged as a robust predictor of intention to use chatbots to reschedule deliveries. This implies that parcel recipients are principally motivated by whether the chatbot can help them arrange delivery more conveniently and reliably. This was inconsistent with Alsarayreh et al. (2025)'s results, although this could have been due to their inclusion of the effect of habit, which may have been confounding. Conversely, and similar to Alsarayreh et al. (2025), our findings showed that perceived ease of use was not a statistically significant predictor. This suggests that, in a task-oriented, logistical interaction, usability concerns may be subordinate to outcome expectations. This is consistent with the utilitarian framing of this context (Arce-Urriza et al., 2025) and diverges from findings in hospitality and retail (Liu et al., 2024; Pillai & Sivathanu, 2020). Consistent with ECM (Bhattacherjee, 2001), satisfaction significantly predicted future intention to use chatbots for rescheduling delivery. It also served as the pathway through which chatbot errors indirectly reduced future intention, despite the direct effect of the error-prone condition being non-significant. This reinforces the need to integrate affective and evaluative constructs into predominantly cognitive models such as UTAUT when studying interactive AI technologies (Arce-Urriza et al., 2025; Chih et al., 2025). The absence of a direct effect of chatbot condition on intention is interesting and merits further research. It could imply that a single interaction with a flawed or intrusive chatbot may not be sufficient to influence behavioral intentions, especially in a utilitarian context and/or users may attribute errors to novelty or technical limitations. (De Sá Siqueira et al., 2024).

In terms of *implications for practitioners*, the significant effect of expected performance suggests that LSPs and chatbot developers should prioritize functionality. A chatbot that reliably helps recipients reschedule or redirect their parcel delivery will drive adoption more effectively than one that is kept too simple. This implies that chatbot development should focus on providing a comprehensive set of options, including delivery times, location, and alternative delivery options such as parcel lockers. Furthermore, the significance of social influence suggests that LSPs can benefit from social comparison mechanisms, for example, by communicating usage rates or customer ratings, which may encourage hesitant recipients to try the chatbot. The mediation findings highlight the importance of satisfaction. Even if chatbot errors

do not directly affect parcel recipients' future usage intentions, they do erode satisfaction. This can potentially erode recipients' usage intentions over time if these errors are not addressed. LSPs and chatbot developers should therefore invest in extensive testing, iterative improvement, and swift referral to human agents when needed. Even though privacy-invasive chatbot behavior was not shown to directly reduce intention in this study, it remains a reputational risk and should be minimized through transparent data practices and minimal data collection.

Several limitations need to be acknowledged. First, data were collected from a student sample at a single university, in a single country, limiting generalizability. Students tend to be younger and more digitally literate than the general population (Kokkinou et al., 2025; Zhou et al., 2020). The non-significance of facilitating conditions and perceived risk may also reflect students' digital fluency. Future research should therefore use more diverse, representative samples. Second, although using live chatbots in experimental chatbot interactions increased ecological validity, these systems remain dynamic and unpredictable (De Sá Siqueira et al., 2024). As such, it is not entirely possible for the researchers to maintain tight control over experimental conditions, as would be desirable in a scenario-based experiment (Kokkinou & Cranage, 2011). The control measures demonstrated that, in general, the manipulations were successful. However, the study could be replicated either by having participants read a simulated interaction or by interacting with the chatbot in the presence of researchers as part of a think-aloud qualitative research design. Additionally, the privacy manipulation, while successfully verified by the control measures, did not produce the expected effects on intention to use. Future research might explore more salient privacy violations.

Acknowledgments

This study is part of the project From Click to Clean: Consumer-Driven Sustainable Last-mile Logistics (LOG.KIEM.02.009) funded by Regieorgaan SiA and supported by the SiA SPRONGgroep LILS (Low Impact Lastmile logisticS SPR.ALG.03.006).

References

Alharbi, N., Ud Din, F., Paul, D., & Sadgrove, E. (2025). Driving AI chatbot adoption: A systematic review of factors, barriers, and future research directions. *Journal of Open Innovation: Technology, Market, and Complexity*, *11*(3), 100590. https://doi.org/10.1016/j.joitmc.2025.100590

Alsarayreh, S., Al-Emran, M., Arpaci, I., & Shaalan, K. (2025). Promoting environmental sustainability through chatbot use in shipment delivery: An integrated model of UTAUT2

and TTF. *Asia-Pacific Journal of Business Administration*, 1–25. https://doi.org/10.1108/APJBA-10-2024-0611

Amaya, J., Encarnación, T., & Cantillo, V. (2025). Sustainable last mile delivery alternatives: Influencing factors and willingness to use. *Transportation Research Part D: Transport and Environment*, *139*, 104574. https://doi.org/10.1016/j.trd.2024.104574

Arce-Urriza, M., Chocarro, R., Cortiñas, M., & Marcos-Matás, G. (2025). From familiarity to acceptance: The impact of Generative Artificial Intelligence on consumer adoption of retail chatbots. *Journal of Retailing and Consumer Services*, *84*, 104234. https://doi.org/10.1016/j.jretconser.2025.104234

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, *54*, 101473. https://doi.org/10.1016/j.tele.2020.101473

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model1. *MIS Quarterly*, *25*(3), 351–370.

Buldeo Rai, H., Broekaert, C., Verlinde, S., & Macharis, C. (2021). Sharing is caring: How non-financial incentives drive sustainable e-commerce delivery. *Transportation Research Part D: Transport and Environment*, *93*, 102794. https://doi.org/10.1016/j.trd.2021.102794

Buldeo Rai, H., Verlinde, S., & Macharis, C. (2019). The "next day, free delivery" myth unravelled: Possibilities for sustainable last mile transport in an omnichannel environment. *International Journal of Retail & Distribution Management*, *47*(1), 39–54. https://doi.org/10.1108/IJRDM-06-2018-0104

Chih, W.-H., Wang, K.-Y., & Banda, H. W. (2025). Chatbots at the frontline: Unveiling antecedents of customers' willingness to accept chatbot intervention in service recovery. *Journal of Retailing and Consumer Services*, *84*, 104254. https://doi.org/10.1016/j.jretconser.2025.104254

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003.

De Sá Siqueira, M. A., Müller, B. C. N., & Bosse, T. (2024). When Do We Accept Mistakes from Chatbots? The Impact of Human-Like Communication on User Experience in Chatbots That Make Mistakes. *International Journal of Human–Computer Interaction*, *40*(11), 2862–2872. https://doi.org/10.1080/10447318.2023.2175158

DHL. (n.d.). *Bezorging wijzingen*. Retrieved February 27, 2026, from https://www.dhlecommerce.nl/nl/consument/bezorging-wijzigen

Gevaers, R., Van de Voorde, E., & Vanelslander, T. (2014). Cost Modelling and Simulation of Last-mile Characteristics in an Innovative B2C Supply Chain Environment with Implications on Urban Areas and Cities. *Procedia - Social and Behavioral Sciences*, *125*, 398–411. https://doi.org/10.1016/j.sbspro.2014.01.1483

Hair, J. F. (Ed.). (1998). *Multivariate data analysis*. Prentice Hall.

Kasilingam, D. L. (2020). Understanding the attitude and intention to use smartphone chatbots for shopping. *Technology in Society*, *62*, 101280. https://doi.org/10.1016/j.techsoc.2020.101280

King, H. (2024, July 12). *Parcel redirections highlight demand for flexibility, DHL reveals*. Parcel and Postal Technology International. https://www.parcelandpostaltechnologyinternational.com/news/parcels/parcel-redirections-highlight-demand-for-flexibility-dhl-reveals.html

Kokkinou, A., & Cranage, D. A. (2011). Modeling human behavior in customer-based processes: The use of scenario-based surveys. *Proceedings of the 2011 Winter Simulation Conference (WSC)*, 683–689. https://doi.org/10.1109/WSC.2011.6147796

Kokkinou, A., Lastauskaitė, A., Quak, H., Karaoulanis, A., Mitas, O., & Mandemakers, A. (2025). Unlocking sustainable last mile deliveries through interventions in the pre-delivery stage. *Cleaner Logistics and Supply Chain*, 100250. https://doi.org/10.1016/j.clscn.2025.100250

Kokkinou, A., & Quak, H. (2026). *Gemak dient de mens (behalve als het niet werkt): Een kijkje in de digitale klaagmuur voor pakketautomaten* (B. Vannieuwenhuyse & S. Weijers, Eds.). https://vervoerslogistiekewerkdagen.com/wp-content/uploads/2026/03/GEMAK-DIENT-DE-MENS-BEHALVE-ALS-HET-NIET-WERKT.pdf

Kokkinou, A., Quak, H., & Mitas, O. (2026). Leveraging fairness to nudge consumers towards more sustainable last mile delivery options. *Transportation Research Interdisciplinary Perspectives*, *37*, 101956. https://doi.org/10.1016/j.trip.2026.101956

Kokkinou, A., Quak, H., Mitas, O., & Mandemakers, A. (2024). Should I wait or should I go? Encouraging customers to make the more sustainable delivery choice. *Research in Transportation Economics*, *103*, 101388. https://doi.org/10.1016/j.retrec.2023.101388

Liu, M., Yang, Y., Ren, Y., Jia, Y., Ma, H., Luo, J., Fang, S., Qi, M., & Zhang, L. (2024). What influences consumer AI chatbot use intention? An application of the extended technology acceptance model. *Journal of Hospitality and Tourism Technology*, *15*(4), 667–689. https://doi.org/10.1108/JHTT-03-2023-0057

Mangiaracina, R., Marchet, G., Perotti, S., & Tumino, A. (2015). A review of the environmental implications of B2C e-commerce: A logistics perspective. *International Journal of Physical Distribution & Logistics Management*, *45*(6), 565–591. https://doi.org/10.1108/IJPDLM-06-2014-0133

Mangiaracina, R., Perego, A., Seghezzi, A., & Tumino, A. (2019). Innovative solutions to increase last-mile delivery efficiency in B2C e-commerce: A literature review. *International Journal of Physical Distribution & Logistics Management*, *49*(9), 901–920. https://doi.org/10.1108/IJPDLM-02-2019-0048

Niemeijer, R., & Buijs, P. (2023). A Greener Last Mile: Reviewing the Carbon Emission Impact of Pickup Points in Last-Mile Parcel Delivery. *Renewable and Sustainable Energy Reviews*, *186*, 113630.

Nogueira, G. P. M., Silva, S. V., Hasenclever, L., & De Assis Rangel, J. J. (2024). Sustainability in B2C e-commerce: A literature review and future research opportunities in last-mile deliveries. *Environment, Development and Sustainability*, *27*(6), 12415–12436. https://doi.org/10.1007/s10668-024-04471-2

Norman, H. (2025, December). Delivering Control. *Parcel and Postal Technology International.* https://parcelandpost.mydigitalpublication.com/december-2025

Parasuraman, A., & Colby, C. L. (2015). An Updated and Streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research*, *18*(1), 59–74. https://doi.org/10.1177/1094670514539730

Pillai, R., & Sivathanu, B. (2020). Adoption of AI-based chatbots for hospitality and tourism. *International Journal of Contemporary Hospitality Management*, *32*(10), 3199–3226. https://doi.org/10.1108/IJCHM-04-2020-0259

PostNL. (n.d.). *Bezorgvoorkeuren*. Retrieved February 27, 2026, from https://www.postnl.nl/ontvangen/pakket-ontvangen/bezorging-pakketten/bezorgvoorkeuren/

Ragni, M., Rudenko, A., Kuhnert, B., & Arras, K. O. (2016). Errare humanum est: Erroneous robots in human-robot interaction. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 501–506. https://doi.org/10.1109/ROMAN.2016.7745164

Rese, A., Ganster, L., & Baier, D. (2020). Chatbots in retailers' customer communication: How to measure their acceptance? *Journal of Retailing and Consumer Services*, *56*, 102176. https://doi.org/10.1016/j.jretconser.2020.102176

Richey, R. G., Chowdhury, S., Davis-Sramek, B., Giannakis, M., & Dwivedi, Y. K. (2023). Artificial intelligence in logistics and supply chain management: A primer and roadmap for research. *Journal of Business Logistics*, *44*(4), 532–549. https://doi.org/10.1111/jbl.12364

Savelsbergh, M., & Van Woensel, T. (2016). 50th Anniversary Invited Article—City Logistics: Challenges and Opportunities. *Transportation Science*, *50*(2), 579–590. https://doi.org/10.1287/trsc.2016.0675

Seegers, S. (2025, February 13). AI Innovations in Last-Mile Delivery: Driving Efficiency by Elevating Customer Satisfaction. *ParcelIndustry.Com*. https://parcelindustry.com/article-6434-AI-Innovations-in-Last-Mile-Delivery-Driving-Efficiency-by-Elevating-Customer-Satisfaction.html

Seghezzi, A., & Mangiaracina, R. (2023). Smart home devices and B2C e-commerce: A way to reduce failed deliveries. *Industrial Management & Data Systems*, *123*(5), 1624–1645. https://doi.org/10.1108/IMDS-10-2022-0651

Toader, D.-C., Boca, G., Toader, R., Măcelaru, M., Toader, C., Ighian, D., & Rădulescu, A. T. (2019). The Effect of Social Presence and Chatbot Errors on Trust. *Sustainability*, *12*(1), 256. https://doi.org/10.3390/su12010256

Vakulenko, Y., Shams, P., Hellström, D., & Hjort, K. (2019). Service innovation in e-commerce last mile delivery: Mapping the e-customer journey. *Journal of Business Research*, *101*, 461–468. https://doi.org/10.1016/j.jbusres.2019.01.016

Venkatesh, Thong, & Xu. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, *36*(1), 157. https://doi.org/10.2307/41410412

Zhou, M., Zhao, L., Kong, N., Campy, K. S., Xu, G., Zhu, G., Cao, X., & Wang, S. (2020). Understanding consumers' behavior to adopt self-service parcel services for last-mile delivery. *Journal of Retailing and Consumer Services*, *52*, 101911. https://doi.org/10.1016/j.jretconser.2019.101911

Developing People-centric Data Governance with Service Design

Jaana Helsing, Tanja Korhonen
Kajaani University of Applied Sciences, Kajaani, Finland
jaana.helsing@gmail.com, tanja.korhonen@kamk.fi

People are in the center when creating Data Governance practices. Data Management can be developed into a user-centric using Service Design principles to address challenges in Data Governance, Data Quality, and data interoperability. This article emphasises the importance of placing people at the center of Data Governance. Through literature review, survey, and co-creation workshops, a service concept was designed to foster collaboration and shared ownership. Findings reveal gaps in Data Governance and Data Management practices, highlighting the need for structured Data Management frameworks and human-centred approaches. The paper contributes to theory and practice by demonstrating how Service Design supports cultural adoption and co-creation in Data Management.

DOI
https://doi.org/10.18690/um.fov.4.2026.22

ISBN
978-961-299-152-4

Keywords:
data management,
data governance,
stakeholder engagement,
human-centered design,
service design

University of Maribor Press

1 Introduction

Data has emerged as a critical strategic asset for organizations seeking to drive innovation, optimize operations, and enhance decision-making capabilities (Wallis, 2021). In contemporary digital transformation initiatives, effective data management is no longer merely a technical function but a strategic enabler that demands cultural change and sustained human engagement (Pedrosa & Araujo Freitas, 2024). Organizations increasingly recognize that robust data governance frameworks are essential for ensuring data quality, interoperability, and trustworthiness: prerequisites for advanced analytics, automation, and evidence-based decision-making.

Despite significant investments in data platforms, governance frameworks, and technical infrastructure, many organizations struggle to realize the full value of their data assets. A persistent gap exists between the availability of sophisticated data management frameworks such as DAMA-DMBOK and their practical adoption in organizational contexts. (Earley & Henderson, 2017) Research reveals that this implementation gap stems primarily from human and cultural factors rather than technical limitations: unclear ownership and accountability, fragmented cross-functional collaboration, resistance to change, and governance processes that fail to align with everyday work practices (Erander & Granat, 2024; Mazzarella, 2024).

In today's digital transformation, Data Management is not just a technical function, it is a strategic enabler that requires cultural change and human involvement. Organisations often struggle with fragmented practices and unclear ownership. (Justy et al., 2023.) They face cultural barriers that prevent them from fully using their data. These barriers include difficulty in moving away from a status quo mindset. Long-standing workplace cultures and fixed ways of working make it less likely for employees to use or share data. (Chakravorty, 2020.) Cultural transformation is critical: Data Governance must be embedded into everyday work, not imposed as bureaucracy.

Service design offers a promising pathway to bridge this gap by placing people at the center of governance development. Service design principles, including co-creation, user journey mapping, prototyping, and ecosystem thinking provide actionable methods for translating abstract governance policies into tangible, user-friendly touchpoints (Figueiredo et al., 2024; Mariani, 2024). By reframing data governance

as a service rather than a compliance mandate, organizations can design governance processes that align with user workflows, foster collaboration, and build shared ownership across business units and functional domains.

This paper addresses the question: *How can service design principles be integrated into data governance development to create people-centric practices that enhance adoption, accountability, and organizational value?* Drawing on a constructive research approach, we present a case study from an energy sector organization where we applied service design methods: literature review, stakeholder surveys, and co-creation workshops, to develop a user-centric data governance operating model. Our research makes three primary contributions: (1) a synthesized framework linking service design interventions to data governance capabilities, (2) empirical evidence of organizational barriers and how co-creative methods address them, and (3) practical guidance and artifacts for embedding people-centric governance into organizational practice.

The remainder of this paper is structured as follows. Section 2 reviews foundational concepts in data management and data governance, emphasizing the human dimension. Section 3 describes our research methodology, including survey design and co-creation workshop approaches. Section 4 presents empirical findings from the current state assessment, survey results, and workshop insights. Section 5 discusses the target state design and service concept development. Section 6 concludes with implications for research and practice, along with recommendations for operationalizing people-centric data governance in institutional settings.

2 Foundations and Development of Data Management

In contemporary organizations, data is widely recognised as a strategic asset that enables evidence-based decision-making, operational efficiency, innovation, and regulatory compliance. Yet, the realisation of this value depends on disciplined Data Management practices that span Data Governance, Data Architecture, Data Modeling, Data Quality, Metadata, Data Security, and lifecycle management (Earley & Henderson, 2017). Previous studies frame Conceptual foundations for understanding data and its management, and a practical, strategy-aligned approach for developing Data Management capabilities. Organisations must pair robust Data Governance with user-centric design and product thinking to achieve sustainable impact from data (DalleMule & Davenport, 2017; Wallis, 2021). Recent enterprise

frameworks emphasize that data governance functions defining clear policies, roles, and data quality standards are central to ensuring data integrity and regulatory alignment (Madhavan, 2024). Governance adoption improves when capabilities are explicitly defined, empirically validated and measured, rather than treated as abstract principles, reinforcing the case for operational, people-centred design of governance services (Merkus et al., 2023). Recent systematic analyses of governance frameworks emphasize the importance of clearly defined roles, decision rights, and maturity models for organizational data governance success (Acev et al., 2025).

2.1 Data Management

Data Management encompasses planning, execution, and oversight across the data lifecycle, ensuring control, protection, and value realization (Earley & Henderson, 2017). The DAMA-DMBOK wheel (see Fig. 1) centers Data Governance and links Data Architecture, Data Modeling, Data integration/interoperability, Data Quality, Metadata, storage/operations, Data Security, Master/Reference data, and analytics/BI as mutually reinforcing capabilities (Earley & Henderson, 2017).

Figure 1: Knowledge Areas of Data Management (Adapted from Earley & Henderson, 2017)

Data Governance defines decision rights, standards, accountability, and escalation mechanisms for data across the enterprise (Abraham et al., 2019; Khatri & Brown, 2010). Critical success factors include executive sponsorship, cross-divisional

participation, defined ownership/stewardship, measurable quality metrics, and ongoing training (Cheong & Chang, 1005, 2007). Operating models can be centralised, decentralised, federated, hybrid, or networked; the "best fit" depends on business strategy and the offense–defense balance (DalleMule & Davenport, 2017; Earley & Henderson, 2017). Recent work shows that co-eliciting governance requirements with stakeholders, covering usability, compliance, lineage, access, and organisational design, helps overcome adoption barriers in multi-stakeholder settings (Haberl, Malin & Thalmann, 2022). Organisations can operationalise this using lightweight service-design artefacts, such as the Data Product Canvas and Data Service Cards, to embed policy, quality and accountability in daily delivery (Fruhwirth et al., 2020; Breitfuss et al., 2020).

Data does not manage itself; people do. Successful adoption requires change management, clear role definitions, and communities of practice to embed new behaviors (Earley & Henderson, 2017; Kosonen, 2019). This human dimension is often underestimated but is decisive for governance maturity and the scale-up of Data Quality and Metadata practices (Heisig, 2009).

2.2 Developing Data Management: Data Strategy and Data Governance

A Data Strategy clarifies data's purpose, availability, organisation, use, and capabilities, aligning defense (risk, compliance, integrity) with offense (growth, innovation, customer value) (DalleMule & Davenport, 2017; Datastrategiaopas, 2021; Earley & Henderson, 2017). It sets vision, principles, objectives, and a roadmap, typically supported by a dedicated Data Management Strategy covering governance, architecture, quality, metadata, master/reference, security, and analytics (Wallis, 2021; Earley & Henderson, 2017).

Knowledge or Data Management efforts fail when culture, measurement, strategy, structure, leadership, IT problems, and lack of standards are not addressed holistically (Liebowitz, 2016). Development must integrate social and organisational factors, not just technology, to improve efficiency (e.g., faster data discovery, fewer manual reconciliations) and reduce risk (Kaario & Peltola, 2008). Unified governance frameworks that integrate cultural change considerations, process alignment, and

stakeholder engagement are essential for organizational adoption, particularly in highly regulated sectors (Dibouliya, 2025).

EU-level instruments such as GDPR, Data Governance Act, Digital Services Act, Digital Markets Act, Data Act, and the AI Act shape expectations for data transparency, portability, fairness, and safety (Traficom, n.d.; Office of the Data Protection Ombudsman, n.d.). Industry-specific regimes (e.g., DORA, PCI-DSS, Solvency II, BCBS 239) further reinforce governance requirements (Finanssivalvonta, n.d.; Earley & Henderson, 2017). The energy sector operates under an extensive regulatory framework encompassing EU internal energy market rules, climate and decarbonisation legislation, security-of-supply requirements, and critical-infrastructure cybersecurity and data regulations, making accountability and Data Management central compliance enablers rather than purely operational concerns (Zhang et al., 2025). Embedding privacy, retention, and access controls in the development lifecycle is non-negotiable.

To translate governance into day-to-day practice, a layered operating model is needed, with each level carrying distinct responsibilities. At the Executive level, leadership sets vision, policy and priorities. At the Strategic Common level, a Data Governance Board establishes shared standards, taxonomies and training. The Tactical level manages subject areas, definitions, stewardship, quality rules and the data lifecycle, while the Operational level applies standards in projects, documents metadata and lineage, monitors quality indicators and manages access. Together, these layers clarify decision rights, collaboration points and stewardship practices, which strengthens adoption. (Cheong & Chang, 2007; DalleMule & Davenport, 2017; Earley & Henderson, 2017)

Effective Data Management produces timely, relevant, and trusted information at lower total cost of ownership and greater speed, thus supporting both defensive and offensive objectives (Hawtin & Lecore, 2011; DalleMule & Davenport, 2017). Quantifiable benefits include reduced manual work, faster access to data, improved compliance posture, stronger collaboration/literacy, and better ROI on data initiatives (Hawtin & Lecore, 2011; Cheong & Chang, 1006, 2007). Without governance, data risks such as inefficiency, quality failures, compliance breaches,

outages, and missed opportunities, compound (Earley & Henderson, 2017; Redman, 1998).

3 Research Methodology

The study employed a constructive research approach, which is well-suited for solving practical problems through the creation of new knowledge and artifacts (Lukka, 2003; Piirainen & Gonzalez, 2013). The goal was to design a user-centric Data Management service for an energy sector company by combining theoretical insights with empirical evidence. This approach ensured that the resulting concept was both academically grounded and operationally relevant. The original reseach data was gathered as part of thesis (Helsing, 2025).

The methodology integrated qualitative and quantitative methods within a mixed-method framework, structured into three phases:

Phase 1: Literature Review. A systematic review of academic and industry sources established the theoretical foundation for Data Management and Service Design. Frameworks such as DAMA-DMBOK and service-dominant logic (Vargo, Maglio, Akaka, 2008), informed the conceptual model.

Phase 2: Empirical Data Collection

- Survey: Distributed to 40–50 participants across business units, IT, and enabling functions
- The survey captured perceptions of current Data Management maturity and improvement priorities. Responses (n=16) were analyzed using descriptive statistics to identify gaps in areas such as Data Quality, Metadata Management, and Data Governance.
- Co-Creation Workshops: Two workshops applied the Six Thinking Hats (De Bono, 1985) method to generate ideas and validate needs. These sessions facilitated cross-functional dialogue, enabling participants to define Data Governance roles, articulate pain points, and envision target-state capabilities.

Phase 3: Service Design Application. Service design principles (Stickdorn & Schneider, 2011) guided the development of the Data Management service concept. Tools such as personas and customer journey mapping represented user needs and designed touchpoints. The Double Diamond process (Design Council, n.d.) structured iterative prototyping and refinement with phases of Discover, Define, Develop and Deliver.

Data Collection and Analysis

Survey data was analysed through descriptive statistics (frequencies, percentages) that summarised maturity levels and priority areas. Visualisations such as bar charts highlighted patterns in responses. Workshop data was analysed using content analysis clustered insights under governance, quality, and usability themes. Outputs informed the design of the Data Governance Operating Model and service components. Findings were synthesised into a target operating model, a roadmap, and measurable objectives using the OKR (Objectives and Key Results) framework.

This methodology ensured alignment between theory and practice, combining evidence-based design with stakeholder co-creation. By embedding service design into Data Management development, the approach delivered a solution that is both technically robust and user-centric.

4 Findings

4.1 Current State of Data Management

The empirical work (survey and workshops) surfaced a consistent set of pain points. Respondents described fragmented practices, unclear ownership, and manual workarounds that undermine trust in data and slow delivery. Key pain points:

- Unclear ownership and accountability: Roles are undefined and stakeholders "don't know whom to contact".
- Data Quality issues: Inaccurate and inconsistent data hinders automation, forecasting, optimisation, and investment planning.
- Metadata and lineage gaps: Context, definitions, and lineage difficult to find.

- Manual work and rework: Repeated data entry, ad-hoc access controls, and governance drop-off after projects.
- Limited cross-unit collaboration: Collaboration exists but is insufficient.

These findings echo the literature's view that Data governance and Data Quality management are cross-divisional and require sustained leadership, standards, and training (Cheong & Chang, 2007; Earley & Henderson, 2017).

4.2 Survey results

The survey (15.5.2025–16.6.2025) reached 40–50 employees and received 16 responses, representing Business (37.5%), Business IT (25%), Enabling functions (18.7%), and Group IT (18.8%). Data challenges were widespread:

- n=15 reported experiencing Data Quality issues several times in the past year; n=0 reported "not at all," confirming persistence and business impact (Redman, 1998; Loshin, 2020).
- Only n=1 model data systematically, while n=9 recognise the need, but lack processes or resources.
- n=10 say they have "some metadata" but no processes or guidelines; n=4 report having none.
- n=11 have only partially identified processes/owners
- Governance practices are emerging: n=10 have "started" practices, n=5 lack both ownership and management practices, and n=1 reports clear ownership and management.

Respondents highlighted need for group-level guidance that is described in table 1.

Urgency ranking (lower average = higher priority): Data Governance (avg 2.7), Data Modeling (3.3), Data Quality (3.4), Data Architecture (3.7). This aligns with DAMA's centrality of governance and foundational role of architecture and modeling (Earley & Henderson, 2017).

Table 1: Need for group-level support

Area	Need for Group-level Support
Data Architecture	66.7%
Metadata Management	60.0%
Data Platforms	60.0%
Data Governance	53.3%
Data Modeling	53.3%
Master / Reference Data	53.3%
Data Quality	40.0%
Data Security	40.0%
Data Flows / Integrations	40.0%
Document / Content Management	40.0%
Data Product Development	33.3%

4.3 Workshop insights

Two workshops were conducted: (1) a Data Governance Working Group session (7 participants) and (2) a Six Thinking Hats session (3 participants). Data Governance Working Group defined itself as a strategic enabler: building a shared language, clarifying roles, reducing duplication, and accelerating data-driven initiatives. 12-Month objectives include establish Data Subject Areas and Data Families, define taxonomy, standards, operationalise governance in real projects, and embed privacy and Data Quality. Value Proposition comes from efficiency, risk mitigation, transparency, collaboration, acceleration, and maturity. These outputs reflect critical success factors from Data Governance literature (Cheong & Chang, 2007; Khatri & Brown, 2010; Abraham et al., 2019).

The synthesis of the Six Hats Workshop was that with white (facts) hat, there are manual processes, lack of collaboration and need for Data Architecture support. Red (feelings) hat showed frustration when Data Governance is not a continuous effort and desire for visible and practical results. The yellow (benefits) hat revealed benefits with trusted, accessible, visible data as well as automation, reusable solutions, clearer governance. Black (risks) hat perceived bureaucracy and unclear value realisation and the Green (ideas) hat embraced communities, business-led roles and reusable data assets. Overall, participants asked for tangible artifacts (playbooks, templates, catalogs, lineage), business-led ownership, and community practices to sustain change, consistent with service design's co-creation methods and Double Diamond process (Design Council, n.d.; Koivisto et al., 2019).

4.4 Target State and Service Concept

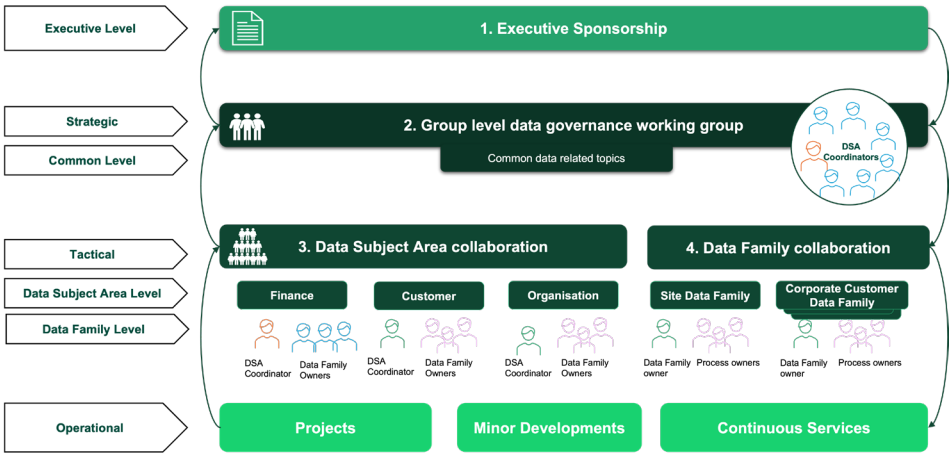

Figure 2: Data Governance Operating Model

Table 2: Design rationale for Data Governance Operating Model

Model element	Empirical finding	Design principle	Resulting design decision
Four-layer operating model	Governance activities dropped after project delivery; lack of continuity	Governance must span strategic, tactical, and operational levels	Operating model structured into four distinct responsibility levels
Data Family Owner role	Survey and workshops revealed unclear data ownership	Accountability should be business-led, not IT-driven	Dedicated business role accountable for critical data
Data Subject Area Coordinator	Fragmented collaboration across data families	Federated coordination improves cross-alignment	Coordinator role introduced at tactical level
Limited number of roles	Resistance to bureaucracy; preference for practical governance	Lightweight governance increases adoption	Role set kept intentionally minimal
Absence of a formal steering group in initial model	Workshop participants expressed concern over bureaucracy and low perceived value of governance forums	Lightweight governance structures improve adoption in low-maturity contexts	Steering group intentionally excluded from initial model; executive sponsorship used instead
Executive sponsorship without steering group	Need for prioritisation and escalation without adding decision layers	Clear sponsorship can provide direction while minimizing overhead	Named sponsor established to support authority and visibility

Using survey and workshop evidence, the study defined a vision and mission for Data Management. Vision is to *be a trusted enabler of data-driven success: data is discoverable, reliable, and readily available*. And mission is *to empower business units with user-centric Data Governance and architecture services, improving data quality, metadata, and master/reference data in collaboration with data family owners and practitioners.*

A Data Governance Operating Model was formed to support decision-making and collaboration around data. This clarifies decision rights and execution pathways, addressing the survey's ownership and collaboration gaps (Earley & Henderson, 2017; DalleMule & Davenport, 2017). The Data Governance Operating Model is presented in Figure 2, and its design rationale is provided in Table 2.

5 Discussion and Conclusions

Outcomes achieved through the study were shared understanding of current maturity and gaps, defined governance operating model and roles with a 12-month agenda and user-centric service concept prototype with personas, journeys, and component design. Implications related to embedding governance in delivery lifecycles (projects/products), making metadata/lineage visible, and productizing dat**a** are pivotal to shift from ad-hoc practices to trusted, reusable, and discoverable data. Without these steps, quality failures and duplication persist (Redman, 1998; Loshin, 2020); with them, data becomes an actionable asset (Nokkala, 2020; DalleMule & Davenport, 2017).

The findings of this study underscore a critical insight: Data Management is not solely a technical challenge, but it is fundamentally a human one. While frameworks, tools, and governance models provide structure, their success depends on the engagement, understanding, and collaboration of people across the organisation. Survey results revealed that gaps in ownership, unclear roles, and fragmented collaboration were among the most pressing issues. These are not technological failures; they are organisational and cultural challenges. Workshops reinforced this perspective. Participants expressed frustration when governance efforts ended after project go-lives and voiced a strong desire for visible, practical results rather than abstract frameworks. Service design methods proved invaluable in translating technical concepts into human-centered solutions.

The research demonstrates that putting people first is the cornerstone of successful Data Management. Organisations often invest heavily in platforms and tools, assuming that technology alone will solve data challenges. However, without clear roles, accountability, and cultural adoption, even the most advanced systems fail to deliver value. Data does not manage itself—people do. Ultimately, the path to trusted, discoverable, and actionable data runs through people. Prioritising human factors alongside technical solutions is essential to unlocking the full potential of data assets and driving innovation, compliance, and competitive advantage.

6 Recommendations and Limitations

Data Governance should be integrated into project and product lifecycles rather than treated as a one-off activity. This approach ensures continuity and prevents governance from "dropping off" after go-live. Organisations should leverage Service Design principles to make Data Governance collaborative and user-friendly. Tangible artifacts such as playbooks, templates, and data catalogs should be developed to support adoption and usability. Implementing a layered operating model with clearly defined roles, addresses ownership gaps and accelerates adoption. Persistent quality issues and metadata gaps identified in the survey highlight the need for systematic modeling, lineage documentation, and quality monitoring. Resources should be allocated to these areas. Encourage cross-functional collaboration through communities and business-led roles to sustain cultural change and embed governance into everyday work.

The survey included 16 respondents, which may not fully represent the entire organisation or industry. Findings should be interpreted as indicative rather than exhaustive. The study focused on an energy sector company; recommendations may require adaptation for other industries with different regulatory and operational contexts. While service design offers a structured approach, translating concepts into practice requires sustained leadership commitment and resources, which were not tested in full-scale implementation. The research emphasised cultural and organisational factors; technology-related challenges (e.g., platform limitations, integration complexity) were outside the study's scope.

References

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data Governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management, 49, 424–438.

Acev, D., Biyani, S., Rieder, F., Aldenhoff, T. T., & Blazevic, M. (2025). Systematic analysis of data governance frameworks and their relevance to data trusts. Management Review Quarterly. https://doi.org/10.1007/s11301-025-00545-1

Breitfuss, G., Fruhwirth, M., Wolf-Brenner, C., Riedl, A., de Reuver, M., Ginthoer, R., & Pimas, O. (2020). Data Service Cards – A supporting tool for data-driven business. In 33rd Bled eConference Proceedings, 599-614. University of Maribor Press.

Chakravorty, R. (2020). Common challenges of data governance. Journal of Securities Operations & Custody, 13(1), 23–43.

Cheong, L. K., & Chang, V. (2007). The Need for Data Governance: A case study. In Proceedings of the 18th Australasian Conference on Information Systems (ACIS 2007). Toowoomba, Australia: University of Southern Queensland, 1001-1006.

DalleMule, L., & Davenport, T. H. (2017). What's Your Data Strategy? Harvard Business Review, 112-121.

Datastrategiaopas (2021). Datastrategia – Miksi ja miten organisaationi tulisi laatia sellainen. Innolink Group Oy. Available 22.2.2025: https://www.innolink.fi/datastrategiaopas/

De Bono, E. (1985). Six Thinking Hats. London: Penguin Books.

Design Council (n.d.). The Double Diamond. Available 16.02.2025: https://www.designcouncil.org.uk/our-resources/the-double-diamond/

Dibouliya, A. (2025). Unified data governance framework for AI-enabled data warehouses in banking. European Modern Studies Journal, 9(4), 67-76. https://doi.org/10.59573/emsj.9(4).2025.7

Earley, S., & Henderson, D. (2017). DAMA-DMBOK: Data Management body of knowledge (2nd ed.). Data Management Association, 20-572.

Erander, L., & Granat, E. (2024). Governing data for artificial intelligence: A qualitative study of the human factors that affect governance of client-data for IT-consulting firms. Lund University.

Figueiredo, R. M. C., Pedrosa, G. V., Gardenghi, J. L. C., Venson, E., & Júdice, M. (2024). Co-design and co-creation in digital public services: A service design approach. *Proceedings of WCGE 2024*. https://doi.org/10.5753/wcge.2024.2505

Finanssivalvonta. (N.d.). Sääntelykokonaisuudet. Available 23.2.2025: https://www.finanssivalvonta.fi/saantely/saantelykokonaisuudet/

Fruhwirth, M., Breitfuss, G., & Pammer-Schindler, V. (2020). The Data Product Canvas – A visual collaborative tool for designing data-driven business models. In 33rd Bled eConference Proceedings, 515-528. University of Maribor Press.

Haberl, A., Malin, C. D., & Thalmann, S. (2022). A framework to identify data governance requirements in open data ecosystems. In 35th Bled eConference Proceedings, 359-374. University of Maribor Press.

Heisig, P. (2009). Harmonisation of Knowledge Management- comparing 160 KM frameworks around the globe. Journal of Knowledge Management, 13, 4-31.

Helsing, J. (2025). Developing data management with service design. Kajaani University of Applied Sciences. https://urn.fi/URN:NBN:fi:amk-2025091724768

Justy, T., Pellegrin-Boucher, E., Lescop, D., Granata, J., & Gupta, S. (2023). On the edge of big data: Drivers and barriers to data analytics adoption in SMEs. Technovation, 127, 102850. https://doi.org/10.1016/j.technovation.2023.102850

Kaario, K., & Peltola, T. (2008). Tiedonhallinta: Avain tietotyön tuottavuuteen. Jyväskylä: Docendo.

Khatri, V., & Brown, C. V. (2010). Designing Data Governance. Communications of the ACM, 53(1), 148–152.

Koivisto, M., Säynäjäkangas, J., & Forsberg, S. (2019). Palvelumuotoilun bisneskirja. Helsinki: Alma Talent Oy.
Kosonen, M. (2019). Tiedolla johtamisen käsikirja. Kaakkois-Suomen Ammattikorkeakoulu. https://urn.fi/URN:ISBN:978-952-344-183-5
Liebowitz, J. (2016). Successes and Failures of Knowledge Management. Harrisburg University of Science and Technology, Harrisburg, Pennsylvania, 9-22.
Loshin, D. (2020). Evaluating the Business Impacts of Poor Data Quality. Knowledge intetgrity, Inc, Silver Spring, Maryland.
Lukka, K. 2003. The Constructive Research Approach, in Ojala, L. & Hilmola, O.-P. (eds.) Case Study Research in Logistics, Publications of Turku School of Economics and Business Administration, Series B, B1:2003, pp. 83–120.
Madhavan, D. (2024). Enterprise data governance: A comprehensive framework for ensuring data integrity, security, and compliance in modern organizations. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 10*(5). https://doi.org/10.32628/cseit241051062
Mariani, I. (2024). Shaping public sector digital transformation through design: Translation approaches on training programs as multi-stakeholder ecosystems. *Proceedings of DRS 2024*. https://doi.org/10.21606/drs.2024.1048
Mazzarella, A. J., III. (2024). *Towards resilient data governance: An industry study on practitioner perspectives on the current state of data governance*. Doctoral dissertation. ProQuest Dissertations Publishing.
Merkus, J., Helms, R. W., & Kusters, R. J. (2023). Data governance capabilities: empirical validation in case studies of large organisations. In A. Pucihar, M. K. Borštnar, R. Bons, G. Ongena, M. Heikkilä, & D. Vidmar (Eds.), 36th Bled eConference Digital Economy and Society: The Balancing Act for Digital Innovation in Times of Instability (1 ed., pp. 35-48). University of Maribor Press. https://doi.org/10.18690/um.fov.4.2023
Nokkala, T. (2020). Governance of platform data: From Canonical Data Models to Federative Interoperability, University of Turku.
Office of the Data Protection Ombudsman (n.d.). Tietosuojalainsäädäntö. Available 23.2.2025: https://tietosuoja.fi/lainsaadanto
Pedrosa, G. V., & Araujo Freitas, L. M. (2024). Toward a human-centered and ecosystemic evaluation framework for public sector digital planning. Journal of Technology Management & Innovation, 20(3), 31–56.
Piirainen, K. A., & Gonzalez, R. A. (2013). Constructive synergy in design science research: A comparative analysis of design science research and the constructive research approach. LTA, 3–4, 206–234.
Redman, T. (1998). The Impact of Poor Data Quality on the Typical Enterprise. Communications of the ACM Volume 41, Issue 2, 79-82.
Stickdorn, M., & Schneider, J. (2011). This is service design thinking: Basic, tools, cases. Bis Publishers.
Traficom (n.d.). EU digital and data statutes in a nutshell. Available 23.2.2025: https://www.traficom.fi/en/communications/data-economy-and-digital-services/eu-digital and-data-statutes-nutshell?
Vargo, S. L., Maglio, P. P, Akaka & M. A. (2008). On value and value co-creation: A service systems and service logic perspective. https://doi.org/10.1016/j.emj.2008.04.003
Wallis, I. (2021). Data Strategy: From definition to execution. Swindon: BCS Learning & Development Limited.
Zhang, S, Maharajan, S., Bygrave, L.A., & Yu, S. (2025). Data Sharing, Privacy and Security Considerations in the Energy Sector: A Review from Technical Landscape to Regulatory Specifications. arXiv:2503.03539v1 [cs.CR], 5 March 2025

Four Gen AI-Enabled Value Logics in Industrial B2B Markets

Eveliina Lakka,[1] Joel Mero,[1] Matti Leppäniemi,[1]
Heikki Karjaluoto[2]

[1] Jyväskylä University School of Business and Economics, Jyväskylä, Finland
eveliina.a.lakka@student.jyu.fi, joel.mero@jyu.fi, matti.leppaniemi@jyu.fi
University of Jyväskylä, Faculty of Information Technology, Jyväskylä, Finland
heikki.karjaluoto@jyu.fi

DOI
https://doi.org/10.18690/um.fov.4.2026.23

ISBN
978-961-299-152-4

Generative artificial intelligence (Gen AI) is rapidly transforming industrial B2B markets, yet firms struggle to translate its technological capabilities into realized value. This qualitative multiple-case study examines Gen AI-enabled value logics in six large global industrial B2B firms based on 15 semi-structured interviews. Using an abductive approach, we identify four archetypal value logics: synthesizing, interactional, predictive, and diagnostic. These logics capture distinct ways in which Gen AI structures dispersed data, enhances communication, translates predictive insights into actionable guidance, and supports diagnostic problem resolution. Drawing on a framework integrating AI intelligence types and human–AI interaction modes as an analytical lens, we show that Gen AI value logics in industrial B2B contexts cluster within thinking intelligence and augmentation or automation configurations, reflecting contextual constraints such as safety requirements and relational norms. By shifting attention from technological capabilities to configuration-level value logics, this study advances understanding of how Gen AI reconfigures value creation in industrial B2B markets and offers a typology to guide managerial implementation.

Keywords:
generative artificial intelligence, value logics, industrial B2B markets, human-AI interaction, value creation

University of Maribor Press

1 Introduction

Generative artificial intelligence (Gen AI) is rapidly transforming how businesses operate and engage with customers across sectors (Teng et al., 2025; Yuan et al., 2025). In industrial B2B markets, Gen AI reshapes how firms configure and realize value creation through context-sensitive analysis, content generation, and scalable knowledge integration (McKinsey & Company, 2025; Haefner et al., 2023). Yet despite these technological advances, many firms struggle to translate AI investments into sustained competitive advantage (Barney and Reeves, 2024; Kemp, 2024; Yuan et al., 2025).

Existing research on AI and Gen AI has examined capabilities and competencies (Ayala et al., 2025; Li et al., 2024; Sjödin et al., 2021), ecosystems and infrastructures (Burström et al., 2021), adoption drivers and barriers (Chen et al., 2022; Maga and Bodlaj, 2025), and implementation considerations (Grewal et al., 2025; Reim et al., 2020). However, the value logics enabled by Gen AI remain underexplored. In this study, value logics refer to archetypal configurations through which firms structure and realize value creation enabled by Gen AI in industrial B2B contexts. This perspective shifts attention from technological features or articulated market promises to recurring patterns in how value is configured and enacted in practice.

Understanding these logics is critical because Gen AI does not merely enhance analytical processes; it reshapes how knowledge is interpreted, communicated, and operationalized across organizational and relational boundaries (Akter et al., 2023; Jorzik et al., 2024). Moreover, the configuration of human–AI interaction plays a central role in determining how value is enacted in industrial settings (Ayala et al., 2025; Cui et al., 2024; Singh et al., 2024).

In this paper, we aim to deepen the understanding of Gen AI-enabled value logics in industrial B2B markets. We develop a conceptual model integrating two dimensions: AI intelligence types, mechanical, thinking, and feeling (Huang and Rust, 2018), and modes of human–AI interaction, augmentation, automation, and autonomous collaboration (Azer and Alexander, 2025; Davenport and Kirby, 2016). Our research question is: What types of Gen AI-enabled value logics emerge in industrial B2B firms? Drawing on a qualitative multiple-case study of six large

industrial firms and 15 interviews, we identify four archetypal value logics: synthesizing, interactional, predictive, and diagnostic.

2 Literature review

2.1 Defining Gen AI

Artificial intelligence can be defined as a system's ability to interpret external data, learn from it, and use these insights to achieve specific goals through adaptive processes (Haenlein and Kaplan, 2019). AI functions as an umbrella term encompassing various algorithmic approaches, including machine learning and deep learning (Strobel et al., 2024). Machine learning enables systems to detect patterns and improve performance over time, while deep learning relies on layered neural networks to learn complex representations from large datasets (Ayala et al., 2025; Naeem et al., 2025).

Generative artificial intelligence is a subset of deep learning distinguished by its ability to produce novel content rather than merely analyze or predict based on existing data (Huang et al., 2024; Singh et al., 2024; Yuan et al., 2025). Compared to conventional AI systems, Gen AI demonstrates stronger adaptability, contextual sensitivity, multimodal data processing capabilities, and the ability to generate human-like outputs (Mariani and Dwivedi, 2024). These capabilities are typically enabled by large-scale generative models trained on extensive datasets and refined through iterative feedback processes (IBM, 2024; Strobel et al., 2024).

The content-generative capacity of Gen AI is central to this study. While conventional AI systems primarily produce analytical outputs such as predictions and classifications, Gen AI generates communicable content that reshapes how knowledge is interpreted and how value is articulated and realized in practice (Grewal et al., 2025; Mariani and Dwivedi, 2024; Strobel et al., 2024). Rather than merely enhancing analytical processes, Gen AI transforms the layer through which insights are expressed and operationalized. In the empirical analysis, we use this content-generation criterion to identify what is distinctly generative within each value logic.

2.2 Value Creation and Value Logics in B2B Markets

Building on the definition introduced above, we situate value logics within prior research on value creation and business model design. Prior studies have examined how firms organize and articulate value creation for customers (Payne et al., 2017; Teece, 2010), while service-dominant logic emphasizes that value emerges through interaction and resource integration rather than being embedded solely in outputs (Vargo and Lusch, 2004, 2008; Vargo et al., 2026). Conceptualizations of value have evolved from exchange-oriented, cost-benefit formulations toward multidimensional and use-oriented views that emphasize experience, relationships, and co-creation (Zeithaml et al., 2020; Eggert et al., 2018). In industrial B2B markets, this shift is reflected in a move toward relational and use-based conceptions of buyer–supplier exchange (Eggert et al., 2018). These perspectives suggest that value is enacted through configurations of resources, processes, and relationships.

Gen AI reconfigures both dimensions of this discussion. First, it reshapes how firms articulate and communicate insights, interpretations, and recommendations in customer interactions (Ruokonen and Ritala, 2025; Wessel et al., 2025). Second, it introduces a non-human actor into value co-creation processes, altering how resources are interpreted, combined, and enacted in industrial B2B relationships (Wessel et al., 2025; Wirtz and Stock-Homburg, 2025). Examining Gen AI through a value logic lens, therefore, shifts attention from technological features to how value creation is configured in practice.

2.3 Enhancing B2B Value Logics with Gen AI

To analyze Gen AI-enabled value logics in industrial B2B markets, we employ an analytical framework grounded in two dimensions: AI intelligence types (Huang and Rust, 2018) and modes of human–AI interaction (Azer and Alexander, 2025; Davenport and Kirby, 2016; Raisch and Krakowski, 2021). Together, these dimensions specify both the nature of value generated and how value is configured between human and AI actors.

Huang and Rust (2018) distinguish among mechanical, thinking, and feeling intelligence. Mechanical AI supports routine tasks, thinking AI enables analytical reasoning and decision support, and feeling AI addresses emotional interpretation.

Gen AI is most closely associated with thinking intelligence due to its capacity to generate novel and context-sensitive content (Sedkaoui and Benaichouba, 2024), while also incorporating elements of mechanical automation and emerging affective responsiveness (Huang and Rust, 2024). These intelligence types delineate the qualitative character of value that Gen AI can help configure.

The second dimension concerns how humans and AI collaborate in value creation. Human–AI interaction ranges from augmentation, where AI supports human decision-making, to automation, where AI performs tasks under human supervision, and to autonomous configurations, where AI operates independently within defined boundaries (Davenport and Kirby, 2016; Raisch and Krakowski, 2021). These modes are not strictly sequential; rather, they may coexist and shift depending on contextual and task-specific constraints.

Integrating these dimensions, we develop a conceptual model presented in Table 1. The matrix outlines theoretically possible configurations of Gen AI-enabled value logics across combinations of intelligence types and interaction modes. We use this framework as an analytical lens for interpreting empirical patterns rather than as a rigid classificatory scheme. In industrial B2B contexts, certain configurations may be more feasible than others (Burström et al., 2021).

Table 1: Analytical Framework for Mapping Gen AI-Enabled Value Logics

	Augmentation (Human-led, AI-assisted)	Automation (AI-led, human-supervised)	Autonomous (AI-led, self-supervised)
Mechanical AI	AI supports humans by preparing structured inputs	AI executes routine process steps	AI runs defined workflows independently within specified boundaries
Thinking AI	AI provides insights to support human decisions	AI produces draft decisions or recommendations for human approval	AI optimizes and adapts dynamically within predefined parameters
Feeling AI	AI interprets emotional cues to help humans tailor responses	AI delivers emotionally attuned responses under human oversight	AI adapts communication patterns independently based on detected emotional cues

3 Methods

3.1 Research Approach

To investigate the types of value logics emerging through industrial B2B firms' use of Gen AI, we applied a qualitative multiple-case study approach (Dubois and Gadde, 2002, 2014; Yin, 2018). Given the limited research on Gen AI-enabled value logics in industrial contexts, qualitative methods are appropriate for such discovery-focused research (Eisenhardt and Graebner, 2007; Siggelkow, 2007). A multiple-case study design enabled the generation of holistic insights into how Gen AI-enabled value logics take shape within their real-world contexts (Yin, 2018).

3.2 Data Collection

Table 2: Overview of Firms and Respondents

Firm	Industry	Workforce	Identified Gen AI Value Logics	Interviewees
A	Elevators and escalators	60,000-65,000	Synthesizing, interactional, predictive, diagnostic	Senior Gen AI Strategic Designer; Product Manager (2)
B	Chemicals	≤ 1,000	Synthesizing	Development Manager; CIO
C	Welding	≤ 1,000	Synthesizing	Service Manager
D	Electrification and automation	≥ 100,000	Synthesizing, interactional, predictive, diagnostic	SVP Service; Head of Sales; Director Digital Ecosystems
E	Mining and metals	15,000-20,000	Synthesizing, diagnostic	SVP Services; Lead Service Designer; Director Data & AI; Design Manager
F	Marine and energy	15,000-20,000	Synthesizing, predictive, diagnostic	Lead Business Innovation Manager; Head of Lifecycle Data Operations

We employed purposive sampling (Patton, 2015) to identify six industrial B2B firms operating in different markets that had begun leveraging Gen AI in concrete initiatives relevant to value creation. The empirical data were collected through semi-structured interviews with informants responsible for developing Gen AI customer interactions and related value creation practices within their organizations. We employed snowball sampling, in which key informants recommended additional experts from their firms for further interviews (Miles *et al.,* 2014). An open-ended thematic interview guide was used to focus on the key objectives, requirements, and

challenges of leveraging Gen AI to create customer value. We conducted 15 interviews with the six case firms, each transcribed verbatim. An overview of the firms, interviewees, and identified value logics is provided in Table 2.

3.3 Data Analysis

We used an abductive analysis approach, iteratively moving between empirical data and theoretical literature to conceptualize Gen AI-enabled value logics (Dubois and Gadde, 2002, 2014). The analysis proceeded in three stages consistent with Yin's (2018) multiple-case study techniques. First, we conducted within-case analysis to develop detailed descriptions of Gen AI-enabled value creation and modes of human-AI interaction within each firm. Second, we conducted a cross-case synthesis to compare cases and identify recurring patterns. Third, we used these cross-case findings to support analytic generalization toward higher-level theoretical insights about Gen AI-enabled value logics. These generalizations were elaborated abductively through iterative movement between data and theory (Dubois and Gadde, 2002, 2014).

4 Findings

Based on the data analysis, we identified four archetypal Gen AI-enabled value logics enacted by industrial B2B firms: synthesizing, interactional, predictive, and diagnostic. Firms frequently enacted multiple value logics at once rather than relying on a single dominant configuration. Table 3 summarizes the core characteristics of each value logic, including their primary intelligence type, interaction mode, generative mechanism, commercialization pattern, and key challenges.

Table 3: Gen AI-enabled Value Logics

	Synthesizing Value	Interactional Value	Predictive Value	Diagnostic Value
Mechanical and Thinking AI	Mechanical AI	Thinking AI	Thinking AI	Thinking AI
Human-AI Interaction	Augmentation & Automation	Augmentation & Automation	Augmentation & Automation	Augmentation & Automation
Overall Goal	Process and structure large volumes of data	Enhance communication and information accessibility	Analyze data to forecast and make recommendations	Identify root causes and help with problem resolution
Distinct Generative Mechanism	Generates structured documents	Generates conversational, context-sensitive responses	Generates narrative explanations and optimization strategies	Generates step-by-step troubleshooting guidance and resolution pathways
Key Customer Benefit	Faster, more comprehensive service	Enhanced access to relevant information	Improved operational efficiency	Self-service problem solving; faster service
Commer-cialization Pattern	Enhances existing services	Embedded or subscription add-ons	New service modules or add-ons	Enhances existing service contracts
Key Challenge	Siloed data	Human preference for personal interaction	Data acquisition and accuracy	Product/process complexity

4.1 Synthesizing Value

The synthesizing value logic focused on structuring large volumes of dispersed data to improve efficiency and accessibility across internal and customer-facing processes. It combined mechanical and thinking intelligence within configurations of augmentation and automation.

Its distinct generative mechanism lies in Gen AI's ability to compose structured documents from unstructured inputs. Rather than merely retrieving or classifying information, Gen AI generated integrated outputs such as draft tenders, sales proposals, and part-identification reports that combined multiple data sources into coherent, human-readable formats. For example, one firm used Gen AI to identify damaged spare parts from photographs to ensure accurate replacements, while another applied it to extract tender requirements and suggest compatible products. As one informant explained:

"Gen AI has been utilized in the sales proposal creation process, especially in simple, high-volume product areas. It enables making customer proposals faster and more easily" (Director, Firm D).

Unlike the other value logics, synthesizing value was predominantly internally oriented. Customer benefits were realized indirectly through faster responses, improved documentation accuracy, and greater service consistency. These capabilities were embedded within existing operational processes rather than introduced as separate commercial initiatives. A key challenge concerned siloed data across legacy systems and organizational units, which constrained effective Gen AI deployment:

"Customer sentiment data has traditionally been owned by marketing and sales. Technical data is owned elsewhere, and combining these two data areas brings many discussions" (Head of Lifecycle Data Operations, Firm F).

Addressing these silos required both technical integration and cross-functional coordination.

4.2 Interactional Value

The interactional value logic centered on enhancing customer communication and improving access to relevant information. It was primarily grounded in thinking intelligence and enacted through augmentation and automation.

Its distinct generative mechanism was Gen AI's ability to produce conversational, context-sensitive responses tailored to specific customer queries. Unlike conventional AI-based search systems that retrieve pre-authored content, Gen AI generated novel responses by integrating multiple knowledge sources and adapting outputs to the framing of each request. For instance, one firm implemented a Gen AI-driven chatbot within its customer support portal, while another used Gen AI to assist call center agents by suggesting context-specific responses. As one informant noted:

"Based on a customer inquiry [to the call center], the [Gen AI-based] tool already provides potential responses, and they [the agents] then choose the one that is most applicable, or they may make some changes. ... We have not made the step yet to automate customer interaction completely. This is due

to the safety requirements of our industry, where we do not want to risk giving customers incorrect advice" (Strategic Designer, Firm A).

Interactional value increased accessibility and responsiveness by enabling customers to obtain relevant information more quickly and with less effort. In most cases, Gen AI operated as an assistant to customer-facing personnel, scaling expertise and promoting consistency across interactions. Even when embedded behind human agents, the value logic remained directly customer-facing. Commercialization typically involved embedding AI capabilities within existing services or offering premium features as add-ons. A key challenge concerned established customer preferences for direct human interaction and strong personal relationships:

"As the customer base is limited, our interaction with customers tends to be very personal. These interactions tend to be personified to our account managers [rather than mediated by technology]" (Director, Firm E).

This preference for human contact required firms to balance AI-enabled efficiency gains with maintaining the relational quality valued by customers.

4.3 Predictive Value

The predictive value logic focused on analyzing operational data to support forward-looking decision-making, particularly in predictive maintenance and resource optimization. It was grounded in thinking intelligence and enacted through augmentation and automation.

While predictive analytics has long been present in industrial B2B markets, the generative contribution of Gen AI did not primarily concern improved forecasting accuracy. Instead, Gen AI transformed how predictive insights were communicated and operationalized. By generating narrative explanations, scenario-based reasoning, and natural-language recommendations, Gen AI enabled customers to understand not only what was predicted, but why it mattered and how different courses of action might unfold. This communicative layer bridged quantitative model outputs and situated decision-making. For example, one firm used Gen AI to analyze weather and energy market data to optimize energy consumption strategies. Another

enhanced predictive maintenance services by refining forecasts and contextualizing recommendations:

"Instead of just telling the customer they need a service, and maybe explaining why, we can now also estimate what would happen if they do not have that service" (Product Manager, Firm A).

The predictive value logic emphasized foresight. Customers could anticipate operational risks, evaluate the consequences of inaction, and make more informed decisions about maintenance and resource allocation. For firms, these capabilities enabled scalable service extensions across installed bases. A key challenge concerned the availability and reliability of operational data:

"We know what the configuration of the asset is. But the challenge [for predictive accuracy] is that while we know the asset and its surroundings well, the actual process or the operational purpose of the asset varies widely" (Director, Firm E).

Such data limitations constrained the reliability and applicability of predictive Gen AI solutions and required firms to invest in improved data integration and governance practices.

4.4 Diagnostic Value

The diagnostic value logic centered on identifying the root causes of operational problems and supporting their resolution. It was grounded in thinking intelligence and enacted through augmentation and automation.

Although diagnostic systems have long relied on conventional AI techniques such as anomaly detection and fault classification, Gen AI added a generative layer by producing step-by-step troubleshooting guidance tailored to specific operational contexts. Rather than simply detecting and classifying malfunctions, Gen AI generated contextual natural-language advice that guided technicians or customers through resolution processes in real time, explaining both how and why particular actions were appropriate. For example, one firm integrated Gen AI into its remote services to support troubleshooting, while another piloted diagnostic tools providing real-time guidance to field technicians:

"The fact that our field service technicians can now get real-time help and advice for their problem-solving when there is an issue at the customer site, that is something we have piloted now with Gen AI" (Strategic Designer, Firm A).

This value logic emphasized faster and more accurate problem resolution. Customers could address certain issues independently and accelerate service encounters requiring on-site intervention. For firms, diagnostic capabilities also generated insights into usage patterns and recurring failure points, improving reliability and reducing troubleshooting costs. These capabilities were typically embedded within existing service contracts rather than introduced as standalone offerings. A key challenge stemmed from product and process complexity:
"In industrial markets, it is rarely the case that you just download something from the internet and start using it. Instead, there is much configuration required to have exactly the right approach and data sources" (Director, Firm D).

The bespoke nature of industrial equipment required significant contextual adaptation of Gen AI systems, limiting the scalability of standardized diagnostic solutions across diverse installed bases.

5 Discussion and Conclusions

5.1 Theoretical Contributions

Drawing on a qualitative multiple-case study, this research makes two contributions. First, we develop a conceptual model for understanding Gen AI-enabled value logics in industrial B2B markets by integrating AI intelligence types (Huang and Rust, 2018) with modes of human–AI interaction (Azer and Alexander, 2025; Davenport and Kirby, 2016; Raisch and Krakowski, 2021). While prior research has examined these constructs separately, we connect them through a value logic lens and apply the framework empirically.

Second, we identify four archetypal Gen AI-enabled value logics: synthesizing, interactional, predictive, and diagnostic. We respond to calls for research that clarifies how AI reshapes value creation and that develops typological understandings of AI-driven configurations (Bahoo et al., 2023; Jorzik et al., 2024). Our findings offer context-specific insight into industrial B2B markets, where

technological capabilities, organizational conditions, and relational norms appear to jointly shape how Gen AI is enacted. A central empirical pattern across our cases is that all four value logics cluster within thinking intelligence and augmentation or automation configurations. Autonomous and feeling-oriented configurations remained largely unoccupied in our data. These patterns suggest that technological capability alone does not determine enacted value logics. In the industrial B2B contexts we studied, factors such as safety requirements, siloed data, product complexity, and customization demands shape how Gen AI is configured in practice, constraining enactment to modes that preserve human oversight and relational trust.

From an organizational perspective, augmentation and automation did not emerge as sequential stages but rather appeared as coexisting configurations within a single value logic. This is consistent with the view that human–AI interaction is dynamically balanced rather than linearly progressing toward autonomy (Raisch and Krakowski, 2021). From a relational perspective, Gen AI introduces a non-human actor into buyer–supplier interactions (Wessel et al., 2025), potentially enabling greater responsiveness and foresight while simultaneously requiring careful preservation of trust-based norms (Kamalaldin et al., 2020). The unoccupied cells in the analytical framework may thus be read as theoretically possible but contextually constrained configurations rather than as inevitable developmental trajectories.

5.2 Managerial Implications

For managers, the typology of Gen AI-enabled value logics offers a structured lens for evaluating AI initiatives. Rather than focusing on technological capabilities, firms can assess how Gen AI is configured in relation to value creation objectives and organizational readiness. The predominance of augmentation and automation suggests that near-term gains are most effectively realized through human–AI collaboration. At the same time, firms should evaluate where greater autonomy is contextually appropriate, particularly in predictive and diagnostic applications where outputs can be validated against measurable operational outcomes. Interactional value logics, which depend on relational trust and contextual judgment, are likely to remain human-centered. For customers, the typology offers distinctions obscured by generic claims about AI-enabled value. Customers may use it to articulate the

value they are seeking and to negotiate the degree of human involvement that aligns with their safety, confidentiality, and relational preferences.

5.3 Limitations and Future Research Areas

This study has several limitations that point to avenues for further research. First, our empirical base of six large industrial firms and 15 interviews is modest, and may explain the limited variation observed across intelligence types and interaction modes. Future research should examine smaller firms and alternative contexts to assess the generalizability of the identified value logics. Second, the cross-sectional design limits insight into how value logics evolve over time. Longitudinal studies could explore how configurations shift as technological capabilities and organizational readiness develop. Third, the distinction between mechanical, thinking, and feeling intelligence warrants reassessment as AI technologies mature. Finally, ethical considerations, including accountability, bias, and trust in AI-mediated interactions, merit investigation, particularly regarding how governance practices shape the legitimacy and development of Gen AI-enabled value logics.

Acknowledgment

This work was supported by the Foundation for Economic Education.

References

Akter, S., Hossain, M.A., Sajib, S., Sultana, S., Rahman, M., Vrontis, D. and McCarthy, G. (2023) "A framework for AI-powered service innovation capability: Review and agenda for future research", *Technovation*, 125, p.102768.

Ayala, N.F., Rodrigues da Silva, J., Cannarozzo Tinoco, M.A., Saccani, N. and Frank, A.G. (2025) "Artificial Intelligence capabilities in Digital Servitization: Identifying digital opportunities for different service types", *International Journal of Production Economics*, 284, p.109604.

Azer, J. and Alexander, M. (2025) "Human-Machine Engagement (HME): conceptualization, typology of forms, antecedents, and consequences", *Journal of Service Research*, 28(1), pp. 112-130.

Bahoo, S., Cucculelli, M. and Qamar, D. (2023) "Artificial intelligence and corporate innovation: a review and research agenda", *Technological Forecasting and Social Change*, 188, p. 122264.

Barney, J.B. and Reeves, M. (2024) "AI won't give you a new sustainable advantage", *Harvard Business Review*, 102(5), pp. 72-79.

Burström, T., Parida, V., Lahti, T. and Wincent, J. (2021) "AI-enabled business-model innovation and transformation in industrial ecosystems: a framework, model and outline for further research", *Journal of Business Research*, 127, pp. 85-95.

Chen, L., Jiang, M., Jia, F. and Liu, G. (2022) "Artificial intelligence adoption in business-to-business marketing: toward a conceptual framework", *The Journal of Business & Industrial Marketing*, 37(5), pp. 1025-1044.

Cui, Y., van Esch, P. and Phelan, S. (2024) "How to build a competitive advantage for your brand using generative AI", *Business Horizons*, 67(5), pp. 583-594.
Davenport, T.H. and Kirby, J. (2016) "Just how smart are smart machines?", *MIT Sloan Management Review*, 57(3), p. 21.
Dubois, A. and Gadde, L.E. (2002) "Systematic combining: an abductive approach to case research", *Journal of Business Research*, 55(7), pp. 553-560.
Dubois, A. and Gadde, L. (2014) "Systematic combining—a decade later", *Journal of Business Research*, 67(6), pp. 1277-1284.
Eggert, A., Ulaga, W., Frow, P., and Payne, A. (2018) "Conceptualizing and communicating value in business markets: From value in exchange to value in use", *Industrial marketing management,* 69, pp. 80-90.
Eisenhardt, K.M. and Graebner, M.E. (2007) "Theory building from cases: opportunities and challenges", *Academy of Management Journal*, 50(1), pp. 25-32.
Grewal, D., Satornino, C.B., Davenport, T. and Guha, A. (2025) "How generative AI is shaping the future of marketing", *Journal of the Academy of Marketing Science*, 53(3), pp. 702-722.
Haefner, N., Parida, V., Gassmann, O. and Wincent, J. (2023) "Implementing and scaling artificial intelligence: a review, framework, and research agenda", *Technological Forecasting and Social Change*, 197, p. 122878.
Haenlein, M. and Kaplan, A. (2019) "A brief history of artificial intelligence: on the past, present, and future of artificial intelligence", *California Management Review*, 61(4), pp. 5-14.
Huang, M. and Rust, R.T. (2018) "Artificial intelligence in service", *Journal of Service Research*, 21(2), pp. 155-172.
Huang, M. and Rust, R.T. (2024) "The caring machine: feeling AI for customer care", *Journal of Marketing*, 88(5), pp. 1-23.
IBM (2024) *What is generative AI?* Available at: https://www.ibm.com/think/topics/generative-ai (Accessed: 20 August 2025).
Jorzik, P., Klein, S.P., Kanbach, D.K. and Kraus, S. (2024) "AI-driven business model innovation: a systematic review and research agenda", *Journal of Business Research*, 182, p. 114764.
Kamalaldin, A., Linde, L., Sjödin, D., & Parida, V. (2020) "Transforming provider-customer relationships in digital servitization: A relational view on digitalization", *Industrial Marketing Management, 89*(C), pp. 306-325.
Kemp, A. (2024) "Competitive advantage through artificial intelligence: toward a theory of situated AI", *The Academy of Management Review*, 49(3), pp. 618-635.
Li, L., Tong, Y., Liu, Y. and Yang, S. (2024) "Artificial Intelligence-Enabled customer value proposition capability and market performance: the moderating role of environmental heterogeneity", *IEEE Transactions on Engineering Management*, 71, pp. 5588-5599.
Maga, S. and Bodlaj, M. (2025) "Drivers and outcomes of chatbot use in the business-to-business context: exploring business customers' perspectives", *The Journal of Business & Industrial Marketing*, 40(1), pp. 250-264.
Mariani, M. and Dwivedi, Y.K. (2024) "Generative artificial intelligence in innovation management: a preview of future research developments", *Journal of Business Research*, 175, p. 114542.
McKinsey & Company (2025) *The state of AI*. Available at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai (Accessed: 20 August 2025).
Miles, M.B., Huberman, A.M. and Saldaña, J. (2014) *Qualitative data analysis: a methods sourcebook*. 3rd edn. Thousand Oaks: Sage.
Naeem, R., Kohtamäki, M. & Parida, V. (2025) "Artificial intelligence enabled product–service innovation: Past achievements and future directions: Artificial intelligence enabled product–service innovation: past", *Review of managerial science, 19*(7), pp. 2149-2192.
Patton, M.Q. (2015) *Qualitative research and evaluation methods*. Thousand Oaks: Sage.
Payne, A., Frow, P. and Eggert, A. (2017) "The customer value proposition: evolution, development, and application in marketing", *Journal of the Academy of Marketing Science*, 45(4), pp. 467-489.

Raisch, S. and Krakowski, S. (2021) "Artificial intelligence and management: the automation–augmentation paradox", *The Academy of Management Review*, 46(1), pp. 192-210.

Reim, W., Åström, J. and Eriksson, O. (2020) "Implementation of artificial intelligence (AI): a roadmap for business model innovation", *AI*, 1(2), pp. 180-191.

Ruokonen, M. and Ritala, P. (2025) "Managing Generative AI for Strategic Advantage", *Research technology management, 68*(4), pp. 11-20.

Sedkaoui, S. and Benaichouba, R. (2024) "Generative AI as a transformative force for innovation: a review of opportunities, applications and challenges", *European Journal of Innovation Management*.

Siggelkow, N. (2007) "Persuasion with case studies", *Academy of Management Journal, 50*(1), pp. 20-24.

Singh, K., Chatterjee, S. and Mariani, M. (2024) "Applications of generative AI and future organizational performance: the mediating role of explorative and exploitative innovation and the moderating role of ethical dilemmas and environmental dynamism", *Technovation*, 133, p. 103021.

Sjödin, D., Parida, V., Palmié, M. and Wincent, J. (2021) "How AI capabilities enable business model innovation: scaling AI through co-evolutionary processes and feedback loops", *Journal of Business Research*, 134, pp. 574-587.

Strobel, G., Banh, L., Möller, F. and Schoormann, T. (2024) "Exploring generative artificial intelligence: a taxonomy and types", in *Proceedings of the 57th Hawaii International Conference on System Sciences*.

Teece, D. J. (2010) Business Models, Business Strategy and Innovation. *Long range planning, 43*(2), pp. 172-194.

Teng, D., Ye, C. and Martinez, V. (2025) "Gen-AI's effects on new value propositions in business model innovation: evidence from information technology industry", *Technovation*, 143, p. 103191.

Vargo, S.L. and Lusch, R.F. (2004) "Evolving to a new dominant logic for marketing", *Journal of Marketing*, 68(1), pp. 1-17.

Vargo, S.L. and Lusch, R.F. (2008) "Service-dominant logic: continuing the evolution", *Journal of the Academy of Marketing Science*, 36(1), pp. 1-10.

Vargo, S. L., Wieland, H., and Akaka, M. A. (2026) "Service-Dominant Logic in the AI Era", *Journal of service research : JSR*.

Wessel, M., Adam, M., Benlian, A., Majchrzak, A. & Thies, F. (2025). "Generative AI and its Transformative Value for Digital Platforms", *Journal of management information systems, 42*(2), pp. 346-369.

Wirtz, J. and Stock-Homburg, R. (2025) "Generative AI Meets Service Robots", *Journal of service research: JSR, 28*(4), pp. 527-543.

Yin, R.K. (2018) *Case study research and applications: design and methods*. 6th edn. Thousand Oaks: SAGE Publications.

Yuan, C., Yin, Y., Wang, S., Zhang, Z. and Moon, H. (2025) "Exploring technology, organization and environmental factors driving firms generative AI adoption intention: the moderating role of trans-parasocial relation", *The Journal of Business & Industrial Marketing*, 40(5), pp. 1146-1159.

Zeithaml, V. A., Verleye, K., Hatak, I., Koller, M., and Zauner, A. (2020) "Three decades of customer value research: Paradigmatic roots and future research avenues", *Journal of Service Research, 23*(4), pp. 409–432.

Applying Antifragility in Design Principles for a Cyber Resilience Maturity Model

Edzo A. Botjes, Tim Huygh, Laury Bollen, Remko W. Helms

Open Universiteit, Heerlen, the Netherlands
edzo.botjes@ou.nl, tim.huygh@ou.nl, laury.bollen@ou.nl, remko.helms@ou.nl

In an increasingly hyperconnected world rising uncertainty threatens both business continuity and society, prompting new regulatory responses from the EU and the US. Traditional, deterministic approaches to cyber resilience which focus merely on returning to a 'status quo' state after a disruption are no longer sufficient. Instead, organizations must evolve toward antifragility: the ability to thrive and exploit opportunities emerging from chaos. Drawing on an iterative design science research methodology this study proposes a theoretically grounded framework consisting of two design goals and twelve design principles to help organizations achieve antifragility within a cyber resilience context and guide the development of future cyber resilience maturity models.

DOI
https://doi.org/10.18690/um.fov.4.2026.24

ISBN
978-961-299-152-4

Keywords:
antifragility, cyber resilience, cyber security, risk management, maturity models

University of Maribor Press

1 Introduction

In today's volatile, uncertain, complex, and ambiguous (VUCA) environments, organizations face unprecedented cyber threats and disruptions (Bennett & Lemoine, 2014; Saeed et al., 2023) which impact business operations, society, the global economy, and public trust (Eling et al., 2021; Reeder & Hall, 2021). Governments, including the EU and US, are addressing these challenges through cyber resilience legislation based on existing frameworks (Carroll, 2024; Tasheva & Kunkel, 2022; Vostoupal et al., 2023).

Traditional deterministic approaches to risk management and cyber security such as NIST CSF v2.0, ISO 27005, ISO 3100, NIST 800-30 and NIST 800-37 (Cerqueira Junior & Arima, 2023; NIST, 2023; Putra & Soewito, 2023; Taherdoost, 2022) are effective in complicated domains ("known unknowns") because of their reliance on checklists and static risk assessments. These frameworks are however less useful in chaotic and complex domains ("unknown unknown") because they fail to adapt to the dynamic nature of current cyber threats and emergent behavior of organizations (S. Ashby et al., 2018; Eling et al., 2021; Herley, 2016).

Predicting unknowns to protect existing value is not viable due to the infinite effort required. Organizations must instead build resilience to create value amid unforeseen disruptions (Adobor & Kudonoo, 2025; Aven, 2015a). This leads to the following **Problem Statement:** *Organizations struggle to create business value in chaotic and unpredictable contexts.* To assist organizations a Cyber Resilience Maturity Model (CRMM) may serve as a valuable tool (e.g., Imran et al., 2022; Sepúlveda Estay et al., 2020). However, despite extensive research (Botjes et al., 2026a), no existing CRMM adequately addresses the essential capability of responding to unforeseen events. In response the present paper proposes a theoretically grounded framework consisting of two design goals and twelve design principles to help organizations achieve antifragility within a cybersecurity context and guide the development of future cyber resilience maturity models. This leads to the **Research Question:** *What are the design goals and design principles for a Cyber Resilience Maturity Model (CRMM) that adequately addresses the capability to gain value in response to unforeseen events?*

To answer this question, we apply an iterative DSR approach (Section 2). We first capture the relevant theoretical constructs identified in our *diagnosis phase* (Figure 1, Section 3). In our *design objective and requirements phase* we subsequently draw on design theory to propose the design goals (Table 1) and design principles (Table 2) for such a CRMM.

2 Methodology

To enhance rigor in developing the CRMM we apply Design Science Research (DSR) throughout the CRMM lifecycle (Adekunle et al., 2022; Becker et al., 2009a). Recognizing our domain as a wicked, socio-technical problem, we follow an iterative DSR to develop our design knowledge (Vom Brocke et al., 2020) relevant to the creation of our knowledge artifact. We integrate the eADR (Mullarkey & Hevner, 2019) and the eDSR (Tuunanen et al., 2024) into five phases: (1) Diagnosis, (2) Design Objectives and Requirements, (3) Design and Development, (4) Implementation, and (5) Evolution. Each phase follows five steps: (a) Problem Formulation, (b) Artifact Creation, (c) Evaluation, (d) Reflection, and (e) Formalization of Learning (Mullarkey & Hevner, 2019). This paper shares the results from the iteration in and between the *diagnosis phase*, and the *design objectives and requirements phase*.

Within the *diagnosis phase* we applied exploratory searches using open databases (Botjes et al., 2026b) and proprietary AI tools (OpenAI ChatGTP and Google Gemini) to identify relevant models and constructs, and in doing so refining a coherent theoretical framework for both the problem- and solution-space (Kuechler & Vaishnavi, 2008; Rivard, 2021).

We define the *design objective and requirements* of the CRMM using these two design theory principles (Gregor, 2006; Gregor & Hevner, 2013): design goals define the artifact's purpose and scope (Gregor & Jones, 2007; Walls et al., 1992) and design principles translate goals into actionable guidelines (Markus et al., 2002). Both of these principles need to be grounded in justificatory knowledge (Baskerville & Pries-Heje, 2010; Fischer et al., 2010). We ensure transparency and traceability from goals to principles (Lukyanenko & Parsons, 2020), in support of the development of new knowledge (Hevner et al., 2024; Maedche et al., 2024), which involves balancing

specificity and generality to maximize re-usability and validation (Gregor et al., 2020).

To ensure replicability, independence, precision, and falsifiability (Recker, 2021) we practice open notetaking, documenting the process, design knowledge, and contributions (Hevner et al., 2024; Maedche et al., 2024), respecting the FAIR-principles (Wilkinson et al., 2016). We used open-source tools for editing, versioning, and reference management (Botjes et al., 2026b). Furthermore, open and proprietary databases are used for the literature review. Finally, proprietary AI (OpenAI ChatGPT, Mistral La Chat and Google Gemini) are used for spelling, grammar and word-reduction.

3 Conceptual Background

3.1 Resilience

Generally, resilience refers to how long an organization takes to recover from disruptions (e.g., context changes) and restore its original function and value (Linkov & Kott, 2019). It is central to business continuity (Steen et al., 2024) because changes in context can alter the value proposition of products or services (Lusch & Nambisan, 2015). For example, negative news about a bank can shift public perception leading to rapid withdrawal of digital assets and potential insolvency (Smeets, 2021). Detailed definitions on resilience vary, e.g., prescriptive or descriptive, as one layer or as a multi layered concept, as a capacity, ability, characteristic, outcome, process, behavior, strategy, approach, or as a type of performance (Hillmann & Guenther, 2021; Hosseini et al., 2016).

We adopt Martin-Breen and Anderies (2011), who distinguishes three sub-types of resilience: (1) Engineering Resilience (which maintains both construction and function through rigorous design); (2) Systems Resilience (which preserves function while adapting construction (e.g., using backup systems)); (3) Complex Adaptive Systems (CAS) Resilience (whichenables changes in both function and construction, by empowering autonomy and mastery). For cyber resilience all three sub-types are relevant, whereas traditional cyber resilience frameworks incorporate only the first two types of resilience.

3.2 Cyber Resilience

Digital transformation has made cyber resilience critical for all organizational stakeholders (Caon et al., 2025). This urgency stems from: (1) rising global cyber incident frequency (Dupont et al., 2023); (2) increasing threats to business continuity (Eling et al., 2021); (3) expanding global and regional regulations (Chiara, 2022; IMF, 2024; Tasheva & Kunkel, 2022; The World Bank, 2023; Vostoupal et al., 2023), which now address cyber resilience across product life-cycles and supply chains; and (4) the evolution of security, from physical, infrastructure, operational, information security to cybersecurity into cyber resilience at the overlap of enterprise risk management and business continuity (Eling et al., 2021; Linkov & Kott, 2019; Munusamy & Khodadi, 2023; Tzavara & Vassiliadis, 2024). Cyber resilience involves managing the complexity of socio-technical and cyber-physical systems, particularly the dynamic interplay between human and machine (Alrumaih et al., 2023; Belinski et al., 2020; Dupont et al., 2023), aimed to uphold business continuity by addressing the organization and its environment (Eling et al., 2021).

3.3 Cyber Resilience Frameworks

Traditional cyber resilience frameworks aim to identify risks and mitigate their negative impact on people, processes, and technology by using Engineering and Systems Resilience to absorb (internal and external) disruptions and restore original functions (Bharathy & McShane, 2014; Hosseini et al., 2016; Ruiz-Martín et al., 2018). However, risk events are inherently unpredictable, subjective, and too numerous to manage comprehensively (Aven, 2015a; 2018; Derbyshire & Wright, 2014; Eling et al., 2021; Fjaeran & Aven, 2021) due to nonlinear dynamic behavior (Boeing, 2016; Lorenz, 1963), knowledge gaps (Heylighen, 2002; Raviv et al., 2022), and black swan events that lie beyond current imagination (Thekdi & Aven, 2024). This means risk events are in the chaotic and complex ("unknown unknown") domain. According to the law of requisite variety (W. R. Ashby, 1958; Beer, 1974), the vast variety of risk events to be absorbed in our hyper-connected, nonlinear dynamic reality, renders traditional risk management approaches and frameworks impractical, resource-intensive, and fundamentally non-viable.

CAS Resilience introduces a proactive dimension to cyber resilience. By increasing internal variety through the adaptation of function and structure it empowers organizations to seize new opportunities and deliver greater value (Axenie et al., 2024; Martin-Breen & Anderies, 2011). It leverages the adaptive cycle of exploitation, conservation, release, reorganization and exploration, driving continuous evolution (Walker et al., 2004; Sundstrom & Allen, 2019). Proactive risk management enables organizations to identify and capitalize on upside risks (Aven, 2015b; Florea & Florea, 2016). By maximizing the frequency of exploitation and exploration cycles organizations can better handle the "unknown unknowns", thereby fostering antifragility.

3.4 Antifragility

Antifragility, introduced by Taleb (2012), goes beyond resilience by enabling systems to thrive in unpredictable environments and creating value in new realities (Adobor, 2024; Sagala & Őri, 2024). It optimizes conditions to trigger CAS Resilience, enabling organizations to adapt and capitalize on unforeseen events through learning (Axenie et al., 2024; Johnson & Gheorghe, 2013). Antifragility ensures organizations remain relevant to stakeholders by adapting their value propositions and internal structures (Adobor & Kudonoo, 2025; Parmar et al., 2010; Spitz & Desbiey, 2025). Given the dynamic internal and external contexts organizations operate in, cyber resilience must integrate antifragility to address the unexpected and sustain business continuity.

3.5 Sensemaking

Because of the continuous change in and around organizations, sensemaking is essential for building resilience (Bennett & Lemoine, 2014; Weick et al., 2005). Sensemaking is the process of interpreting reality between subjective and objective perspectives (Mukumbang et al., 2023; Mutch, 2010). Sensemaking guides decision-making in ambiguous contexts, helping organizations navigate uncertainty and adapt their actions effectively (Snowden, 2005; Thompson, 1995; Weick, 1995). It enables organizations to achieve sustainable competitive advantages through organizational learning and by creating, acquiring and integrating knowledge (Alashwal et al., 2017; Al-Shehab et al., 2005).

3.6 Organizational Learning

Managing uncertainty is the core objective of enterprise risk management, business continuity management, and cyber resilience. To achieve this, organizational learning must be embedded in cyber resilience from the boardroom down to the operational level (Dupont et al., 2023; Park et al., 2013). This process relies on continuous sensemaking by translating observations into actionable strategies and leveraging feedback loops to foster flexibility, improvement, and self-healing (Argyris & Schön, 1978; Huang et al., 2022; Nonaka & Takeuchi, 1995).

Organizational learning operates in two complementary domains: (1) empirical realism, where knowledge is treated as objective, and (2) social constructivism, where shared understanding emerges through interaction (Brøns Kringelum & Brix, 2021). For a cyber resilience maturity model to be effective it must address both the knowledge itself and the social interactions required to create collectively understood knowledge and integrate it across the organization (Belinski et al., 2020; Raviv et al., 2022).

3.7 Maturity Models

Maturity models are ubiquitous across business and academia, numbering in the hundreds (Becker et al., 2009a; Monteiro & Maciel, 2020; Wendler, 2012). Maturity Models serve as a structured approach to sensemaking and as a knowledge artifact and they aid practitioners, expert groups, and scholars in achieving excellence, fostering collective expertise, and developing theories (Buckle, 2018; Oltra-Rodríguez et al., 2025). They define discrete maturity levels that represent qualitative and quantitative stages of progression for a defined class of objects, e.g., increasing organizational capability within a specific context (Becker et al., 2009a; Lasrado, 2018; Wendler, 2012).

Maturity models typically include: [a] levels (usually 3-6), [b] descriptions of each level, [c] number of dimensions, [d] factors for each dimension, and [e] performance descriptions for activities/elements at each level (Becker et al., 2009a). They are constructed as either staged fixed-level models, continuous fixed-level models, focus area models, or set-theoretic approach for maturity models (Monteiro & Maciel, 2020; Adekunle et al., 2022). As summarized by Lasrado (2018, p. 23): "[maturity

assessment models] represent theories about how organizational capabilities evolve in a stage-by-stage manner along an anticipated, desired, or logical maturation path".

At their core, maturity models assume that a desired future state (normative/prescriptive) can be achieved by effectively applying knowledge and following identified improvement steps from the current state (descriptive) (Kazanjian & Drazin, 1989; Pullen, 2007; Solli-Sæther & Gottschalk, 2010). In this way, they aim to support organizational development, process improvement, and benchmarking (internal or external) of the object under study. However, their effectiveness and ability to fulfill their intended purpose remain contentious (Adekunle et al., 2022; Buckle, 2018). Critiques often highlight challenges in generalization and validation, particularly in VUCA contexts (i.e., wicked problem) where rigor in design and validation is difficult to achieve (Adekunle et al., 2022; Becker et al., 2009b).

3.8 Conceptual Model

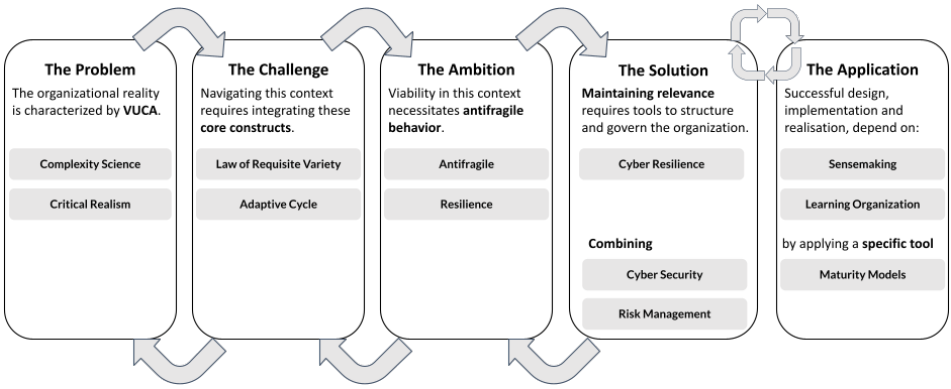

Figure 1: Conceptual model

Navigating within VUCA environments constitutes a wicked socio-technical problem that necessitates a deep understanding of the dynamic interplay between humans, machines, and their broader environment. Before formulating specific design guidelines it is essential to synthesize the theoretical foundations into a coherent structure that encompasses both the problem and solution spaces. Through the iterative exploration of these spaces during the *diagnosis phase* we developed a conceptual model (Figure 1) incorporating the core constructs

previously discussed. This model serves as a logical bridge to the design process: it informs the design goals and design principles of the proposed Cyber Resilience Maturity Model (CRMM) and establishes the foundation for the subsequent *design objectives and requirements phase*.

4 Design Goals and Principles

To answer our research question during the *diagnosis phase* we elicited the relevant constructs as part of a conceptual model (Figure 1, Section 3) of the problem and solution space. During the *design objectives and requirements phase* we formulate two design goals as design objective (Table 1), and 12 actionable design principles as requirements (Table 2) following the prescribed syntax.

4.1 Design Goals

Business continuity is not only about sustaining operations but also about maintaining relevance to all stakeholders. In complex environments static stability leads to obsolescence. Thus, organizations must dynamically adjust to preserve their relevance (e.g., Eling et al., 2021; Reeder and Hall, 2021; Saeed et al., 2023; Snowden, 2005). This we capture in Design Goal 01 (DG 01).

Reactive adaptation alone is insufficient for long-term survival. Organizations must convert volatility into value to remain relevant. In a context of continuous chaos the capacity to derive value from disorder requires antifragile behavior (e.g., Adobor, 2024; Axenie et al., 2024; Johnson and Gheorghe, 2013; Sagala and Őri, 2024). To stipulate this we formulated Design Goal 02 (DG 02).

Table 1: Design Goals

DG 01.	Support organizations in staying relevant for all stakeholders by adapting to change to ensure business continuity. *e.g., Eling et al., 2021; Reeder and Hall, 2021; Saeed et al., 2023; Snowden, 2005*
DG 02.	Support organizations in achieving antifragile behavior: to derive business value from the continuous chaos in and around them. *e.g., Adobor, 2024; Axenie et al., 2024; Johnson and Gheorghe, 2013; Sagala and Őri, 2024*

4.2 Design Principles

Design principles on supporting the user to envision how to uphold business continuity

Cyber resilience, just like cyber security and risk management, aims to preserve business continuity. This is why the first design principles are generic, defining the general function of the CRMM. Business continuity relies on managing known risks ("known unknowns", DP 01) while acknowledging the boundaries of our knowledge ("unknown unknowns", DP 02) (e.g., S. Ashby et al., 2018; Metcalf et al., 2021; Snowden, 2005; Spitz and Desbiey, 2025; Taleb, 2012; Weick et al., 2005). These domains contrast with the baseline of "known knowns". By recognizing the limits of what we know we mitigate the fragility that arises from false certainty. Upholding business continuity starts with preparing the organization to deal with these two domains. How the risks can be identified and converted into action is addressed in more detail in the other design principles.

Design principles on supporting the user in conserving the current situation

One way to maintain business continuity is to preserve the current situation. Engineering Resilience, Systems Resilience, and standard risk management frameworks (including cybersecurity) focus on identifying risks (DP 03) and protecting against negative impacts (DP 04) (e.g., Aven, 2018; Dupont et al., 2023; Eling et al., 2021; Florea and Florea, 2016; Mizrak, 2023). This approach aims to attenuate variety and aligns with the conservation phase of the adaptive cycle.

Design principles on supporting the user in improving the current situation

Value creation requires mechanisms for exploration, reorganization, and the exploitation phases of the adaptive cycle. This involves exploring upside risk and opportunities (DP 05) and capitalizing on them (DP 06), amplifying variety to enable Complex Adaptive Systems (CAS) Resilience (e.g., Adobor and Kudonoo, 2025; Sundstrom and Allen, 2019; Walker et al., 2004).

Design principles on supporting the user in improving sensemaking

Central to all design principles is the requirement for user agency and emancipation, specifically empowering individuals to determine their actions (DP 07) (e.g., Bennett and Lemoine, 2014; Mutch, 2010; Weick et al., 2005). However, in socially interconnected organizations individual agency must align with coordinated collaboration (DP 08) (e.g., Dupont et al., 2023; Snowden, 2005; Thompson, 1995). Through collective coordination, organizations can intentionally reshape their context (DP 09), enabling double-loop organizational learning and structural adaptation (e.g., Adekunle et al., 2022; Brøns Kringelum and Brix, 2021; Lasrado, 2018).

Design principles on supporting the user in improving learning

Learning includes an increased understanding of the interplay between objective reality (empirical realism) and inter-subjective reality (social constructivism). These perspectives define the scope of both the individual user's practice of personal mastery (DP 10) (e.g., Belinski et al., 2020; Brøns Kringelum and Brix, 2021; Raviv et al., 2022) as well as the team's learning discipline (DP 11) (e.g, Belinski et al., 2020; Brøns Kringelum and Brix, 2021; Raviv et al., 2022).

Design principles on supporting the user in becoming antifragile

The cultivation of learning disciplines relies on the response to volatility. While Engineering and Systems Resilience seek to resist shocks and CAS Resilience aims to adapt, antifragility extends these capacities to gain value from disorder and stress. Because novelty resides in the unknown, shielding individuals from exposure leads to systemic fragility. Therefore, to transition from surviving to thriving, individuals need support in their transition from acknowledging (DP 02) to actively embracing exposure to the unknown (DP 12) (e.g., Beer, 1974; Sundstrom and Allen, 2019; Taleb, 2012). This engagement with volatility generates learning signals, which increase the opportunities of exploiting opportunities, and consequently securing the upside of uncertainty.

Table 2: Design Principles

	Design principles on supporting the user to envision how to uphold business continuity
DP 01	Support individuals (actors) to determine how to prepare their organization (users) to uphold business continuity (aim) when their organization is confronted with known unknown events (context). *e.g., S. Ashby et al., 2018; Metcalf et al., 2021; Snowden, 2005; Weick et al., 2005*
DP 02	Support individuals (actors) to determine how to prepare their organization (actors) to uphold business continuity (aim) when their organization is confronted with unknown unknown events (context). *e.g., Metcalf et al., 2021; Snowden, 2005; Spitz and Desbiey, 2025; Taleb, 2012*
	Design principles on supporting the user in conserving the current situation
DP 03	Support individuals (actors) to uphold business continuity by identifying downside risks, also known as known threats, (aim) of their organization (user). *e.g., Aven, 2018; Dupont et al., 2023; Florea and Florea, 2016*
DP 04	Support individuals (actors) to limit the negative impact of materialized threats, also known as risk events, on people, process and technology (aim) of their organization (user). *e.g., Eling et al., 2021; Mizrak, 2023*
	Design principles on supporting the user in improving the current situation
DP 05	Support individuals (actors) to uphold business continuity by identifying upside risks, also known as known opportunities, (aim) of their organization (user). *e.g., Sundstrom and Allen, 2019; Walker et al., 2004*
DP 06	Support individuals (actors) in exploiting opportunities, as changing the function and construction of a system to optimize the value proposition in a new reality (aim), in their organization (user). *e.g., Adobor and Kudonoo, 2025; Sundstrom and Allen, 2019; Walker et al., 2004*
	Design principles on supporting the user in improving sensemaking
DP 07	Support individuals (actor, user) in their sensemaking to determine their next action (aim). *e.g., Bennett and Lemoine, 2014; Mutch, 2010; Weick et al., 2005*
DP 08	Support individuals (actors) in their sensemaking to determine their next action (aim) as part of their organization (user). *e.g., Dupont et al., 2023; Snowden, 2005; Thompson, 1995*
DP 09	Support individuals (actors) to determine the growth-path from the current state to a desired state (aim) of their organization (user). *e.g., Adekunle et al., 2022; Brøns Kringelum and Brix, 2021; Lasrado, 2018*
	Design principles on supporting the user in improving learning
DP 10	Support individuals (actor, users) in their learning (aim) from an empirical realist and a social constructivist viewpoint (mechanism). *e.g., Belinski et al., 2020; Brøns Kringelum and Brix, 2021; Raviv et al., 2022*
DP 11	Support individuals (actors) as part of the group (user), in their organizational learning (aim) from an empirical realist and a social constructivist viewpoint (mechanism). *e.g., Belinski et al., 2020; Brøns Kringelum and Brix, 2021; Raviv et al., 2022*
	Design principles on supporting the user in becoming antifragile
DP 12	Support individuals (actors) as part of an organization (user) in embracing exposure to the unknown to increase the possibilities of exploiting opportunities and learning (aim). *e.g., Beer, 1974; Sundstrom and Allen, 2019; Taleb, 2012*

Collectively these design principles address the reduction of fragility (by acknowledging the unknown), the improvement of resilience (by limiting the impact of the unknown), adaption (by navigating the unknown) and the achievement of antifragility (feeding on the unknown); these are summarized below in Table 2.

5 Conclusion, Discussion, and Future Research

Stakeholders in the cyber resilience domain require a sensemaking artifact to navigate complexity. Existing frameworks and models in cyber resilience and related fields (such as risk management and cybersecurity) focus primarily on risk prediction and mitigation, but they often overlook organizational learning and value creation. Consequently, these frameworks and models fail to foster antifragility, as they do not account for the potential to grow from disruptions. This mismatch leaves stakeholders ill-equipped to address the VUCA nature of their environments. This requires new cyber resilience artifacts (e.g., maturity models) for which we proposed design goals (Table 1) and design principles (Table 2) in our research.

The design goals and principles are intentionally generic to ensure a broad applicability across domains such as cybersecurity, business continuity, risk management, and enterprise governance. This high-level approach addresses a common critique of specific knowledge artifacts (such as rigid risk processes) which lose value when the organizational context shifts. By prioritizing generic principles before building specific artifacts we ensure the model remains relevant across diverse environments. This generic approach fits the statement variety eats variety from the law of requisite variety (Figure 1). The design goals and design principles therefore support the actors and users during their iterations of the adaptive cycle phases, enabling them to absorb the shifting variety within both internal and external organizational contexts.

The practical implication of this approach is that the model that results from the suggested design principles should be agnostic to changes in the (shared) reality of what individuals perceive as real and should even support these changes. The set of design goals and principles also affects existing risk management and cybersecurity frameworks since none of the existing frameworks meets all these requirements. Therefore, these frameworks are fragile in the face of changes in context and changes within the user and actors.

A key limitation of our research is the lack of empirical validation. We have not yet interviewed various experts and practitioners to validate the proposed design goals and principles. In future research we will attempt to validate the design goals and principles with experts. Additionally we will continue our literature review to catalog relevant cyber resilience frameworks in an open database, combined with insights into their coverage of our proposed design principles. Furthermore we seek to foster collaboration on this database among researchers and practitioners in the cyber security and risk management domains with the goal of creating a validated dataset and an openly available body of knowledge. Concurrently, we will develop a new Cyber Resilience Maturity Model (CRMM) to validate the operationalization of the validated design principles.

References

Adekunle, S. A., Aigbavboa, C., Ejohwomu, O., Ikuabe, M., & Ogunbayo, B. (2022). A critical review of maturity model development in the digitisation era [Publisher: MDPI AG Publisher: Multidisciplinary Digital Publishing Institute]. Buildings, 12(6), 858. https://doi.org/10.3390/buildings12060858

Adobor, H. (2024). How organizations can benefit from volatility: The promise of antifragility and some cautionary notes. Organizational Dynamics, 101098. https://doi.org/10.1016/j.orgdyn.2024.101098

Adobor, H., & Kudonoo, E. C. (2025). Antifragility and organizations: An organizational design perspective. Leadership & Organization Development Journal. https://doi.org/10.1108/LODJ-03-2024-0185

Alashwal, A. M., Abdul-Rahman, H., & Asef, A. (2017). Influence of organizational learning and firm size on risk management maturity [Publisher: American Society of Civil Engineers (ASCE)]. Journal of Management in Engineering, 33(6), 04017034. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000553

Alrumaih, T. N., Alenazi, M. J., AlSowaygh, N. A., Humayed, A. A., & Alablani, I. A. (2023). Cyber resilience in industrial networks: A state of the art, challenges, and future directions [Publisher: Elsevier BV]. Journal of King Saud University - Computer and Information Sciences, 35(9), 101781. https://doi.org/10.1016/j.jksuci.2023.101781

Al-Shehab, A. J., Hughes, R. T., & Winstanley, G. (2005). Facilitating organisational learning through causal mapping techniques in IS/IT project risk management [ISSN: 0302-9743 Journal Abbreviation: Wissensmanagement]. In K.-D. Althoff, A. Dengel, R. Bergmann, M. Nick, & T. Roth-Berghofer (Eds.), Professional knowledge management (pp. 145–154, Vol. 3782). Springer Berlin Heidelberg. https://doi.org/10.1007/11590019_18

Argyris, C., & Schön, D. A. (1978). Organizational learning: A theory of action perspective. Addison-Wesley.

Ashby, S., Buck, T., Nöth-Zahn, S., & Peisl, T. (2018). Emerging IT risks: Insights from german banking [Publisher: Springer Science and Business Media LLC]. The Geneva Papers on Risk and Insurance – Issues and Practice, 43(2), 180–207. https://doi.org/10.1057/s41288-018-0081-8

Ashby, W. R. (1958). Requisite variety and its implications for the control of complex systems [ISBN: 978-1489907202 Publisher: Springer US]. Cybernetica, 1(2), 83–99. https://doi.org/10.1007/978-1-4899-0718-9_28

Aven, T. (2015a). Implications of black swans to the foundations and practice of risk assessment and management [Publisher: Elsevier]. Reliability Engineering & System Safety, 134, 83–91. https://doi.org/10.1016/j.ress.2014.10.004

Aven, T. (2015b). The concept of antifragility and its implications for the practice of risk analysis [Publisher: Wiley]. Risk Analysis, 35(3), 476–483. https://doi.org/10.1111/risa.12279

Aven, T. (2018). The call for a shift from risk to resilience: What does it mean? [Publisher: Wiley]. Risk Analysis, 39(6), risa.13247. https://doi.org/10.1111/risa.13247

Axenie, C., López-Corona, O., Makridis, M. A., Akbarzadeh, M., Saveriano, M., Stancu, A., & West, J. (2024). Antifragility in complex dynamical systems [Publisher: Springer Science and Business Media LLC]. Nature Partner Journals (NPJ) Complexity, 1(12). https://doi.org/10.1038/s44260-024-00014-y

Baskerville, R., & Pries-Heje, J. (2010). Explanatory design theory. Business & Information Systems Engineering, 2(5), 271–282. https://doi.org/10.1007/s12599-010-0118-4

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009a). Developing maturity models for IT management: A procedure model and its application [Publisher: Springer Nature]. Business & Information Systems Engineering, 1(3), 213–222. https://doi.org/10.1007/s12599-009-0044-5

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009b). Entwicklung von Reifegradmodellen für das IT-Management: Vorgehensmodell und praktische Anwendung [Publisher: Springer Nature]. WIRTSCHAFTSINFORMATIK, 51(3), 249–260. https://doi.org/10.1007/s11576-009-0167-9 Beer, S. (1974, January). Designing freedom (Vol. 27) [ISSN: 0030-3623 Publication Title: Operational Research Quarterly (1970-1977)]. John Wiley & Sons. https://doi.org/10.2307/3008834

Belinski, R., Peixe, A. M., Frederico, G. F., & Garza-Reyes, J. A. (2020). Organizational learning and industry 4.0: Findings from a systematic literature review and research agenda [Publisher: Emerald]. Benchmarking: An International Journal, 27(8), 2435–2457. https://doi.org/10.1108/bij-04-2020-0158

Bennett, N., & Lemoine, G. J. (2014). What a difference a word makes: Understanding threats to performance in a VUCA world. [Publisher: Elsevier BV]. Business Horizons, 57(3), 311–317. https://doi.org/10.2139/ssrn.2406676

Bharathy, G. K., & McShane, M. K. (2014). Applying a systems model to enterprise risk management [Publisher: RELX Group (Netherlands)]. Engineering Management Journal, 26, 38–46. https://doi.org/10.1080/10429247.2014.11432027

Boeing, G. (2016). Visual analysis of nonlinear dynamical systems: Chaos, fractals, self-similarity and the limits of prediction [Publisher: Multidisciplinary Digital Publishing Institute]. Systems, 4(4), 37. https://doi.org/10.3390/systems4040037

Botjes, E. A., Huygh, T., Bollen, L., & Helms, R. (2026a, January 21). Cyber resilience maturity model keywords. https://doi.org/10.5281/ZENODO.18323583

Botjes, E. A., Huygh, T., Bollen, L., & Helms, R. (2026b, February 4). Open research infrastructure - CRMM. Zenodo. https://doi.org/10.5281/ZENODO.18487444

Brøns Kringelum, L., & Brix, J. (2021). Critical realism and organizational learning [Publisher: Emerald Publishing Limited]. The Learning Organization, 28(1), 32–45. https://doi.org/10.1108/TLO-03-2020-0035

Buckle, P. (2018). Maturity models for systems thinking [Publisher: MDPI AG]. Systems, 6(2), 23. https://doi.org/10.3390/systems6020023

Caon, M., Panetti, E., Meier, E., & Baldegger, R. J. (2025). The dual effect of COVID-19: Diverse digitalization approaches in micro- and small businesses. Journal of the International Council for Small Business, 6(2), 216–228. https://doi.org/10.1080/26437015.2024.2404184

Carroll, J. (2024). The u.s. national cybersecurity strategy: A vehicle with an international journey [Publisher: Academic Conferences International Ltd]. European Conference on Cyber Warfare and Security, 23(1), 107–115. https://doi.org/10.34190/eccws.23.1.2300

Cerqueira Junior, A. S., & Arima, C. H. (2023). Cyber risk management and ISO 27005 applied in organizations: A systematic literature review [Publisher: Faculdade Novo Milênio]. REVISTA FOCO, 16(2), e1188. https://doi.org/10.54751/revistafoco.v16n2-215

Chiara, P. G. (2022). The cyber resilience act: The EU commission's proposal for a horizontal regulation on cybersecurity for products with digital elements: An introduction. International Cybersecurity Law Review, 3(2), 255–272. https://doi.org/10.1365/s43439-022-00067-6

Derbyshire, J., & Wright, G. (2014). Preparing for the future: Development of an 'antifragile' methodology that complements scenario planning by omitting causation. [Publisher: Elsevier BV]. Technological Forecasting and Social Change, 82, 215–225. https://doi.org/10.1016/j.techfore. 2013.07.001

Dupont, B., Shearing, C., Bernier, M., & Leukfeldt, R. (2023). The tensions of cyber-resilience: From sensemaking to practice. Computers & Security, 132, 103372. https://doi.org/10.1016/j.cose.2023.103372

Eling, M., McShane, M., & Nguyen, T. (2021). Cyber risk management: History and future research directions [Publisher: Blackwell Publishing Ltd]. Risk Management and Insurance Review, 24(1), 93–125. https://doi.org/10.1111/rmir.12169

Fischer, C., Winter, R., & Wortmann, F. (2010). Design theory. Business & Information Systems Engineering, 2(6), 387–390. https://doi.org/10.1007/s12599-010-0128-2

Fjaeran, L., & Aven, T. (2021). Creating conditions for critical trust – how an uncertainty-based risk perspective relates to dimensions and types of trust [Publisher: Elsevier BV]. Safety Science, 133, 105008. https://doi.org/10.1016/j.ssci.2020.105008

Florea, R., & Florea, R. (2016). Internal audit and risk management. ISO 31000 and ERM approaches. Economy Transdisciplinarity Cognition, 19(1), 72–77. https : / / www. ugb . ro / etc / etc2016no1 / 13 _ Florea _Radu__Florea_Ramona.PDF

Gregor, S. (2006). The nature of theory in information systems. MIS Quarterly, 30(3), 611–642. https://doi.org/10.2307/25148742

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. MIS Quarterly, 37(2), 337–355. https://doi.org/10.25300/MISQ/2013/37.2.01

Gregor, S., & Jones, D. (2007). The anatomy of a design theory. Journal of the Association for Information Systems, 8(5), 312–335. https://doi.org/10.17705/1jais.00129

Gregor, S., Kruse, L., & Seidel, S. (2020). Research perspectives: The anatomy of a design principle. Journal of the Association for Information Systems, 21, 1622–1652. https://doi.org/10.17705/1jais.00649

Herley, C. (2016). Unfalsifiability of security claims. Proceedings of the National Academy of Sciences, 113(23), 6415–6420. https://doi.org/10.1073/pnas.1517797113

Hevner, A. R., Parsons, J., Brendel, A. B., Lukyanenko, R., Tiefenbeck, V., Tremblay, M. C., & Vom Brocke, J. (2024). Transparency in design science research. Decision Support Systems, 182, 114236. https://doi.org/10.1016/j.dss.2024.114236

Heylighen, F. (2002). The science of self-organization and adaptivity. In L. D. Kiel (Ed.), Knowledge management, organizational intelligence and learning, and complexity. (Vol. 3). EOLSS Publishers Co Ltd. https://researchportal.vub.be/en/publications/the- science- of- self-organization-and-adaptivity

Hillmann, J., & Guenther, E. (2021). Organizational resilience: A valuable construct for management research? International Journal of Management Reviews, 23(1), 7–44. https://doi.org/10.1111/ijmr.12239

Hosseini, S., Barker, K., & Ramirez-Marquez, J. E. (2016). A review of definitions and measures of system resilience [Publisher: Elsevier BV]. Reliability Engineering & System Safety, 145, 47–61. https://doi.org/10.1016/j.ress.2015.08.006

Huang, Y., Huang, L., & Zhu, Q. (2022). Reinforcement learning for feedback-enabled cyber resilience [Publisher: Elsevier BV]. Annual 20 Reviews in Control, 53, 273–295. https://doi.org/10.1016/j.arcontrol.2022.01.001

IMF. (2024, April). Cyber risk: A growing concern for macrofinancial stability (chapter 3). In Global financial stability report, april 2024: The last mile: Financial vulnerabilities and risks (p. 26). International Monetary Fund. https://doi.org/10.5089/9798400257704.082

Imran, H., Salama, M., Turner, C., & Fattah, S. (2022). Cybersecurity risk management frameworks in the oil and gas sector: A systematic literature review [ISSN: 2367-3389 Series Title: Lecture Notes in Networks and Systems]. In K. Arai (Ed.), Advances in information and communication (pp. 871–894, Vol. 439). Springer International Publishing. https://doi.org/10.1007/978-3-030-98015-3_59

Johnson, J., & Gheorghe, A. V. (2013). Antifragility analysis and measurement framework for systems of systems [Publisher: Springer Science and Business Media LLC]. International Journal of Disaster Risk Science, 4(4), 159–168. https://doi.org/10.1007/s13753-013-0017-7

Kazanjian, R. K., & Drazin, R. (1989). An empirical test of a stage of growth progression model. Management Science, 35(12), 1489–1503. https://doi.org/10.1287/mnsc.35.12.1489

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: Anatomy of a research project. European Journal of Information Systems, 17(5), 489–504. https://doi.org/10.1057/ejis.2008.40

Lasrado, L. A. (2018). Set-theoretic approach to maturity models [Doctoral dissertation, Copenhagen Business School] [ISSN:0906-6934 OCLC: 1175940135 ISBN: 978-87-93579-77-4, 978-87-93579-76-7]. https://www.econstor.eu/bitstream/10419/209064/1/cbs-phd2018-15.pdf

Linkov, I., & Kott, A. (2019). Fundamental concepts of cyber resilience: Introduction and overview [Journal Abbreviation: Cyber Resilience of Systems and Networks]. In A. Kott & I. Linkov (Eds.), Cyber resilience of systems and networks (pp. 1–25, Vol. abs/1806.02852). Springer International Publishing. https://doi.org/10.1007/978-3-319-77492-3_1

Lorenz, E. N. (1963). Deterministic nonperiodic flow [Publisher: American Meteorological Society]. Journal of the atmospheric sciences, 20(2), 130–141. https://journals.ametsoc.org/downloadpdf/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.pdf

Lukyanenko, R., & Parsons, J. (2020). Research perspectives: Design theory indeterminacy: What is it, how can it be reduced, and why did the polar bear drown? Journal of the Association for Information Systems, 21(5), 1343–1369. https://doi.org/10.17705/1jais.00639

Lusch, R. F., & Nambisan, S. (2015). Service innovation: A service-dominant logic perspective. MIS Quarterly, 39(1), 155–175. https://doi.org/10.25300/MISQ/2015/39.1.07

Maedche, A., Elshan, E., Höhle, H., Lehrer, C., Recker, J., Sunyaev, A., Sturm, B., & Werth, O. (2024). Open science: Towards greater transparency and openness in science [Publisher: Springer Nature]. Business & Information Systems Engineering. https://doi.org/10.1007/s12599-024-00858-7

Markus, M. L., Majchrzak, A., & Gasser, L. (2002). A design theory for systems that support emergent knowledge processes. MIS Quarterly, 26(3), 179–212. https://aisel.aisnet.org/misq/vol26/iss3/4/

Martin-Breen, P., & Anderies, J. M. (2011). The bellagio initiative, background paper, resilience: A literature review [event-place: Brighton:IDS]. Resilience: A Literature Review. http://opendocs.ids.ac.uk/opendocs/handle/123456789/3692

Metcalf, G. S., Kijima, K., & Deguchi, H. (Eds.). (2021). Handbook of systems sciences. Springer Singapore. https://doi.org/10.1007/978- 981- 15-0720-5

Mizrak, F. (2023). Integrating cybersecurity risk management into strategic management: A comprehensive literature review. Pressacademia, 3. https://doi.org/10.17261/Pressacademia.2023.1807

Monteiro, E. L., & Maciel, R. S. P. (2020). Maturity models architecture: A large systematic mapping [Publisher: Sociedade Brasileira de Computacao - SB]. iSys - Brazilian Journal of Information Systems, 13(2), 110–140. https://doi.org/10.5753/isys.2020.761

Mukumbang, F. C., De Souza, D. E., & Eastwood, J. G. (2023). The contributions of scientific realism and critical realism to realist evaluation [Publisher: Taylor & Francis]. Journal of Critical Realism, 22(3), 504–524. https://doi.org/10.1080/14767430.2023.2217052

Mullarkey, M. T., & Hevner, A. R. (2019). An elaborated action design research process model (P. Ågerfalk, Ed.) [Publisher: Informa UK Limited]. European Journal of Information Systems, 28(1), 6–20. https://doi.org/10.1080/0960085X.2018.1451811

Munusamy, T., & Khodadi, T. (2023). Building cyber resilience: Key factors for enhancing organizational cyber security [Publisher: Multimedia University]. Journal of Informatics and Web Engineering, 2(2), 59–71. https://doi.org/10.33093/jiwe.2023.2.2.5

Mutch, A. (2010). Technology, organization, and structure—a morphogenetic approach [Publisher: Institute for Operations Research and the Management Sciences (INFORMS)]. Organization science, 21(2), 507–520.

NIST. (2023, August 8). The NIST cybersecurity framework 2.0 (NIST CSWP 29 ipd) (Author long name: National Institute of Standards and Technology; Backup Publisher: National Institute of Standards and Technology). NIST. Gaithersburg, MD. https://doi.org/10.6028/NIST.CSWP.29.ipd

Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company. How japanese companies create the dynamics of innovation. Oxford University Press.

Oltra-Rodríguez, M. A., Stonehouse, P., Afonso-Alonso, N., & Holgado-Terriza, J. A. (2025). Disciplined delivery and organizational design maturity: A socio-technical evolutionary journey. Systems, 13(5), 374. https://doi.org/10.3390/systems13050374

Park, J., Seager, T. P., Rao, P. S. C., Rao, P. S. C., Rao, P. S. C., Rao, P. S. C., Convertino, M., Linkov, I., Linkov, I., Linkov, I., & Linkov, I. (2013). Integrating risk and resilience approaches to catastrophe management in engineering systems [Publisher: Wiley]. Risk Analysis, 33(3), 356–367. https://doi.org/10.1111/j.1539-6924.2012.01885.x

Parmar, B. L., Freeman, R. E., Harrison, J. S., Wicks, A. C., Purnell, L., & De Colle, S. (2010). Stakeholder theory: The State of the Art [Publisher: Routledge]. Academy of Management Annals, 4(1), 403–445. https://doi.org/10.5465/19416520.2010.495581

Pullen, W. (2007). A public sector HPT maturity model. Performance Improvement, 46(4), 9–15. https://doi.org/10.1002/pfi.119

Putra, A. P., & Soewito, B. (2023). Integrated methodology for information security risk management using ISO 27005:2018 and NIST SP 800-30 for insurance sector [Publisher: Science and Information Organization]. International Journal of Advanced Computer Science and Applications, 14(4). https://doi.org/10.14569/IJACSA.2023.0140468

Raviv, L., Lupyan, G., & Green, S. C. (2022). How variability shapes learning and generalization [Publisher: Elsevier BV]. Trends in Cognitive Sciences, 26(6), 462–483. https://doi.org/10.1016/j.tics.2022.03.007

Recker, J. (2021). Scientific research in information systems: A beginner's guide [ISSN: 2196-8705 Publication Title: Progress in IS]. Springer International Publishing. https://doi.org/10.1007/978-3-030-85436-2

Reeder, J. R., & Hall, T. (2021). Cybersecurity's pearl harbor moment: Lessons learned from the colonial pipeline ransomware attack. The Cyber Defense Review, 6(3), 15–40. https://www.jstor.org/stable/48631153

Rivard, S. (2021). Theory building is neither an art nor a science. it is a craft. Journal of Information Technology, 36(3), 316–328. https://doi.org/10.1177/0268396220911938

Ruiz-Martín, C., Lopez-Paredes, A., & Wainer, G. (2018). What we know and do not know about organizational resilience [Publisher: Universitat Politecnica de Valencia]. International Journal of Production Management and Engineering, 6(1), 11 https://doi.org/10.4995/ijpme.2018.7898

Saeed, S., Altamimi, S. A., Alkayyal, N. A., Alshehri, E., & Alabbad, D. A. (2023). Digital transformation and cybersecurity challenges for businesses resilience: Issues and recommendations. Sensors, 23(15), 6666. https://doi.org/10.3390/s23156666

Sagala, G. H., & Őri, D. (2024). Antifragility, resilience and collaborative networks of SMEs: A theoretical foundation [Publisher: Emerald Publishing Limited]. European Journal of Innovation Management. https://doi.org/10.1108/EJIM-09-2023-0797

Sepúlveda Estay, D. A., Sahay, R., Barfod, M. B., & Jensen, C. D. (2020). A systematic review of cyber-resilience assessment frameworks [Publisher: Elsevier BV]. Computers & Security, 97, 101996. https://doi.org/10.1016/j.cose.2020.101996

Smeets, M. E. (2021, August). Fear drop and fear change. perceptions of security in the 21st century: Their formation, trends, and impact in society [Doctoral dissertation, University of Cardiff].

Snowden, D. (2005). Strategy in the context of uncertainty [Publisher: Emerald]. Handbook of Business Strategy, 6(1), 47–54. https://doi.org/10.1108/08944310510556955

Solli-Sæther, H., & Gottschalk, P. (2010). The modeling process for stage models [Publisher: Taylor & Francis]. Journal of Organizational Computing and Electronic Commerce, 20(3), 279–293. https://doi.org/10.1080/10919392.2010.494535

Spitz, R., & Desbiey, O. (2025). The future of risk and insurability in the era of systemic disruption, unpredictability and artificial intelligence. The Journal of Operational Risk. https://doi.org/10.21314/JOP.2025.003

Steen, R., Haug, O. J., & Patriarca, R. (2024). Business continuity and resilience management: A conceptual framework. Journal of Contingencies and Crisis Management, 32(1), e12501. https://doi.org/10.1111/1468-5973.12501

Sundstrom, S. M., & Allen, C. R. (2019). The adaptive cycle: More than a metaphor. Ecological Complexity, 39, 100767. https://doi.org/10.1016/j.ecocom.2019.100767

Taherdoost, H. (2022). Cybersecurity vs. information security [Publisher: Elsevier BV]. Procedia Computer Science, 215, 483–487. https://doi.org/10.1016/j.procs.2022.12.050

Taleb, N. N. (2012, November). Antifragile: Things that gain from disorder [ISSN: 1059-5422 Publication Title: Competitiveness Review: An International Business Journal]. Random House. https://doi.org/10.1108/10595421311319852

Tasheva, I., & Kunkel, I. (2022). In a hyperconnected world, is the EU cybersecurity framework connected? [Publisher: SAGE Publications]. European View, 21(2), 186–195. https://doi.org/10.1177/17816858221136106

The World Bank. (2023, June). Sectoral cybersecurity maturity model (SCMM). Worldbank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099062623085028392/p17263707c36b702309f7303dbb7266e1cf

Thekdi, S., & Aven, T. (2024, October 7). Decoding black swans and other historic risk events: Themes of progress and opportunity for risk science (1st ed.). CRC Press. https://doi.org/10.1201/9781003437031

Thompson, F. (1995). Business strategy and the boyd cycle [Publisher: Wiley]. Journal of Contingencies and Crisis Management, 3(2), 81–90. https://doi.org/10.1111/j.1468-5973.1995.tb00060.x

Tuunanen, T., Winter, R., & Vom Brocke, J. (2024). Dealing with complexity in design science research: A methodology using design echelons. MIS Quarterly, 48(2), 427–458. https://doi.org/10.25300/MISQ/2023/16700

Tzavara, V., & Vassiliadis, S. (2024). Tracing the evolution of cyber resilience: A historical and conceptual review [Publisher: Springer Science and Business Media LLC]. International Journal of Information Security, 23(3), 1695–1719. https://doi.org/10.1007/s10207-023-00811-x

Vom Brocke, J., Winter, R., Hevner, A., & Maedche, A. (2020). Special issue editorial – accumulation and evolution of design knowledge in design science research: A journey through time and space. Journal of the Association for Information Systems, 21(3), 520–544. https://doi.org/10.17705/1jais.00611

Vostoupal, J., Stupka, V., Harašta, J., Kasl, F., Loutocký, P., & Malinka, K. (2023). The legal aspects of cybersecurity vulnerability disclosure: To the NIS 2 and beyond. https://doi.org/10.2139/ssrn.4640775

Walker, B. H., Holling, C. S., Carpenter, S. R., & Kinzig, A. (2004). Resilience, adaptability and transformability in social–ecological systems [Publisher: The Resilience Alliance]. Ecology and society, 9(2).

Walls, J. G., Widmeyer, G. R., & Sawy, O. A. E. (1992). Building an information system design theory for vigilant EIS. Information Systems Research, 3(1), 36–59. http://www.jstor.org/stable/23010780

Weick, K. E. (1995). Sensemaking in organizations. Sage Publications.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. Organization Science, 16(4), 409–421. https://doi.org/10.1287/orsc.1050.0133

Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study [Publisher: Elsevier BV]. Information and Software Technology, 54(12), 1317–1339. https://doi.org/10.1016/j.infsof.2012.07.007

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3(1), 160018. https://doi.org/10.1038/sdata.2016.18

About the authors

Edzo A. Botjes MSc is a PhD candidate in cyber resilience at the Open Universiteit, the Netherlands. His research interests are on how to improve business continuity and focuses on the overlap between risk management, cyber security, enterprise architecture, enterprise governance, resilience, and is specialised in antifragility. He is supervising several MSc students at the Antwerp Management School, and is an external teacher for the MSc course Enterprise Architecture at the University of Applied Sciences Utrecht. His research activities resulted in (cited) publications.

Prof.dr.ir. **Remko Helms** is a professor in Business Analytics at Open Universiteit (The Netherlands) in the Information Systems group. His teaching and research interests are mainly focused on how organizations realize value from exploring and exploiting data, including topics such as data analytics processes, data analytics strategy, data governance and analytics governance. He is supervising several Phd students and groups of master students on these topics. His work appeared in different Information Systems and Knowledge Management conferences as well as journals. He is a member of Association for Information Systems (AIS).

Dr. **Laury Bollen** is an associate professor of Information Management at the Open Universiteit, the Netherlands. He holds a PhD in Business Economics from Maastricht University. His research interest focuses on IT Governance, Information Security Governance and Platform Governance.
His research activities have resulted in publications in a broad range of international academic journals including Information & Management, the Journal of Information Systems, Information Systems Management, the Journal of Cleaner Production, the European Accounting Review, and the Journal of Business Ethics.

Dr. **Tim Huygh** is assistant professor at the department of information science at the faculty of science of the Open Universiteit, the Netherlands. He is also a visiting professor at the Antwerp Management School. He received his PhD in 2019 on IT governance from the University of Antwerp. His research interests include the governance and (strategic) management of information and technology, information security governance, and cyber resilience. His research has been published in SCI-indexed journals like Information Systems Journal (ISJ) and Decision Support Systems (DSS), and various conference proceedings including the International Conference on Information Systems (ICIS) and the Hawaii International Conference on System Sciences (HICSS). He also co-authored three books on the topic of IT governance (published by Springer). Since 2020, he is co-chairing the minitrack "IT governance and its mechanisms" at HICSS.

Understanding Trust in GenAI: The Roles of Reliability, Explanation Quality, and Confidence in Judgment

Fatima Bilal, Anssi Öörni, Jozsef Mezei
Åbo Akademi University, Turku, Finland
fatima.bilal@abo.fi, anssi.oorni@abo.fi, jozsef.mezei@abo.fi

Generative Artificial Intelligence (GenAI) is increasingly utilised in academic contexts; however, concerns about reliability and transparency raise questions about user trust and adoption. Drawing on Trust in Technology Theory, this study examines how perceived reliability and explanation quality influence trust in GenAI and whether trust mediates their effects on intention to use AI in an academic context. This research study also examines the role of confidence in judgment as a user-level factor, using survey-based data from 262 participants and hierarchical regression with a mediation analysis. The results show that perceived reliability and explanation quality significantly enhance trust in AI. However, trust partially mediates their relationship with intention to use GenAI. In contrast, confidence in judgment does not significantly predict trust or intention, highlighting a conceptual distinction between users' self-confidence and trust in AI systems. The findings highlight the importance of perceived reliability and explanation quality in fostering calibrated trust and responsible adoption of GenAI by students and teachers in educational settings.

DOI
https://doi.org/10.18690/um.fov.4.2026.25

ISBN
978-961-299-152-4

Keywords:
trust in AI, perceived reliability, explanation quality, confidence in judgment, adoption intention, generative ai (gen AI)

University of Maribor Press

1 Introduction

Generative Artificial Intelligence (GenAI) is defined as an advanced branch of Artificial Intelligence (AI) that includes creating and refining content, including text, music, and images, based on the input provided by the user or data that is accessible from the internet. GenAI utilises deep learning models for the identification of structures and patterns of large datasets and produces content and results by using large dataset patterns (Miranda et al., 2025). After its emergence in the year 2018, GenAI has evolved and been adopted in academic settings and professional contexts as systems based on large language models are now used for decision support, generating ideas, summarising input, drafting, and producing final results. The new and improved algorithms of AI, including reinforcement learning and deep learning, are able to process data in natural form and are capable of mining raw and unstructured data, including images and text (Dwivedi et al., 2023). Considering the academic settings, the use of GenAI has been widely adopted and has transformed academic learning; however, GenAI offers both opportunities and challenges in the development of students, impacting the level of users' trust in GenAI (Vieriu et al., 2025).

Trust is characterised as the key factor in human-AI interaction, as it refers to users' expectations considering how the AI system will generate results and outputs under risks on uncertain conditions (Vereschak et al., 2021). Past researchers (Lee & See, 2004; Lockey et al., 2021) have identified and discussed key factors of users' trust in technology by reviewing empirical methodologies, including reliability and explainability of the system. In an educational context and setting, the factor of AI explanation quality is important as students rely on AI-based output and results instead of its critical evaluation. The current literature on the use of GenAI has emphasised that students need to have a clear understanding and capability for successful use of GenAI in learning (Bećirović et al., 2025). However, the existing research on GenAI only lacks integration of user-level and system-level factors and examines trust cues solely. Similarly, the current research, including (Ursavaş, 2025; Zhou, 2024), is based on limited testing and examination of explanation quality, perceived reliability, and users' trust and confidence in GenAI, which also creates doubt in comprehension and understanding of the role of users' confidence. Existing empirical studies on trust in GenAI, including Rittenberg & Holland (2024), suggested that a higher level of self-confidence of AI users might result in calibrated

trust in automated systems. For addressing this gap, this research study aims to investigate how perceived reliability and explanation quality impact trust in GenAI and how trust mediates the impact of perceived reliability and explanation quality in terms of intention to use GenAI. This research also examines how confidence in AI judgments predicts the usage intention and trust of users.

2 Theoretical Background

2.1 Trust in Generative AI in Academic Contexts

As stated by the study of Lim et al. (2023), GenAI generates humanised content in response to the input of prompts and commands by the user by using large-scale deep learning models. In the academic context, the use of GenAI tools is significantly increasing for writing assessments, carrying out research synthesis, and retrieving information. However, the output produced by GenAI is subject to inaccuracies and errors, which result in concerns related to the responsible use of AI and reliability. In line with this, the research by Lee & See (2004) stated that trust plays a key role in examining how users can rely on AI systems and other factors that impact the use of AI systems by the users. In support of this notion, the research by Siau & Wang (2018) reviewed and examined trust in machine learning, robotics, and AI. The study reflected that trust is characterised as a key factor in determining the acceptance of AI, along with its development and continuous development in AI. In academic contexts, over-trust in AI links to acceptance of inaccurate outputs; however, under-trust might result in preventing users from receiving AI assistance. The research by Gilkson & Wooley (2020) reflected on the trust of human users in the use of technology and stated that initial research on the acceptance and use of new digital tools and technology relied more on user reactions to new technological features. Therefore, critical understanding and implications of trust in GenAI explain the overall academic level adoption and behaviour of users in GenAI in the academic context.

2.2 System-Based Antecedents of Trust

The study by McKnight et al. (2002) examined the integrative model proposed by McKnight, which incorporated trust types including disposition to trust, trusting beliefs, and trusting intentions. This framework was related to the framework of the

Theory of Reasoned Action (TRA), which resulted in beliefs that form attitude, which creates behavioural intentions and forms overall behaviour. The overall trust in TRA in GenAI is related to the reliability and behaviour of the system. In AI-based systems, competence is measured in terms of how the system is able to perform reliably and generate output and results in accurate terms, which are the system-based antecedents of the trust in the AI system by the users.

2.2.1 Perceived Reliability

Perceived reliability of an AI system is defined as the construct that impacts the level of trust of AI users, as it evaluates the consistency and accuracy of the output and results generated by AI systems. Perceived reliability in research, including Lee & See (2004), is related to two key factors: predictability and competence of the system. Therefore, in terms of GenAI research, users are more likely to perceive GenAI as reliable when predictability and competence level are higher, which results in users' willingness to rely on the system. In line with this, the recent study by Nelson et al. (2025) suggested that students are using AI after assessing key risks and benefits associated with the use of GenAI in academic writing. The study reflected that besides factors including perceived reliability and competence of GenAI tools, including ChatGPT, they are under constant pressure to achieve high grades, which makes them attempt to use GenAI in academic tasks and assessments. Therefore, the accuracy of GenAI is the key determinant in shaping the perceived reliability of the users.

2.2.2 Explanation Quality

In addition to perceived reliability, explanation quality is characterised as another key factor in shaping users' perceptions of AI systems. Explanation quality in GenAI refers to the comprehension and clarity in the overall system reasoning to generate output and results. Therefore, a high level of explanation results in increasing overall transparency, the system's reasoning and comprehension for AI users, and reducing ambiguity in data and results. The study by Hoff and Bashir (2015) reviewed factors that influence trust and stated that when AI users are able to comprehend and understand how the AI system works and how results and outputs are generated, they are more likely to refer to AI as credible and trustworthy. In support of this, the recent research by Hadra et al. (2026) reviewed that accuracy and reliability are

significant in terms of the use of GenAI in the academic context, as students are more likely to use results and output generated by GenAI when the output is non-plagiarized and authentic.

2.3 User-Based Antecedents: Confidence in Judgment

Besides system characteristics and trust of the user in GenAI, user-based antecedents are also critical in influencing trust in AI by confidence in the overall judgement regarding the produced output and results. Confidence in judgment by an AI user is the belief of an individual in their ability to evaluate whether the output generated by AI is accurate and reliable or not. In alignment with previous research, including Hoff & Bashir (2015) and Lee & See (2004), trust is defined as the willingness of users to rely on technology and systems, whereas confidence level is related to self-assurance and evaluation capability of the user in terms of AI-generated output. In line with this, Li et al. (2025) stated that there exists a complex relationship between confidence level and trust of the user in AI, as accuracy and verification of AI output and results lead to a higher level of trust; however, a high level of self-confidence results in reduced reliance, critical evaluation, and overall reliability. Therefore, examining the relationship between trust in AI and judgment requires user-level understanding in the context of system-level cues that shape overall trust.

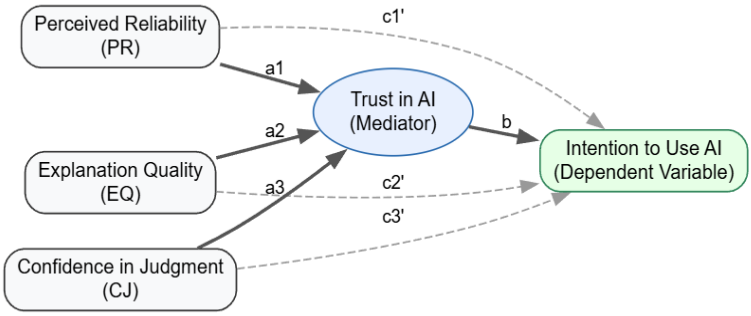

Figure 1: Conceptual Model

3 Hypotheses

As reported by Nazaretsky et al. (2025), trust performs a key role in determining whether users can rely on and adopt AI output and results in academic contexts. Prior research in the context of human and AI interaction stated that trust acts as a mechanism by which system characteristics influence the behavioural intention of AI users.

3.1 Perceived Reliability

Perceived reliability reflects users' belief that the AI system produces accurate and consistent outputs. Reliability signals competence and predictability, which are foundational elements of trust in automation. When users perceive GenAI as reliable, uncertainty is reduced, and trust is strengthened.

H1: Perceived reliability is positively associated with trust in AI.

H2: Trust in AI mediates the relationship between perceived reliability and intention to use AI in academic work.

3.2 Explanation Quality

Because generative AI systems often operate as "black boxes," explanation quality is a key trust-building cue. Clear and understandable explanations enhance transparency and reduce ambiguity, thereby strengthening users' trust in the system.

H3: *Explanation quality is positively associated with trust in AI.*

H4: *Trust in AI mediates the relationship between explanation quality and intention to use AI in academic work.*

Confidence in judgment refers to users' belief in their ability to evaluate AI outputs. The relationship between confidence and trust is theoretically ambiguous: confident users may either trust AI more (due to perceived evaluative ability) or less (due to increased scrutiny). Therefore, confidence is examined as an exploratory predictor of trust.

H5 (Exploratory): Confidence in judgment is associated with trust in AI. (See Appendix: 1).

4 Methodology

4.1 Research Design

This research study utilised a scenario-based online survey design for examining how users evaluate GenAI in academic contexts. The study did not measure the general attitude of AI users towards AI, but it arranged a standardised AI-based interaction for the participants before responding to survey questions. The arrangement of AI-based interaction was carried out to ensure that all research participants evaluated the AI output under controlled conditions, which led to strengthening internal validity. Under controlled conditions, research participants were provided with academic prompts, including a writing task and research-related claims accompanied by a GenAI response, which comprised a short and brief explanation of the integration and reasoning of GenAI in the academic context. Research participants were asked to read the provided prompts and AI response in detail by supposing that they had used the system for real and practical academic purposes. After this, the research participants utilised their own judgement in terms of evaluating AI reliability, explanation quality, and trust in AI regarding academic tasks. This post-task evaluation approach follows established practices in human–AI interaction research, where perceptions and trust are measured immediately after system exposure.

Although the study adopts a cross-sectional survey design, the inclusion of a standardized scenario introduces a quasi-experimental element by ensuring that all participants evaluate the same AI-generated output under controlled conditions. This approach allows for consistency in stimulus exposure while capturing participants' perceptions in a realistic decision-making context. A fully manipulated experimental design was not implemented in this study in order to prioritise ecological validity and to capture naturally occurring variations in user perceptions. Future research could extend this approach by experimentally manipulating system characteristics such as reliability levels or explanation styles to establish stronger causal inference.

4.2 Participants and Sampling

The total number of research participants recruited under this study was 262 (*N=262*), and they were recruited by use of convenience sampling from social media channels and university study groups for understanding their trust in GenAI. The total sample of 262 participants comprised young adults and university students, and they were from different fields, including social science, engineering, and business, and mainly from three countries: Sweden, the United Kingdom (UK), and Finland. Research participants were kept well-informed regarding their voluntary participation and their anonymity throughout the research process and data analysis. Before the collection of data, informed consent was obtained from all participants in electronic form. In the survey questionnaire, the basic demographic information collected from participants included their gender, field of study/education, and their current highest degree level. A seven-point scale was arranged in the survey questionnaire for respondents to indicate their previous experience of use of AI technologies, including 1=very low experience and 7=very high experience. The research constructs and variables were added in a subsequent section of the research survey in this study.

4.3 Measures

In the research study, all the specified constructs were measured by use of a seven-point Likert scale in which a low score indicated limited endorsement and relevance of the construct, whereas a high score indicated stronger endorsement of the constructs in the research study. All measurement items were adapted from prior validated scales in the trust in technology and AI literature (e.g., McKnight et al., 2002; Hoff & Bashir, 2015; Lee & See, 2004), with minor wording adjustments to fit the GenAI academic context.

4.3.1 Trust in AI

Trust in AI was measured in the research study by use of a multi-item scale which was adopted under the GenAI academic context (For example. *"I believe that GenAI provides reliable output, and I completely trusted using the AI-provided recommendations"*. The level of Cronbach's $\alpha = .70$, indicating acceptable internal consistency.

4.3.2 Intention to Use AI

Intention to use AI was measured in this research study as a multi-item scale, which was adopted from existing research in terms of technology acceptance and adoption (For example, *"I would use GenAI in real academic tasks"*. The level of Cronbach's α was .70 which indicates that internal consistency in terms of the intention of using AI was acceptable.

4.3.3 Perceived Reliability

Perceived reliability was measured in this research study by the use of a single-item indicator, which utilised the perception of participants in terms of accuracy and consistency of AI (For example, "*I perceived that AI was reliable in providing results*") The rationale for utilising perceived reliability as single-item measures which were allocated in controlled conditions when construct variable was observed in the specified task.

4.3.4 Explanation Quality

Explanation quality in this research for examining trust in the use of GenAI was examined by the use of a single item by evaluation of the AI reasoning in terms of its usability and clarity (For example, *"GenAI was able to explain the reasoning clearly")*. This construct of explanation quality was able to examine the evaluation of participants in terms of the provided explanation in the standardised response of GenAI.

4.3.5 Confidence in Judgment

Confidence in judgment was measured using a single-item self-assessment (For example, "*I am confident in my ability to evaluate whether the AI's answer is correct*"), which was rated from 1 (not at all confident) to 7 (very confident).

4.4 Data Analysis

In this research study, data analysis was performed through the use of statistical software IBM SPSS (Statistical Package for Social Sciences) Version 28. The data collected from 262 research respondents were included in the dataset and analysed

using SPSS. Initially, descriptive statistics and correlation tests were done for the evaluation of initial relationships between the research variables. Secondly, multiple regression analysis was performed hierarchically for the assessment of intention to use GenAI by predictors and constructs. In the provided Model 1, demographic controls included AI experience of users, gender, current degree level, and field of study. In Model 2, the explanatory power of research variables was examined in terms of constructs including perceived reliability, explanation quality, and trust in AI. For examining the mediating role of trust in the developed hypothesis, the researcher carried out mediation analysis using Hayes's PROCESS macro (Model 4) by inclusion of 95% confidence intervals and 5,000 bootstrap samples. The construct of trust in AI was characterised as a mediator between the perceived reliability of GenAI and intention to use, and explanation quality and intention to use AI by the user. In the research hypothesis testing, the mediation model included control variables, which were statistically analysed by use of a two-tailed test, and the significance level was evaluated at an estimation level of 0.05. Research participants' evaluation of AI output in controlled conditions was used in single-item indicators for assessing perceived reliability and explanation quality. The psychological constructs, including trust and user intention to use GenAI, were measured in this research by the development of multi-item scales. For minimising the risk of bias, participants examined the provided AI scenario before responding to the survey questions. The use of Harman's Single-factor test in the analysis indicated that the majority of the variance (<40%) did not result from a single factor, which resulted in the common method variance being limited in terms of explaining the relationship between all the included variables in the research study.

5 Empirical Findings

5.1 Descriptive Statistics

The first section of empirical findings discusses descriptive statistics, which were analysed under this research study. Table 2 in the provided appendix section summarises that research respondents' responses with a low to moderate level of trust in AI, including a mean (M) = 2.55 and a standard deviation (SD) value of 1.49. The level of perceived reliability resulted in M=2.70 and SD=1.72. Similarly, the observed intention to use AI by users in academic work was low, with M=2.43 and SD=1.79. For the following construct of explanation quality, the results were

M=2.43 and SD=1.90, which suggested that explanation quality is a critical factor and GenAI can be used in academic tasks after caution and diligence (See Appendix 1).

5.2 Bivariate Correlations

Under bivariate correlation analysis, Pearson Correlation was performed, which indicated that strong user intention to use AI in academic contexts was one of the key factors in the overall positive perception of users associated with the use of GenAI. Intention to use GenAI was positively related to trust in AI ($r=.40$, $p<0.001$), explanation quality ($r=.64$, $p<0.001$), and perceived reliability ($r=.60$, $p<0.001$). Results suggested that the reliability of GenAI and explanation quality are related to forming trust. These results are consistent with past research on AI experience with calibration perspective having intention to use AI ($r=-45$, $p<0.001$), which indicated that the higher the level of user experience in GenAI, the more critical evaluation of the AI system is carried out (See Appendix: 3).

5.3 Multiple Regression

A hierarchical multiple regression analysis was performed to assess the predictive capacity of demographic and AI perception variables on the utilisation of AI in the academic context. Model 1, comprising solely demographic variables (gender, field, degree, experience with AI), accounted for 28% of the variance in the intention to use AI, $R^2 = .280$, $F(4, 256) = 24.91$, $p < .001$. Key predictors were field of study ($\beta = .154$, $p = .004$), degree level ($\beta = .240$, $p < .001$), and familiarity with AI ($\beta = -.379$, $p < .001$) (See Appendix: 4). Model 2, incorporating variables such as trust, reliability, and explanation quality, accounted for an additional 28.5% of the variance, resulting in a total explained variance of 56.5%, $R^2 = .565$, $F(8, 252) = 40.96$, $p < .001$. Among the new predictors, reliability ($\beta = .238$, $p < .001$), explanation quality ($\beta = .233$, $p < .001$),. This suggests that perceptions of perceived reliability and explanation quality are central determiners of academic use of AI tools, surpassing the influence of demographic factors (See Appendix: 5)

Although trust in AI was included in a final model, the impact was not significant ($\beta = .062$, $p = .218$), which shows that it does not uniquely predict the intention to use AI. The ANOVA confirmed that both models were statistically significant: Model

1, $F(4, 256) = 24.91$, $p < .001$; Model 2, $F(8, 252) = 40.96$, $p < .001$. This indicates that the regression models reliably predicted intention to use AI. Trust in AI did not remain a statistically significant independent predictor of intention in the regression model; this result likely reflects shared variance between trust, perceived reliability, and explanation quality. Variance inflation factors (VIFs) ranged between 1.4 and 2.3, indicating no critical multicollinearity concerns. Mediation analysis isolates indirect pathways rather than unique direct contributions, confirming that trust functions as a transmission mechanism linking system attributes to the behavioural intention of the AI user.

5.4 Mediation analyses (PROCESS Model 4, 5,000 bootstraps)

For examining how trust in AI mediates the relationship between confidence judgment and intention of using AI, we used 5,000 bootstrap samples and a 95% confidence interval to run Hayes' PROCESS macro (Model 4; Hayes, 2022). The independent variable (X) was confidence judgment; the mediator (M) was trust in AI; and the dependent variable (Y) was intention to use AI. The path from confidence judgment to trust (a-path) was not significant, nor was the direct effect on intention. The indirect effect (a×b) was also non-significant as the CI includes zero, indicating no mediation. Thus, did not influence intention via trust. Therefore, Hypothesis 5 is not supported (See Appendix: 6).

Trust positively mediates the association between perceived reliability and intention. The direct effect remains strong and significant, indicating partial mediation. Hypothesis 2 is supported (see Appendix 7). Significant partial mediation: Clearer reasoning builds Trust and directly boosts Intention; both routes matter. Perceived explanation clarity significantly increases trust, which in turn predicts intention. The indirect effect is large, while the direct effect remains robust, indicating partial mediation, which results in hypothesis 3 being supported (See Appendix: 8).

6 Discussion

The findings of this study should be interpreted as associative rather than strictly causal due to the cross-sectional nature of the research design. The results obtained from data analysis indicated that perceived reliability and explanation quality are the key variables in determining user trust in GenAI in academic contexts. The findings

obtained under this research study are consistent with the past research (Lee & See, 2004; McKnight et al., 2002), which also related that when accurate and consistent outputs and results are generated by GenAI, users are more likely to develop trust level and use AI in academic contexts. Therefore, reliability and transparency of the AI system are characterised as key antecedents of trust in the use of GenAI. Similarly, the findings also resulted in a mean level of trust and intention to use AI, which suggested that users are more cautious in using GenAI in academic settings due to concerns related to the reliability of results and explanation quality. Considering educational and academic contexts, the responsible use and adoption of AI is related to the transparency and ability of the AI system to generate results.

The findings obtained under this research study indicate that trust is a key element in the academic context with relation to the use of GenAI, and AI users are responsible for carrying out final decisions after generating results and output from GenAI. In an academic context, explanation quality reduces the level of uncertainty and increases the level of transparency for the AI users. These findings align with the research by Siau & Wang (2018), which also stated that reliability and explanation quality result in forming behavioural intention of the AI users. Similarly, under the data analysis, when explanation quality and reliability were tested simultaneously in the regression model, trust resulted as a transmission mechanism, which is accompanied by adoption intention to use GenAI output and results. The mediation analysis performed in this research study resulted in the finding that trust is positively associated with both explanation quality and perceived reliability for the user's intention to use AI in academic contexts. Therefore, the findings obtained from this research support that trust is characterised as a mechanism that forms a relationship between the AI system and the intended user behaviour.

The findings from this research indicate that, in academic contexts, reliability and explanation quality are likely to enhance AI users' intention to use AI systems and to build trust. These findings align with the research of Li & Aral (2025), which examined how the level of trust affects overall engagement among AI users. Users are more likely to use GenAI output and results if they find it reliable and usable academically. Similarly, for the other construct, including confidence in judgement, it has been found that confidence in judgement by the AI user does not directly result in the prediction of intention to use AI or overall trust level, which suggests the ability of AI users in their own evaluation does not always result in generating a

higher level of trust in provided AI systems and tools. This theoretical and practical distinction is critical as trust relates to the willingness of users to rely on an external system, whereas confidence is related to the user and his/her own ability to use AI. These findings supported the past research carried out by Agudo et al. (2024) that even when users who are fully aware and confident in their own analysis and judgment are not always likely to rely on AI systems and their results. Therefore, the self-efficacy of AI users is different from system trust, and user-level confidence is increased, and trust is formed in academic contexts when system-level cues, including explanation quality and perceived reliability, are higher.

The findings of this research study also emphasise the significance of calibrated trust of AI users and how the trust level is increased and promoted by enhancing perceived reliability and explanation quality. The results obtained reflected that developers and analysts are well aware of the fact that improvement in perceived reliability and explanation quality does not always result in increased reliance and use of the AI system due to self-efficacy and self-confidence of the users. This is in alignment with the research carried out by Ding et al. (2023) that AI users may develop over-reliance on AI systems in the absence of self-efficacy. Therefore, AI systems are likely to be trusted and utilised when the trust level is positively impacted by perceived reliability and explanation quality.

7 Summary, Implications, and Outlook

The findings demonstrated that perceived reliability and explanation quality are central determinants of trust. When AI systems provide consistent outputs and clear reasoning, users are more likely to develop trust and express an intention to integrate these tools into academic tasks. In contrast, confidence in judgment did not significantly influence trust in AI. This finding highlights an important theoretical distinction: trust in AI and confidence in one's own evaluative ability are related but conceptually distinct constructs. Trust reflects a willingness to rely on the system, whereas confidence reflects belief in one's own competence. The absence of a significant relationship suggests that enhancing system transparency and reliability may be more influential for adoption than strengthening users' self-confidence alone.

With the significant rise in use of GenAI in the academic context, the ability of users to produce results and output from AI systems has significantly increased; however, it has been characterised as a key challenge for users to evaluate its quality and perceived reliability. This study aimed to examine the system-level constructs, including perceived reliability of GenAI and explanation quality, which increases trust in AI and also increases the intention of the user to utilise GenAI in academic work. The findings obtained under this research study revealed that perceived reliability and explanation quality of the AI system determine the level of trust of users in GenAI consistency in AI-based output results in the development of trust and users' intention to use GenAI. However, confidence of the user in their own judgement and evaluation does not impact the level of trust in AI. Therefore, for increasing the reliance on GenAI output and its usability, it is important to enhance the overall system transparency and reliability to generate accurate results.

7.1 Theoretical Contributions

In terms of theoretical contribution, this research study contributed to the GenAI context in three key ways. Firstly, it provided a clear difference between the perceived reliability (competence) and explanation quality (transparency) system cues and how they form trust level in an academic context in performing tasks and assessments. Secondly, the findings of this research study also emphasised comprehension and implications of trust as a mechanism for AI users and how it mediates the influence of perceived reliability and explanation quality. Thirdly, this research study was able to draw a conclusion that trust in AI is conceptually distinct from confidence in personal evaluation and judgment, and self-confidence is not directly associated with the usability and trust prediction in the academic context.

7.2 Limitations

Irrespective of its critical insights and findings on trust in the use of GenAI in the academic context, this research study has certain limitations. The sample size of 262 participants was obtained through convenience sampling and was restricted to students and academic users, which limited the generalizability of this research to other AI users. Secondly, the research study made use of self-reported measures of AI, including trust level, perceived reliability, and intention of the user to use AI. The self-reported measures are prone to common method variance and social bias,

which could have impacted the results of this research. Thirdly, the controlled standardised protocol and design might have impacted the long-term development of trust in the research study for the AI users. An additional limitation relates to potential omitted variables such as personality traits (e.g., openness, conscientiousness) or individual differences in risk perception, which may influence both trust in AI and intention to use. Similarly, cultural factors such as uncertainty avoidance or technological familiarity could shape user perceptions. Future studies should incorporate these variables to better isolate the effects observed in this research. Although the sample includes participants from multiple countries, it remains relatively homogeneous and limited in size. Future research should employ larger and more culturally diverse samples to enhance generalizability and capture global patterns of AI trust and adoption.

7.3 Future Research Directions

In terms of future research on trust in the use of GenAI in the academic context, researchers and practitioners are required to perform an in-depth examination of how AI forms trust in terms of users' cultural attitudes, educational systems, and institutional norms. Researchers can also perform comparative analysis-based studies to understand c differences in terms of trust in AI and trust calibration. Similarly, longitudinal research can be carried out to examine how reliance, evaluation, judgment, and level of trust are evolved over time and how AI users acquire experience and insights in GenAI systems. Future research studies are also required to investigate the allocation of AI functions and users' decisions in responsible terms in an AI-assisted environment. Thus, understanding of AI-based recommendations and judgment will help in the development of tools and models in optimising the level of trust calibration and reducing over- and under-reliance on GenAI in the academic context. Future research could also compare different generations of AI models (e.g., earlier vs. more advanced versions) to examine how improvements in model capability influence user trust and reliance.

Proofreading statement:

The authors confirm that this manuscript has been carefully proofread and revised for language quality, grammar, and overall clarity prior to submission.

References

Bećirović, S., Polz, E., & Tinkel, I. (2025). Exploring students' AI literacy and its effects on their AI output quality, self-efficacy, and academic performance. *Smart Learning Environments*, *12*(1), 29. https://link.springer.com/article/10.1186/s40561-025-00384-3

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., et al. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. International Journal of Information Management, 71, 102642

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Hadra, M., Cambridge, K., & Mesbah, M. (2026). Evaluating the accuracy and reliability of AI content detectors in academic contexts. *International Journal for Educational Integrity*, *22*(1), 4.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Li, H., & Aral, S. (2025). Human trust in AI search: A large-scale experiment. *arXiv preprint*. https://arxiv.org/abs/2504.06435

Li, J., Yang, Y., Liao, Q. V., Zhang, J., & Lee, Y. C. (2025, April). As confidence aligns: understanding the effect of AI confidence on human self-confidence in human-AI decision making. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). https://doi.org/10.1145/3706598.3713336

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. The International Journal of Management Education, 21(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research, 13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Miranda, F. J., & Chamorro-Mera, A. (2025). Exploring the adoption of generative artificial intelligence tools among university teachers. *Higher Education Research & Development*, 1–17. https://doi.org/10.1080/07294360.2025.2559648

Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2025). The critical role of trust in adopting AI-powered educational technology for learning: An instrument for measuring student perceptions. *Computers and Education: Artificial Intelligence*, *8*, 100368. https://doi.org/10.1016/j.caeai.2025.100368

Nelson, A. S., Santamaría, P. V., Javens, J. S., & Ricaurte, M. (2025). Students' perceptions of generative artificial intelligence (GenAI) use in academic writing in English as a foreign language. *Education Sciences*, *15*(5), 611.

Rittenberg, B. S., Holland, C. W., Barnhart, G. E., Gaudreau, S. M., & Neyedli, H. F. (2024). Trust with increasing and decreasing reliability. *Human Factors, 66*(12), 2569–2589. https://doi.org/10.1177/00187208231200471

Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *ACM SIGMIS Database, 49*(4), 99–104. https://doi.org/10.1145/3242734.3242753

Ursavaş, Ö. F., Yalçın, Y., İslamoğlu, H., Bakır-Yalçın, E., & Cukurova, M. (2025). Rethinking the importance of social norms in generative AI adoption: Investigating the acceptance and use of generative AI among higher education students. *International Journal of Educational Technology in Higher Education, 22*(1), 38. https://doi.org/10.1186/s41239-025-00467-3

Vereschak, O., Bailly, G., & Caramiaux, B. (2021). How to evaluate trust in AI-assisted decision-making? A survey of empirical methodologies. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–39. https://doi.org/10.1145/3476068

Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences*, *15*(3), 343. https://doi.org/10.3390/educsci15030343

Zhou, T., & Lu, H. (2024). The effect of trust on user adoption of AI-generated content. *Electron. Libr.*, 43, 61-76. https://doi.org/10.1108/el-08-2024-0244

Appendixes

Appendix: 1

Table 1. Hypothesis development.

Hypothesis	Statement
H1	Perceived reliability positively is associated with trust in AI.
H2	Trust in AI mediates the relationship between perceived reliability and intention to use AI in academic work.
H3	Explanation quality positively is associated with trust in AI.
H4	Trust in AI mediates the relationship between explanation quality and intention to use AI in academic work.
H5 (Exploratory)	Confidence in judgment is associated with trust in AI.

Appendix: 2

Table 2: Descriptive Statistics (N = 262)

Variable	M	SD
Trust in AI	2.55	1.49
Perceived Reliability	2.70	1.72
Explanation Quality	2.43	1.90
Intention to Use AI	2.43	1.79

Appendix: 3

Table 3: Correlations Among Study Variables

Variable	1	2	3	4	5	6	7
1. Gender	—						
2. Field	.078	—					
3. Experience with AI	.140*	−.083	—				
4. Trust in AI	.132*	.054	−.209**	—			
5. Perceived Reliability	−.046	.074	−.369**	.510**	—		
6. Explanation Quality	−.028	.201**	−.442**	.364**	.511**	—	
7. Intention to Use AI	−.063	.201**	−.450**	.404**	.596**	.637**	—

$p < .05$, ** $p < .01$

Appendix: 4

Table 4: Hierarchical Regression Summary

Predictor	Model 1 β (p)	Model 2 β (p)
Gender	.022 (.684)	—
Field of Study	.154 (.004)	—
Experience with AI	–.379 (<.001)	—
Degree Level	.240 (<.001)	—
Perceived Reliability	—	.238 (<.001)
Explanation Quality	—	.233 (<.001)
Trust in AI	—	ns†

Appendix: 5

Table 5: Anova Regression

Model Statistics	Model 1	Model 2
R^2	.280	.565
Adjusted R^2	.269	.551
Standard Error	1.531	1.200
F (df1, df2)	24.91 (4, 256)	40.96 (8, 252)
p-value	< .001	< .001

Appendix: 6

Table 6: Mediation analysis

Path	B (Unstd.)	SE	T	p	95% CI [LLCI, ULCI]
Path a: Confidence → Trust	–0.097	0.061	–1.59	.113	[–0.217, 0.023]
Path b: Trust → Intention	0.628	0.040	15.68	.000	[0.549, 0.707]
Direct Effect (c'): Confidence → Intention	–0.061	0.040	–1.55	.123	[–0.139, 0.017]
Indirect Effect (a×b)	–0.061	0.043	–	–	[–0.149, 0.020]

Appendix: 7

Table 7: Mediation analysis

Path	B	SE	T	P	95% CI
X → M	.4420	.0463	9.55	< .001	[.351, .533]
M → Y	.3479	.0308	11.29	< .001	[.287, .409]
Direct (c') X → Y	.4871	.0267	18.23	< .001	[.435, .540]
Indirect (a×b)	.1538	.0269	—	—	[.105, .209]

Appendix: 8

Table 8: (AI Explanation → Trust → IntentionExplanation →t partial mediation

Path	B	SE	t	P	95% CI
X → M	.2858	.0454	6.30	< .001	[.197, .375]
M → Y	.4274	.0245	17.41	< .001	[.379, .476]
Direct (c') X → Y	.4463	.0193	23.15	< .001	[.408, .484]
Indirect (a×b)	.1222	.0248	—	—	[.075, .173]

Table 9: Summary of Hypothesis

Hypothesis	Result
H1: Perceived reliability positively is associated with trust in AI.	Supported
H2: Trust mediates the relationship between perceived reliability and intention to use AI.	Supported (partial mediation)
H3: Explanation quality positively is associated with trust in AI.	Supported
H4: Trust mediates the relationship between explanation quality and intention to use AI.	Supported (partial mediation)
H5 (Exploratory): Confidence in judgment is associated with trust in AI.	Not supported

DOI
https://doi.org/
10.18690/um.fov.4.2026.26

ISBN
978-961-299-152-4

When Security Meets Intelligence: Trust in AI, Blockchain Security, and Investor Behavior

Fatima Bilal,[1] Anssi Öörni,[1]
Raghava Rao Mukkamala,[2] Bad Re Alam[1]

[1] Åbo Akademi University, Turku, Finland
fatima.bilal@abo.fi, anssi.oorni@abo.fi, badre.alam@abo.fi
[2] Copenhagen Business School, Copenhagen, Denmark
rrm.digi@cbs.dk

Despite the increasing integration of Artificial Intelligence (AI) into cryptocurrency trading environments, concerns remain regarding security, trust, and investor decision-making. Blockchain infrastructures are often described as trust-minimizing systems; however, AI-based trading tools increase reliance on opaque algorithmic outputs. This research study examines how blockchain security, AI accuracy, and AI experience shape trust in AI and behavioural intention in terms of use of an AI-based trading system. For achieving research outcomes, this research study recruited 215 cryptocurrency investors using a structured survey questionnaire and a model for testing trust, trading intentions, and system perceptions. The findings from this research study suggest that blockchain security and perceived accuracy enhance trust in AI. In contrast, AI experience negatively impacts trust in AI, as experienced users exhibit algorithmic scepticism. The findings further related that in the prediction of behavioural intention of users, perceived accuracy of AI and structural security have a direct impact on intention, and trust is not a key determining predictor in the research model. Therefore, structural assurance performs a key role in shaping investors' intentions in a financial market.

Keywords:
trust in AI,
blockchain security,
cryptocurrency trading,
behavioral intention,
algorithmic skepticism

University of Maribor Press

1 Introduction

The integration of deep learning and machine learning by use of Artificial Intelligence (AI)and blockchain technologies has transformed the financial markets, as traditional financial forecasting techniques and mechanisms have struggled to predict price fluctuations caused by changes in external sentiment and non-linear dynamics. Therefore, AI and blockchain technology have emerged as key analytical tools for financial analysis and investors (Alnami et al., 2025). AI-enabled trading algorithms facilitate investors by identifying patterns, analysing a large volume of data, and predicting changes due to a highly volatile environment. Similarly, blockchain technology enables secure digital infrastructure and transactions by providing algorithm-based decision support and cryptographic verification mechanisms (Ying et al., 2017). AI and blockchain together create a hybrid technological environment for financial users and investors (Lu, 2019).

Irrespective of technological advances and development in terms of blockchain and AI, the critical aspect of the human dimension of trust remains central for AI users. Blockchain is referred to as a system that minimizes the requirement of institutional and interpersonal trust levels due to a cryptographic verification mechanism (Kumar et al., 2023). However, an AI system requires users and investors to rely on such models and mechanisms that can not be easily interpreted due to complex language and mechanisms. While blockchain systems reduce the reliance of financial and AI users on social trust, it also results in algorithmic dependency and uncertainty for the users. Therefore, in complex trading environments that comprise cryptocurrency, there exist financial risks and high volatility for financial analysts and investors.

In financial decision environments and systems that are characterised by a high level of uncertainty and volatility, trust functions as a key mechanism that reduces underlying risk and enables decisions by AI users and analysts in an ambiguous environment. In digital financial ecosystems, measures including encryption, technological reliability, and safeguards ensure users' confidence in system output and reliability. Previous research, including Roger C. Mayer et al. (1995), provides a foundational model of trust formation, highlighting ability, benevolence, and integrity as key determinants of trust in uncertain environments. Although not specific to cryptocurrency, this framework is widely applied in digital and

technological contexts to explain how users evaluate trust under risk. However, there exists a research gap in terms of addressing how structural measures in blockchain infrastructure form investors' intention in terms of AI performance and user experience. and the role of trust in an AI-enabled blockchain environment.

There also exists a research gap in terms of determining trust as the primary driver of adoption and how performance and structural measures directly impact the behavioural intention. Therefore, this research aims to address this research gap by examining how perceived AI accuracy, AI experience, and blockchain security impact trust in AI and behavioural intention of users for user AI-based cryptocurrency trading systems. For this purpose, To address this gap, the study uses survey data from 215 active cryptocurrency investors to examine how perceived AI accuracy, blockchain security, and AI experience influence trust and behavioral intention. This research tested a model by differentiating between blockchain security (structural assurance), perceived accuracy (cognitive accuracy), and AI experience (user characteristics).

The study is guided by the following hypotheses:

H_1: *Perceived accuracy of AI-based trading systems positively influences trust in AI.*

H_2*: AI experience negatively influences trust in AI.*

H_3: *Perceived blockchain security positively influences trust in AI.*

H_4: *Trust in AI positively influences behavioral intention to use AI-based trading systems.*

H_5: *Perceived accuracy positively influences behavioral intention.*

H_6: *AI experience negatively influences behavioral intention.*

H_7: *Perceived blockchain security positively influences behavioral intention.*

This research study contributes to the existing knowledge and research in Information Systems (AI) and blockchain in three different ways. Initially, it creates a difference between algorithmic performance and structural security as key drivers

of user intention. Secondly, it analyses that the AI experience of a user may result in reducing trust level, which suggests that experienced users of AI are more prone to scepticism related to AI systems and mechanisms. Thirdly, this research study also relates that thought trust is critical in forming the perception of users, but it can not completely determine the adoption of AI by the users. Therefore, the findings of this research study provide key insights for AI developers, financial regulators, and analysts in AI-enabled trading systems in the financial ecosystem and environment, which are decentralised.

2 Theoretical foundation:

2.1 Trust Formation in Hybrid AI–Blockchain Environments

The integration of AI and blockchain creates a hybrid socio-technical environment in which AI provides predictive decision support, while blockchain ensures secure, decentralized, and verifiable transactions. The study by Lu et al. (2019) stated that in cryptocurrency-related trading, AI-based systems are capable of generating predictive recommendations, whereas blockchain infrastructure provides immutability, consensus mechanisms, and cryptographic verifications. In line with this, Kumar et al (2023) stated that blockchain is referred to as a system that reduces trust; therefore, AI users are recommended to use algorithmic outputs and results. Trust is characterised as a central determining factor in adopting technology and the digital environment. This is aligned with the research of Mayer et al. (1995), which stated that trust facilitates the overall decision-making process and reduces perceived uncertainty of AI systems. However, the research emphasised that trust in AI-enabled blockchain systems is different from the traditional level of systems, which comprises interpersonal trust and includes confidence level in algorithmic performance as cognitive trust and confidence in system infrastructure as structural assurance. In AI-enabled cryptocurrency trading systems, trust formation results from belief-based examination of system performance and security measures instead of the underlying behaviour intention model. Similarly, in high-risk systems, users' intention and evaluation of the AI system are based on the characteristics, including infrastructure security and predictive accuracy. This has also been supported by the study of McKnight et al. (2002), which stated that trust formation is influenced by structural assurance mechanisms and users' beliefs about system competence and reliability (McKnight et al., 2002), which relates to the way in which security

measures and users' performance beliefs form trust and behavioural intention of users in decentralised financial systems.

2.2 Perceived Accuracy and Trust in AI

Perceived accuracy is a perception among users that there is trust in the reliability, precision, and consistency of predictions made by AI-based trading. Performance indicators especially become salient in risky financial markets like cryptocurrency markets. When results are seen to reflect market performance, the user is more likely to develop trust in AI systems. Studies on algorithmic decision-making indicate that perceived system competence is one of the key antecedents of trust development (Siau & Wang, 2018). The predictability of performance and reliability in the output of predictive analytics are cognitive indicators that create less uncertainty and build dependency (Dietvorst et al., 2015). Perceived system performance, according to a belief-based approach, plays a role in positive evaluative judgment, which subsequently defines the formation of trust. As long as investors believe that AI-generated suggestions are correct, they will consider using the system and trust it with its predictive power.

H1: Perceived accuracy of AI-based trading systems positively influences trust in AI.

2.3 AI Experience and Trust Formation

AI experience captures the extent to which users are familiar with AI systems and predictive tools. Traditional adoption models often assume that familiarity reduces uncertainty and increases positive evaluations. However, in algorithmic decision environments, experience may also increase critical awareness of system limitations. Experienced users may better understand model imperfections, biases, or predictive errors, potentially leading to more cautious evaluations. In financial markets characterized by volatility and high stakes, such skepticism may influence trust formation differently than in general consumer technology contexts. Studies on algorithm aversion indicate that exposure to algorithmic errors can reduce reliance on automated systems, particularly among knowledgeable users (Logg et al., 2019). Thus, experience may produce either familiarity-based trust or heightened critical scrutiny.

H2: AI experience negatively influences trust in AI.

2.4 Blockchain Security as Structural Assurance

Blockchain security reflects users' perceptions of integrity, transparency, and immutability of the underlying infrastructure (McKnight et al., 2002). In contrast to AI accuracy, which signals predictive performance, blockchain security signals structural assurance. Structural assurance refers to the perception that technical and institutional safeguards are in place to ensure safe transactions (McKnight et al., 2002). In cryptocurrency trading ecosystems, blockchain security reduces concerns about data manipulation, transaction fraud, and system tampering. Such assurances may strengthen confidence in AI-driven trading environments by reinforcing perceptions of system integrity (Tylcanaz-Prieto et al., 2023).

Therefore, blockchain security is expected to positively shape trust in AI-enabled trading systems.

H3: Perceived blockchain security positively influences trust in AI.

2.5 Trust and Behavioral Intention

Trust in AI is hypothesised to impact the behavioural intention of the users. When users trust AI-based trading systems, they are more likely to express an intention to rely on AI-generated recommendations.

H4: Trust in AI positively influences behavioral intention to use AI-based trading systems.

2.6 Direct Effects of System Perceptions on Intention

Beyond trust-mediated effects, system perceptions may directly shape behavioral intention. In high-risk financial contexts, investors respond directly to performance and security measures without completely relying on general trust evaluation.

Perceived accuracy may directly encourage intention by signaling expected utility and predictive advantage. Similarly, blockchain security may directly influence intention by reducing perceived transactional risk and reinforcing structural reliability.

H5: Perceived accuracy positively influences behavioral intention.

H6: AI experience influences behavioral intention.

H7: Perceived blockchain security positively influences behavioral intention.

3 Methodlogy

3.1 Research Model Development

These three streams together imply a conceptual framework for trust formation (Figure 1). Soomro et al. (2024) find that trust has a strong positive effect on cryptocurrency adoption intention. Thus, in our model, Trust mediates the effects of PA, BS, and AE on Behavioral Intention. Figure 1 (below) diagrams these hypothesised relationships.

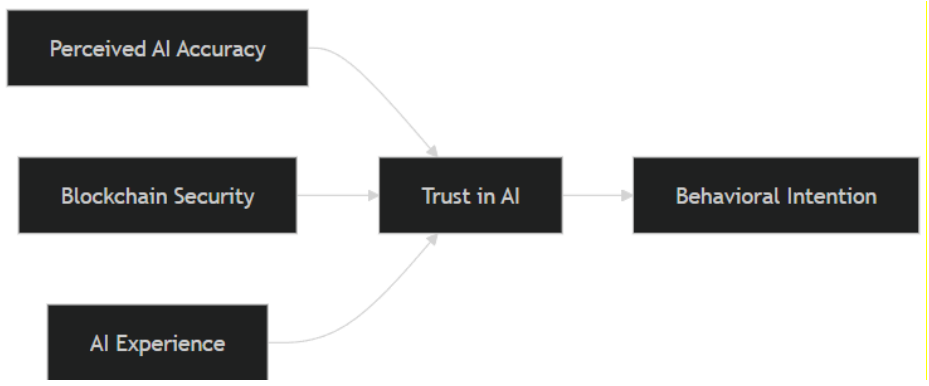

Figure 1: Research framework.

3.2 Research Design and Sample

In this research study, a quantitative survey design was utilised for examining the existing relationship in perceived AI accuracy, Blockchain security, trust in AI, AI experience, and behavioural intention of users for using AI-based cryptocurrency trading systems. For this purpose, data was collected from a sample of 215 active Bitcoin traders and cryptocurrency investors having sufficient knowledge and insights into trading and financial systems. Participants were recruited from online

financial and cryptocurrency communities, including trading discussion groups, social media platforms, such as Telegram and Discord, and blockchain forums. The participation in the research study was voluntary, and all participants were informed regarding their voluntary participation and right to withdraw from the study. Participant Information Sheet (PIS) was distributed among the research participants, which included research purpose, confidentiality measures, and the right to withdraw from the research at anytime. PIS also included that no personal or identifiable information was included in the research process. The final research sample included research participants having different levels of trading frequency, gender, age, AI familiarity, and knowledge of the cryptocurrency market.

3.3 Measurement Instruments

All constructs were measured using multi-item scales adapted from established technology adoption and trust literature, contextualised to AI-based cryptocurrency trading. Responses were recorded using a **7-point Likert scale (1 = strongly disagree, 7 = strongly agree)**. Perceived Accuracy (PA), AI Experience (AE), and Blockchain Security (BS) are used as key constructs in the model.

The constructs included:

Perceived Accuracy (PA, 3 items)

Example item: "I believe the AI system provides accurate recommendations."

AI Experience (AE, 3 items)

Example item: "I have sufficient experience using AI systems in decision-making."

Blockchain Security (BS, 5 items)

Example item: "The blockchain ensures that my transactions are secure and tamper-proof."

Trust in AI (TA, 8 items)

Example item: "I can rely on the AI system to perform reliably."

Behavioural Intention (BI, 3 items)

Example item: "I intend to use the AI system frequently in the future."

Demographic variables such as age, gender, trading frequency, and trading history were also collected for descriptive purposes.

3.4 Data Analysis Procedure

In this research study, data analysis was performed in different stages by using R-Studio. A non-probability convenience sampling approach was used, as participants were recruited from online cryptocurrency communities such as Telegram, Discord, and blockchain forums. Initially, descriptive statistics including mean, minimum and maximum value, and standard deviation were calculated by evaluation of the distribution characteristics. Secondly, Cronbach's Alpha was utilised for the evaluation of internal consistency reliability for all the constructs of the research study. For examining the bivariate relationships existing among blockchain security, AI experience, trust in AI, and behavioural intention of users, the research study utilised the Pearson Correlation Coefficient. Afterwards, for a developed hypothesis, multiple regression was performed, and a regression model was developed and used for predicting the level of trust in AI from the perceived accuracy of the AI system, blockchain security, and AI experience. For examining statistical significance in data, standardised coefficients including p-values and t-values were used. Mediation analysis conducted under the research by use of Bootstrap (5000 samples) examined whether trust in AI mediates the relationship between the behavioural intention and system-level perceptions, including AI experience and perceived accuracy. Indirect impact was considered to be significant if the 95% confidence interval in the bootstrap did not include zero. Harman's single-factor-level test was conducted in the research to examine the common method bias existing in the model and data.

3.5 Analytical Rationale

In terms of analytical rationale, the research approach allowed analysis of both the direct impact of system perception on the behavioural intention of AI users and the indirect impact of operating through trust in AI. This system structure enabled researchers to differentiate the perceived accuracy (cognitive performance) and AI experience (user-related experience) in shaping the overall trust and AI users' trading intentions within AI-enabled blockchain and decentralised financial systems.

4 Empirical Findings

4.1 Descriptive Statistics and Correlations

Under descriptive statistics and correlation analysis, the research participants reported that there exists a high level of blockchain security (M=6.10) and the behavioural intention to use AI-based trading systems and platforms is (M=6.00). However, the research further reported that trust in AI (M=5.30) and perceived accuracy (M=5.80) were also high, which indicated that participants reported favorable perceptions of AI-enabled trading systems. The overall experience of AI was reported as moderate (M=4.60), which indicated that the level of participants' familiarity with the AI system varies.

4.2 Pearson Correlation among Constructs

The second finding of this research relates to the Pearson correlation among the constructs of the research. It has been analysed that perceived accuracy and blockchain security are positively related to the behavioural intention of the user and trust in AI. The overall AI experience of users has a negative relationship with the intention and trust of AI users, which suggests that AI users critically evaluate AI systems when they have a high level of experience. Therefore, there exists a significant level of correlation among the research constructs (Please refer to Appendix: 2).

4.3 Regression Analysis: Predicting Trust in AI

For examining predicting trust in AI and determinants of trust in AI, multiple regression analysis was performed by including AI experience, perceived accuracy, and blockchain security as predicting variables. The result of the analysis suggested that trust in AI is positively related to perceived accuracy, which supported Hypothesis 1 (H_1) of this research study. Secondly, blockchain security also resulted as a key predictor in the model, which supported Hypothesis 3 (H_3). AI experience has a significant negative impact on trust level, which supported Hypothesis 2 (H_2), which suggested that skepticism in AI-based trading systems results from experienced AI users (Please refer to Appendix: 3). Perceived accuracy significantly predicted trust ($\beta = 0.299$, $p < .001$), followed by blockchain security ($\beta = 0.423$, $p < .001$), while AI experience showed a negative effect ($\beta = -0.148$, $p < .001$).

4.4 Regression Analysis: Predicting Behavioral Intention

For predicting behavioural intention, regression analysis was carried out by examining predictors of behavioural intention in using an AI-based trading system. Blockchain security is a key predictor of behavioural intention in a single AI system, which overall supported Hypothesis 7 (H_7) of the research study. The regression analysis also resulted in the finding that perceived accuracy has a significant and positive impact on behavioural intention, which supported Hypothesis 5 (H_5) of this research study and results indicate that AI experience does not have a statistically significant effect on behavioral intention; therefore, H6 is not supported. Therefore, trust in AI does not result in significant prediction of behavioural intention when blockchain security and perceived accuracy are included in the analysis model. Therefore, the findings suggest that Hypothesis 4 (H_4) is not supported by the regression analysis in the model.

4.5 Mediation Analysis

This section of the research study discusses mediation analysis carried out to examine whether trust in AI mediates the impact of perceived accuracy and AI experience on behavioural intention. A bootstrap mediation analysis of 5000 samples was conducted, which resulted in trust in AI only mediating the relationship between the behavioural intention and perceived accuracy to a limited extent. The obtained

results included an indirect effect (0.26), which represented a total of 38% of the overall mediation total effect. In terms of AI experience, the indirect impact of trust on negative but significant results indicates that a high level of user experience results in reducing the trust level, which results in reducing the overall behavioural intention. As given in the appendix section, the total impact of AI experience on intention is not significant, which further suggests that AI experience does not directly impact the adoption behaviour. The findings and analysis from mediation analysis indicated that trust acts as a key determiner in mediating the role in impacting perceived accuracy, but does not determine the overall behavioural intention when risk aversion and security factors are considered.

4.6 Hypothesis Testing Summary

Table 6: Hypothesis Testing Results

Hypothesis	Relationship	Result
H_1	PA → TA	Supported
H_2	AE → TA	Supported (negative effect)
H_3	BS → TA	Supported
H_4	TA → BI	Not Supported
H_5	PA → BI	Supported
H_6	AE → BI	Not Supported
H_7	BS → BI	Supported

5 Discussion and Contributions

5.1 Discussion of Findings

This research study examined how blockchain security, AI accuracy, and AI experience shape trust in AI and behavioural intention for using decentralised financial systems and AI-based cryptocurrency systems. The findings from this study indicate that there are distinct structures for forming user intention and trust in AI-enabled blockchain. The initial findings included that blockchain security and perceived accuracy positively impact trust in AI, resulting in enhancing the overall level of trust in AI by the users. Thus, blockchain security has emerged as one of the strongest predictor for determining trust level and suggests that in decentralised financial markets, structural assurance and system credibility perform a key role in shaping user confidence. This is supported by the existing research, including McKnight et al. (2002), which stated that structural-level security measures and

system mechanisms result in minimising uncertainty in the digital environment. Therefore, perceived accuracy and structural security emerge as key antecedents of trust in AI systems. While transparency is often discussed in prior literature, it was not directly measured in this study and should be examined in future research.

The research findings further related that there is a negative relationship between AI experience and trust of users in AI. In contrast to the conventional assumption, confidence level is not impacted by familiarity. However, the more experienced AI users are, the more sceptical they are in actual terms of use of AI systems. The findings obtained from this research study align with the existing research on aversion, including Dietvorst et al. (2015), which suggests that users may reduce reliance on algorithm systems after identification of errors and inaccuracies. Therefore, in decentralised financial systems, including trading and cryptocurrency, AI users may apply stricter evaluation criteria and standards for system-level automation. Thirdly, in the prediction of behavioural intention of users, perceived accuracy and blockchain security and perceived accuracy are key determinants, whereas trust level in AI is not significant in the applied model. The results obtained suggested that investors are more likely to relate their behavioural intention to the characteristics of the system instead of generalised perceptions related to trust. The mediation analysis carried out under this research also relates that trust partially mediates the relationship between the perceived accuracy and behavioural intention. However, trust does not result in the explanation of adoption behaviour for AI users. Therefore, the findings of the research study suggest that trust is not the only driver for AI adoption in the cryptocurrency financial markets.

5.2 Theoretical Contributions

This research study contributed to research in information systems in three key ways. Firstly, the research study creates a distinction in terms of cognitive and structural drivers of users' intention. The research related to the past research of Mayer et al. (1995) in terms of technology adoption and how trust acts as a key variable in technology. The findings obtained from this research study demonstrated that mechanisms related to structural security directly impact the intention of the user and overall trust evaluation. The findings further suggested that in decentralised financial markets, structural level assurance is directed towards shaping the overall behavioural intention of the users. This study also contributed to reconceptualising

the overall role of trust in the adoption of AI by suggesting that trust functions as a transmission mechanism linking performance perceptions to user intention. However, it can not be considered as the only determining factor in the adoption of a structurally secured environment. Thirdly, the research further strengthened the existing research work, including Dietvorst et al. (2015), in the context that AI-experienced users can not completely trust the recommendations provided by AI, as critical scrutiny is required due to familiarity with AI tools and systems. This also integrates user heterogeneity in AI adoption and suggests that user expertise does not directly contribute to improving overall acceptance of AI systems.

5.3 Practical Implications

In terms of practical implications, the findings of this research study suggest that technical developers and platform designers are required to emphasize the performance quality of AI systems and structural security. The user intention can be enhanced through carrying out a transparent audit, an encryption mechanism, and blockchain security. Similarly, trust of AI users and user adoption intention can be improved through testing of transparency and improving overall perceived AI accuracy. For experienced AI users, increased transparency may help contextualize algorithmic errors and improve understanding of system limitations, rather than directly eliminating errors.

6 Limitations and Future Research

Irrespective of its significant contribution in terms of research in blockchain security, investor behaviour, and trust in AI, this research study is subject to certain limitations. The cross-sectional survey design adopted under this research study did not allow causal inference in the research. Future research can address this limitation and use longitudinal or experimental research designs for examining trust calibration by AI users. The second key limitation in the research study was the online recruitment research sample from online cryptocurrency communities, which might have limited the generalisability of the research. Therefore, future research needs to recruit cross-national samples and institutional investors. The third limitation of the research related to the adopted research model is that it only considered antecedents of trust and intention. Additional variables and constructions, including explanation quality, social influence, perceived risk, and regulatory assurance, need to be taken

into account in future research for improving understanding and implications of AI adoption in decentralised financial markets.

7 Conclusion

The findings and analysis performed under this research study conclude that perceived performance of the AI system and structural security are directly related to influencing the behavioral intention of AI users in a trading environment; however, trust acts as an evaluator and directs the impact of system-level perception on users' intention. Blockchain security has emerged as a key driver of user intention, and AI experience is related to algorithm checking. Thus, the overall findings of this research study enriched the underpinning theory and understanding in terms of the adoption of technology and trust in the hybrid AI-blockchain eco-system.

Proofreading statement:

The authors confirm that this manuscript has been carefully proofread and revised for language quality, grammar, and overall clarity prior to submission.

References

Alnami, H., Mohzary, M., Assiri, B., & Zangoti, H. (2025). An integrated framework for cryptocurrency price forecasting and anomaly detection using machine learning. *Applied Sciences, 15*(4), 1864. https://doi.org/10.3390/app15041864

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126. https://doi.org/10.1037/xge0000033

Kumar, S., Lim, W. M., Sivarajah, U., & Kaur, J. (2023). Artificial intelligence and blockchain integration in business: Trends from a bibliometric-content analysis. Information Systems Frontiers, 25(2), 871–896. https://doi.org/10.1007/s10796-022-10279-0

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Lu, Y. (2019). The blockchain: State-of-the-art and research challenges. Journal of Industrial Information Integration, 15, 80–90. https://doi.org/10.1016/j.jii.2019.04.002

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of Management Review, 20(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. Information Systems Research, 13(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. Journal of the Association for Information Systems, 19(12), 120–143. https://doi.org/10.17705/1jais.00525

Soomro, B. A., Shah, N., & Abdelwahed, N. A. A. (2024). Intention to adopt cryptocurrency: A robust contribution of trust and the theory of planned behavior. Journal of Economic and Administrative Sciences, 40(2), 419–433. https://doi.org/10.1108/JEAS-10-2021-0204

Tylcanaz-Prieto, A. B., Cortez-Ordóñez, A., & Lee, C. W. (2023). Influence of customer perception factors on AI-enabled customer experience in the Ecuadorian banking environment. *Sustainability, 15*(16), 12441. https://doi.org/10.3390/su151612441

Ying, X., Zhang, Y., & Li, Q. (2017). Machine learning techniques for cryptocurrency market prediction. *Journal of Financial Data Science, 1*(1), 65–79.

Appendix: 1

Table 1: Descriptive Statistics of Constructs (N = 215)

Construct	Mean	SD	Min	Max
Perceived Accuracy (PA)	5.80	0.85	2.67	7.00
AI Experience (AE)	4.60	0.97	2.33	7.00
Blockchain Security (BS)	6.10	0.71	3.80	7.00
Trust in AI (TA)	5.30	0.72	3.38	6.88
Behavioral Intention (BI)	6.00	0.55	4.67	7.00

Appendix: 2

Table 2: Correlations Among Constructs

Variable	PA	AE	BS	TA	BI
PA	1				
AE	-0.23	1			
BS	0.80	-0.17	1		
TA	0.73	-0.35	0.73	1	
BI	0.65	-0.20	0.66	0.58	1

*p < .05, **p < .01.

Appendix: 3

Table 3: Regression Results Predicting Trust in AI (R^2 = .63)

Predictor	B	SE	t	p
Perceived Accuracy (PA)	0.299	0.060	5.02	< .001
AI Experience (AE)	-0.148	0.032	-4.67	< .001
Blockchain Security (BS)	0.423	0.070	6.01	< .001

The model explains 63% of the variance in trust in AI (R^2 = .63), indicating strong explanatory power.

Appendix: 4

Table 4: Regression Results Predicting Behavioral Intention (R^2 = .49)

Predictor	B	SE	T	P
Perceived Accuracy (PA)	0.189	0.057	3.28	.001
AI Experience (AE)	-0.027	0.030	-0.89	.373
Blockchain Security (BS)	0.277	0.069	3.99	< .001
Trust in AI (TA)	0.070	0.063	1.11	.268

The model explains 49% of the variance in behavioral intention (R^2 = .49).

Appendix: 5

Table 5: Mediation Results (Bootstrap, 5,000 samples, N = 215)

Path	Effect Type	Estimate	95% CI	p
PA → TA → BI	Indirect	0.26	[0.16, 0.38]	< .001
PA → BI	Direct	0.43	[0.25, 0.60]	< .001
AE → TA → BI	Indirect	-0.12	[-0.20, -0.05]	< .001
AE → BI	Direct	0.11	[0.00, 0.21]	.049

Bridging the Gap: Teaching Data Analytics to Business Students using R

Alinda Kokkinou,[1] Reyhaneh Mazinani,[2]
Hans van Gils[2]

[1] Breda University of Applied Sciences, Academy of Built Environment and Logistics, Breda, the Netherlands
kokkinou.a@buas.nl
[2] Avans University of Applied Sciences, Avans School of Management and Finance, Breda, the Netherlands
reyhaneh.mazinani@ing.com, hap.vangils@avans.nl

Business curricula are increasingly offering data analytics courses. However, training business students in data-analytic thinking and exposing them to new software requires balancing technical skills with domain-specific knowledge in ways that keep students motivated. In-class observations, student evaluations, and student reflections were used to develop and iteratively improve a data analytics course that overcomes these challenges. A survey and retrospective interviews supplemented these to formulate the lessons learned presented in this paper. These include using the software R as a Blackbox application and creating an atmosphere of collaborative learning. The use of relevant datasets and examples also appealed to students' intrinsic motivation for their subject.

DOI
https://doi.org/10.18690/um.fov.4.2026.27

ISBN
978-961-299-152-4

Keywords:
data analytics, situated learning, domain experts, business education, analytic thinking

University of Maribor Press

1 Introduction

Organizations are successfully leveraging the benefits of big data analytics to develop better products and services, reduce costs, mitigate supply chain risks, increase productivity, and gain a competitive advantage (Kokkinou et al., 2023). The resulting focus on data-driven decision-making has increased demand not only for expert data analysts (U.S. News & World Report, 2024), but also for business professionals equipped with skills in data science and analytics (Gharehgozli et al., 2024; King, 2022). In larger organizations, training in such skills is essential because data analysts and business professionals must collaborate to define relevant business problems or questions (Provost & Fawcett, 2013; Vance et al., 2022). Ideally, both parties should possess a solid understanding of each other's domain to enable meaningful dialogue and informed decision-making (Provost & Fawcett, 2013). In Small and Medium Enterprises (SMEs) lacking dedicated data analytics resources, such training is necessary, as business professionals must independently engage in data-driven decision-making without support from data scientists or analysts.

In response to the scarcity of data analytics expertise, organizations are increasingly turning to trainees and recent graduates (Janssen, 2022; Kokkinou, 2023). This trend highlights the need to develop data analytics courses and programs tailored to business management students that familiarize them with data analytic thinking and expose them to software suitable for processing large data sets (Birt et al. 2023; Gharehgozli et al. 2024; King 2022). However, training business students in data analytics presents challenges (Msweli et al., 2023). Business majors often lack the prerequisite knowledge of statistics and computer programming required by many existing data analytics classes (King, 2022; Wang and Gu, 2018).

Despite extensive attention to best practices in developing data analytics courses and curricula for technical students, limited guidance exists for designing courses specifically for business students (Pan and Mazzei, 2023), particularly regarding software integration (Birt et al., 2023). The present study, therefore, describes how this challenge was tackled in the development of two data analytics courses targeted at (a) finance and (b) supply chain majors at two universities of applied sciences. Both courses were iteratively improved based on feedback from students, colleagues, and industry practitioners over five academic years.

In the subsequent sections, a brief theoretical background is provided on data analytics and on other courses requiring programming. Subsequently, the course design is described, including the choice of software, course and project assessment, and the selection of instructional materials. Lessons learned that can help instructors develop similar courses are presented, followed by the conclusion.

2 Theoretical Foundations

Universities are increasingly offering data analytics and data science courses in response to both industry and student demands (Gharehgozli et al., 2024). Such courses are intended to develop students' data analytic thinking and data acumen. Given that data analytics and data science are multidisciplinary, it is not surprising that these courses have been offered by university departments as diverse as business schools (Wang and Gu, 2018) and statistics departments (Hicks and Irizarry, 2018), with little consensus on where such courses belong. This is further complicated by the fact that, as in the academic and professional contexts, the fields of data analytics and data science courses overlap considerably (King, 2022). Areas of overlap include (1) data mining algorithms, (2) data visualization and storytelling, and (3) the integration of computing into the course, in particular R and Python (Hicks and Irizarry, 2018; King, 2022; Msweli et al., 2023).

A significant concern when integrating computing into a course for students without a computer programming background is that learning to program for the first time can be challenging for many (Echeverría et al., 2017; Msweli et al., 2023). Several educational theories and corresponding pedagogical approaches contribute to our understanding of how to keep students motivated through this relatively steep initial learning curve. From a social constructivist perspective, learning is viewed as a primarily social activity, and pedagogical approaches should therefore emphasize students' engagement with others (Barak, 2017). The situated learning perspective further stipulates that learning takes place within authentic activities (Vargas et al. 2024). Hicks and Irizarry (2018) similarly challenge traditional statistics courses for focusing on mathematics rather than on using data to solve real-world problems or answer questions. To bridge this gap, they propose several teaching principles emphasizing the use of examples, such as building the course around case studies, structuring course activities to mimic a data analyst's experience, and demonstrating the importance of critical thinking through relevant examples.

To help students make the connection between the software and the rationale for teaching it, R is frequently taught alongside statistics (Tucker et al., 2022) or data analytics courses (Hicks and Irizarry, 2018; Wang and Gu, 2018). Some instructors prefer to have students type all commands and datasets by hand, and only later move to using scripts. Other instructors prefer that students start their learning with preexisting scripts. Most emphasize the importance of providing extensive support to students to prevent them from becoming frustrated or discouraged (Hicks and Irizarry, 2018; Tucker et al., 2022). Despite the prevalence of best practices in (1) teaching data analytics and (2) teaching statistical software, there is a lack of guidance on teaching data analytics in conjunction with statistical software to non-experts.

3 Course Design

In this section, we describe how the aforementioned educational theories and pedagogical approaches were used to design and iteratively improve two data analytics course(s), (a) for 4th year Bachelor of Science in Finance & Control (BsF&C) and (b) one-year Master of Supply Chain Management (MsSCM) students at two Dutch Universities of Applied Sciences. For each course, we first describe its scope and motivation and how it fits within the curriculum. In the Dutch educational system, the BsF&C is a four-year program comprising 240 European Transfer Credits (ECTs). While this program had been running since 2011, the data analytics course was introduced as part of an extensive curriculum redevelopment in 2020 in response to a quickly changing profession, requiring auditors, accountants, consultants, and other professionals to evolve their skill sets (Deloitte, 2020). Students entering the course are expected to have completed a basic statistics course in their first year, amongst other coursework. The course is worth 8 ECTs (in the European Credit Transfer system, 1 ECT is equivalent to a study load of 27 hours). It spans 10 weeks and is assessed through an individual project. The MsSCM program was developed in response to industry requests for supply chain specialists who can recognize and implement innovations in their field and act at a strategic level. Students entering the course are expected to possess prior knowledge of logistics and supply chain management, statistics, accounting, finance, and operations management. Both courses (one within BsF&C and one within MsSCM) focus on students mastering data analytics procedures and on acquainting them with the broader context of data analytics. We describe the course design for the two courses here. We first clarify the choice to use the software R, as it has implications

for course design. We subsequently discuss the assessment, the topics covered, and the instruction materials. As both courses have undergone multiple Plan-Do-Check-Act iterations at their respective institutions, the main changes and the motivations will be explained.

3.1. Choice of Software R

A deliberate choice was made to use the R software with RStudio. Software such as R and Python offers several advantages over tools like Excel Solver, including greater versatility and the ability to handle complex analyses. Whereas Python is a general-purpose programming language, R is an open-source software environment for statistical computing and graphics ("The R Project for Statistical Computing," 2021). Both R and Python have rich ecosystems of libraries (Python) and packages (R) built around a range of modern statistical methods, making them highly versatile and dynamic (Firth et al., 2021). Although R is perceived as having a steeper learning curve, RStudio, an integrated development environment for R, helps make R more accessible. It offers a console and an editor that support direct code execution, making it much easier to write and edit scripts ("RStudio," 2021). The tools for debugging and workspace management also make it easier for users with little coding experience.

While training students to use software such as Python and R required significant time and effort, it was considered preferable for several reasons. First, doing so would expose students to new software, equipping them with transferable skills (Sankaran et al., 2023). Second, it would introduce students to the basics of programming. Third, many videos and tutorials were available for students interested in improving their knowledge of Python or R beyond the course objectives (Firth et al., 2021). Fourth, because Python and R are open-source, students interested in implementing these tools at their place of employment can do so without needing budget approval. The final decision to use R instead of Python was made based on two additional, context-specific reasons. First, the lecturers had more experience with R and thus were better equipped to help students who encountered bugs. Second, the MsSCM students also needed to learn new software for their research methods course, and R was a better fit than Python because of its explicit orientation toward statistics (Sankaran et al., 2023). The decision to use R instead of Excel Solver or Python was revisited several times, as students perceived

learning R as a driver of workload. Course evaluations showed that despite R being challenging, students felt a sense of achievement and saw R as valuable after completing the course.

3.2. Course Project & Assessment

Consistent with the principles of situated learning (Vargas et al., 2024), students in both courses were assessed individually through a project that mimicked how data analytics teams solve real-world problems. In line with the principles of collaborative learning (Barak, 2017), students were permitted to collaborate on specific tasks, such as data preparation, and to support one another in overcoming challenges, such as debugging. However, despite working with the same dataset, each student was required to work independently. This included formulating their own business objective and corresponding data mining objective. Students also had to conduct their own modeling and evaluate their models. They could choose their own path by focusing on different target variables or by using distinct subsets of the data.

In the first year of the BsF&C course, student groups were allowed to select their own dataset. However, this frequently caused problems with (a) students who were trying to find a dataset that would be "easy" to analyze, and (b) students who were insecure and trying to make sure they found the best possible dataset. In the subsequent two years, lecturers cooperated with Tony's Chocolonely, a rapidly growing chocolate-producing social enterprise. This gave students a real-life case and a corresponding dataset. An important advantage was that all students received the dataset at the outset of the project. The brand was familiar to most students and appealing because of its focus on social sustainability. However, the dataset primarily focused on marketing and sales, and the students indicated they did not find it intrinsically motivating. Furthermore, students reported spending extensive time preparing the data, and that it was too large for their computers to handle. In hindsight, the dataset was too large for most students to handle and not relevant to their experience. In subsequent years, an in-between solution was chosen: a real-life dataset that better fit the students' domain expertise was "prepped" by the lecturers to make it more manageable for the students. A logistics-themed dataset from the second-hand clothing company Vinted was used. Similarly, the brand was familiar to students, thereby increasing their intrinsic motivation. Unlike in prior years, the lecturers conducted some data preparation in advance. This reduced students'

workload without affecting learning outcomes. Preparing the data for the students improved their intrinsic motivation to work on a real-life case, increased their self-efficacy, and reduced their frustration.

The project was assessed through an individual portfolio comprising a consulting report and a methods report. In the consulting report, students identified key business decisions and related metrics. They also summarized the analysis process and presented their key findings and recommendations. Their consulting reports also needed to include appropriate visualization. In a separate methods report, students detailed their analysis using Cross-Industry Standard Process for Data Mining (CRISP-DM), a broadly recognized framework (Data Science Process Alliance, 20-23; Wirth and Hipp, 2000). Initially, the MsSCM students submitted multiple separate assignments intended to test their knowledge of various data mining techniques. However, in the fall of 2024, the multiple assessments were replaced by a single portfolio to reduce students' perceived workload.

3.3 Instruction, Book & Other Materials

Both courses consisted of weekly teacher-led sessions and workshops. The weekly teacher-led sessions were used to present theoretical concepts and to link them to students' future professions through examples. The CRISP-DM framework was introduced to the students during the first week and was frequently used to structure projects and illustrate how data analytics could be linked to business problems. The business understanding and deployment steps were particularly relevant to the course objectives, as they bridged the gap between domain and technical expertise. Domain-specific examples (e.g., supplier selection and procurement) were used to illustrate this process. In subsequent weeks, students were gradually introduced to various statistical concepts, starting with descriptive statistics and regression, and then moving on to supervised and unsupervised data mining techniques. The instructor-led sessions focused on explaining theoretical concepts, illustrating them with domain-specific examples, and providing demonstrations. Students also had ample time to ask questions. The work sessions were self-led. The instructor began each session by explaining the purpose of the activities and addressing any questions. Students were provided with annotated R scripts corresponding to exercises in the prescribed book. The instructors reviewed the book publisher scripts and annotated

them to provide additional examples and explanations. As the course progressed, students had to fill in an increasing number of commands themselves.

As business analytics typically requires teams of experts from separate fields to work together, cooperation was integral to the course design (Vance et al., 2022). During the work sessions, students worked in groups of three. They were encouraged to solve problems together, but instructors were available to answer questions. In week 5, students were introduced to the project, and progressively more class time was allocated for them to work together and ask questions about it. Thus, students encountering problems could seek help and avoid getting discouraged. The book prescribed for the course was "Business Analytics" by Jaggia et al. (2020). This book was selected for its clarity of explanation, extensive examples of R-based analysis, and the provision of R scripts.

4 Methods

To assess the course design and student achievement, multiple methods were employed. Both universities of applied sciences have established PDCA processes at the course and program levels. For both courses, student roundtables were organized halfway through and at the end, supplemented by formal, anonymous written evaluations. Furthermore, both study programs have an industry advisory board, with industry professionals serving as external advisors. For example, at its June 2024 meeting, the MsSCM industry advisory board reviewed the scope, content, and relevance of the data analytics course. Discussion points included the choice to use R, the importance of discussing ethics, legislation, and data governance, and examples of data analytics applications. In addition to formal PDCA activities, students were asked to share their reflection reports (BsF&C) and/or participate in reflective interviews (MsSCM) at the end of the course, and at the end of their study program, resulting in a volunteer sample. The authors analyzed the anonymized roundtable discussion notes, course evaluation comments, industry advisory meeting notes, reflections, and interview transcripts. The full overview is shown in Table 1. It is important to note that the authors were also the lecturers of record, introducing the possibility of bias. Pedagogical advisors of the respective programs were therefore asked to review the data and independently verify the analysis and conclusions. This entailed, among other things, ensuring that all results received equal attention and that no social desirability bias was introduced through

selective data collection and interpretation. For all research activities, informed consent was obtained from all participating students. As all materials used were anonymized, ethics committee approval was waived by both institutions.

Table 1: Overview of information per cohort per study program

Cohort	BsF&C	MsSCM
Fall 2020	1st iteration, course designed by 1st and 2nd author. Course given online. Student feedback obtained during course execution through student roundtables, student written evaluation at the end of course, student roundtable at the end of the course (all organized by semester coordinator). Lecturer observation and reflections.	Course not offered.
Spring 2020	Course not offered.	1st iteration, authors not involved
Fall 2021	2nd iteration Course given in person by 3rd author and one colleague. Student feedback obtained through student written evaluation at the end of course, student roundtable at the end of the course (all organized by semester coordinator). Lecturer observation and reflections.	2nd iteration, course extensively revised by 1st author. Course given by 1st author in person. Course given in person, with the exception of the last two weeks. Student feedback obtained during course execution through student roundtables, student written evaluation at the end of course, student roundtable at the end of the course (all organized by program coordinator). Lecturer observation and reflections. (n =21)
Fall 2022	3rd iteration Course given in person by 3rd author and one colleague. Student feedback obtained through student written evaluation at the end of course, student roundtable at the end of the course (all organized by semester coordinator). Lecturer observation and reflections.	3rd iteration. Course given in person by the 1st author. Student feedback obtained during course execution through student roundtables, student written evaluation at the end of course, student roundtable at the end of the course (all organized by program coordinator). Lecturer observation and reflections. (n =28)
Fall 2023	4th iteration Course given in person by 3rd author and one colleague. Student feedback obtained through student written evaluation at the end of course, student roundtable at the end of the course (all organized by semester coordinator). Lecturer observation and reflections.	4th iteration Course given in person by the 1st author. Student feedback obtained during course execution through student roundtables, student written evaluation at the end of course, student roundtable at the end of the course (all organized by programme coordinator). Lecturer observation and reflections. (n =32)
Fall 2024	No change	No change

Source: Authors' Own Work

To examine the course's impact on data analytic thinking, data were collected in 2024 via a census approach to compare students' ratings of their data analytic thinking skills before and after the course (see Figure 3). For this purpose, an existing four-item scale of organizational data analytic capability was adapted to the individual learner context (Srinivasan and Swink, 2018). Sample items included "I easily combine and integrate information from many data sources to use in my decision making" and "I use advanced analytical techniques (e.g., simulation, optimization, regression) to improve decision making," anchored with 1 (completely disagree) to 7 (completely agree). The scale was tested for reliability using Cronbach Alpha and found to exhibit sufficient reliability (α=0.728, n=75).

5 Results

Across the two courses (BsF&C and MsSCM) and over the four years during which they were administered, student performance was consistently high, with passing rates ranging from 85% to 100% after a second opportunity. The majority of students obtained a grade equivalent to "very good" (8) or "good" (7) in the Dutch system, where 5.5 is considered a sufficient grade. A comparison with other courses offered concurrently showed that while students performed better in the data analytics course, they also exhibited higher degrees of attendance, and higher overall satisfaction with the course.

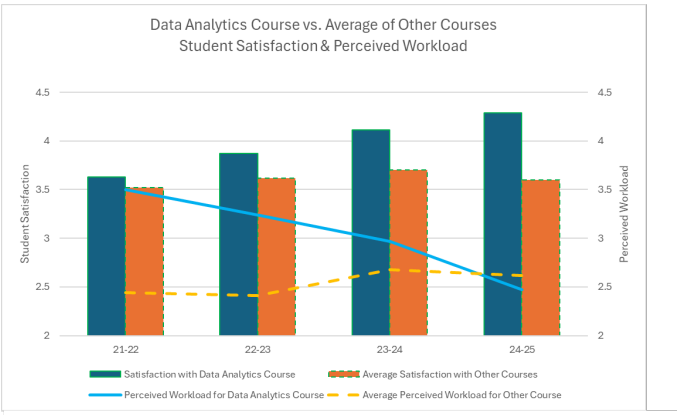

Figure 1: Comparison of Student Overall Satisfaction and Perceived Workload (MsSCM, cohorts 2-5)

Source: Own

As the course evaluation and perceived workload scores of the MsSCM across the last four cohorts indicate, students in earlier cohorts evaluated the course workload as higher than that of other concurrent courses, despite being more satisfied (Figure 1). Each year, measures were taken to decrease students' perceived workload. This was achieved primarily by reviewing formative and summative assessment moments and critically reviewing the assignment in relation to the learning objectives. As Figure 2 shows, the data analytics course contributed to improved students' data-analytic thinking, the primary learning outcome.

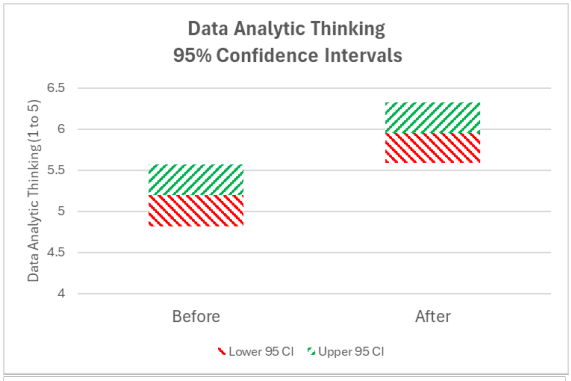

Figure 2 - Data analytic thinking MsSCM students before and after data analytics course (4th year bachelor students vs MsSCM alumni)

Source: Own

5 Lessons Learned

Taken together, the experiences of our students, our own observations, and those of other instructors documented in the academic literature yielded several best practices that could assist others in developing data analytics courses for business students. In the section below, we discuss these and use student quotes from the course evaluations, reflection reports, and individual interviews as illustrations.

5.1 Treat R as a Black Box Application

While one of the strengths of the software R is its versatility, the learning curve for students without a background in computer programming or data analytics is steep (Echeverría et al., 2017; Msweli et al., 2023). To motivate these students, initially treating R as a black box application will help keep them motivated (Firth et al., 2021). This can be achieved by providing students with annotated scripts that create beautiful and informative graphs, thereby contributing to a "*wow look how awesome*" self-efficacy effect. At the start of each course, it is necessary to confirm that these scripts remain bug-free, especially after software and package updates. Instructors also need to ensure that the scripts provided to students can accommodate different date formats and decimal systems, especially when teaching international students. Students can thereby experience the advantages of R, with a less steep learning curve, less frustration, and thus a heightened sense of self-efficacy. As one student indicated on the mid-term student evaluation: "*R was very hard and I doubt if we are going to use that.*" However, after completing the course, this early resistance was replaced with a sense of achievement, as the student had "conquered" something difficult, stating, "*I really liked that we were forced to obtain some programming skills.*"

5.2 Create an Atmosphere of Collaborative Learning

Encouraging students to work together helps them mimic the professional world in which they will be operating. Furthermore, shared difficulties such as dealing with bugs in R code create camaraderie. Consistent with social constructivist theory, collaborative learning is therefore an important tool in supporting students as they learn data analytics (Firth et al., 2021); Vance et al., 2022). However, it remains crucial for instructors to be available to help students get "unstuck" in real time. Thus, supervising students as they work together and practice the concepts is a good use of class time (Firth et al., 2021). The students appreciated the opportunity to ask questions in real time. One student mentioned: "*We were guided a lot. You need to struggle to learn. We were able to ask questions quickly in learning to program.*" Whenever possible, teaching assistants should support instructors in responding to student questions (Hicks and Irizarry, 2018).

5.3 Use Relevant but Manageable Datasets and Examples

Consistent with the situated learning perspective (Vargas et al., 2024), course objectives should be clearly linked to the students' future professions, and examples and exercises should relate to sectors in which they are likely to work. Datasets should be chosen carefully and sources from companies that are attractive to students in product, service, or mission. The students then feel that they are working on relevant and timely business issues. As one student indicated on the course evaluation, "*it was great to work on a real-world case which was very timely.*" However, care needs to be taken to offer datasets that match their skills, to ensure an appropriate balance between effort and learning outcomes. By working on relevant cases and datasets, students can apply their existing domain knowledge and skills to their data analytics exercises, helping them make sense of course content. Students thereby see the added value of being acquainted with R for their future profession. As one student explained, "*The added value of R is learning to code as a finance specialist. This is very useful, even the basics, when you talk with the business analyst who built the model, and you can discuss. What I saw during my internship is that working with large datasets is part of business nowadays. That is what a controller does.*" Ideally, instructors should also be domain experts with data analytics expertise as they can help students bridge the gap between their domain expertise and newly acquired data analytics knowledge (Hicks and Irizarry, 2018; Tucker et al., 2022), as difficult as that may be (Msweli et al., 2023).

5.4 Introduce and Emphasize the Use of a Framework

Students greatly benefited from the systematic and efficient workflow enabled by the application of the CRISP-DM framework (Msweli et al., 2023). It allows students to translate business needs and objectives into questions that could be answered with data, and, conversely, to realize that results only hold value if they are implemented in the business context. The structured approach of CRISP-DM helped students in developing data analytic thinking (Provost and Fawcett, 2013) by clearly distinguishing between the domain expertise they already possessed and the technical expertise they were acquiring (Wirth and Hipp, 2000). Students also appreciated the challenge of embracing a new approach, and how it led them to gain more confidence: "Another main thing that I learned from this project is that it is good to be thrown into the deep sometimes with a challenge, or project in this case, that requires a whole new approach than what we have been used to in the last

couple of years […]. I also feel that I can use this is my future career as well, because you can always get in a situation in which you do not specifically know how to approach it because your knowledge on the matter is scarce." Furthermore, graduates of the program reported that their knowledge of CRISP-DM was beneficial in their professional careers. As one graduate explained: "*it is great that when CRISP-DM comes up, I can contribute to the discussion*." For instructors who use their business experience to illustrate concepts, the CRISP-DM framework provides a way to structure their storytelling.

6 Conclusion and Future Work

Interest in data analytics is booming, and the need for business students to be proficient in this field remains strong. Academic programs aimed at domain experts should offer data analytics courses that leverage students' domain knowledge while accommodating their limited background in computer programming and statistics, thereby maintaining motivation through technically challenging yet accessible content. Drawing from the academic literature on the topic, student reflections, and our own experiences, we presented several lessons learned that may help other instructors develop similar courses. It is important to note that the course was developed in the context of the Dutch educational system, which is characterized by high industry involvement and oversight. The Dutch culture also values data-driven decision-making. Together, these factors ensured the needed legitimacy to invest resources in developing this course. This impetus and resources may not be available in other settings. However, we also acknowledge that the data science and data analytics fields are still developing rapidly, and therefore continuous monitoring and redesign of data analytics courses is necessary. For example, generative Artificial Intelligence (AI) tools (e.g. ChatGPT and Microsoft Copilot) are increasingly being used by students to help them code in R.

Our initial experience with allowing students to use AI to generate code is that it can help students who have already mastered basic programming concepts make progress more quickly. However, when students lack a solid understanding of basic R syntax, they struggle to use prompts effectively and to apply AI-generated code to their assignments. Although the quality of AI-generated code is rapidly improving, hallucinations (references to non-existent packages or functions) remain a persistent issue. This can lead to frustration among both students and instructors, as they often

are unaware that hallucination is at play. Based on anecdotal evidence, we anticipate that as AI-generated code continues to improve, it will become an increasingly relevant educational tool. For now, however, further research is needed to explore how generative AI can best support student learning and to examine its associated ethical implications (Becker et al., 2023). Future course development should therefore critically assess the added value of these tools and consider how to integrate them effectively to enhance student learning outcomes.

References

Barak, M. (2017), "Science Teacher Education in the Twenty-First Century: a Pedagogical Framework for Technology-Integrated Social Constructivism," Research in Science Education, 47, 283–303. https://doi.org/10.1007/s11165-015-9501-y.

Becker, B. A., Denny, P., Finnie-Ansley, J., Luxton-Reilly, A., Prather, J., and Santos, E. A. (2023), "Programming Is Hard - Or at Least It Used to Be: Educational Opportunities and Challenges of AI Code Generation," in Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1, Toronto ON Canada: ACM, 500–506. https://doi.org/10.1145/3545945.3569759.

Birt, J., Safari, M., and De Castro, V. B. (2023), "Critical analysis of integration of ICT and data analytics into the accounting curriculum: A multidimensional perspective," Accounting & Finance, 63, 4037–4063. https://doi.org/10.1111/acfi.13084.

Data Science Process Alliance. (2023). What is CRISP DM? https://www.datascience-pm.com/crisp-dm-2/

Deloitte (2020), "The Future of Controlling," CFO Insights.

Echeverría, L., Cobos, R., Machuca, L., and Claros, I. (2017), "Using collaborative learning scenarios to teach programming to non-CS majors," Computer Applications in Engineering Education, 25, 719–731. https://doi.org/10.1002/cae.21832.

Gharehgozli, A., Gupta, A., and Paik, S.-K. (2024), "Developing an undergraduate business analytics program for a public state-funded business school," Journal of Education for Business, 99, 11–19. https://doi.org/10.1080/08832323.2023.2248348.

Hicks, S. C., and Irizarry, R. A. (2018), "A Guide to Teaching Data Science," The American Statistician, 72, 382–391. https://doi.org/10.1080/00031305.2017.1356747.

Jaggia, S., Kelly, A., and Lertwachara, K. (2020), Business analytics, New York: McGraw-Hill Education.

Janssen, N. (2022), "The Data Science Talent Gap: Why It Exists And What Businesses Can Do About It," Forbes.

King, A. Z. (2022), "Data analytics in Association to Advance Collegiate Schools of Business–accredited U.S. university accounting programs: A quantitative research study," Journal of Education for Business, 97, 320–328. https://doi.org/10.1080/08832323.2021.1953430.

Kokkinou, A. (2023), "Planting the seed: introducing knowledge of digitalization using continuous improvement capabilities," in Operational Excellence in the era of Industry 4.0 and Industry 5.0 Proceedings from the Ninth International Conference on Lean Six Sigma, eds. J. Antony, J. Alsadi, A. Elmougi, M. Ramadan, and M. Bader, Newcastle, UK: Northumbria University, 37–54.

Kokkinou, A., Mandemakers, A., and Mitas, O. (2023), "Developing Resilient and Robust Supply Chains through Data Analytic Capability," Continuity & Resilience Review, 5, 320–342. https://doi.org/10.1108/CRR-07-2023-0013.

Firth, D.R., Triche, J., and Lucus, D.J. (2021) "A framework for teaching an undergraduate data analytics class," International Journal of Information and Operations Management Education, 7 https://doi.org/10.1504/ijiome.2021.10037382

Msweli, N. T., Mawela, T., and Twinomurinzi, H. (2023), "Transdisciplinary teaching practices for data science education: A comprehensive framework for integrating disciplines," Social Sciences & Humanities Open, 8, 100628. https://doi.org/10.1016/j.ssaho.2023.100628.

Pan, K., and Mazzei, M. (2023), "Preparing Accounting Students for Careers Utilizing Data Analytics: Curriculum Innovation Steps and Challenges," AIS Educator Journal, 18, 1–19. https://doi.org/10.3194/1935-8156-18.1.1.

Provost, F., and Fawcett, T. (2013), Data science for business: what you need to know about data mining and data-analytic thinking, Beijing Köln: O'Reilly.

"RStudio" (2021), RStudio, Available athttps://www.rstudio.com/products/rstudio/.

Sankaran, S., Sankaran, K., and Bui, T. (2023), "Student satisfaction with R vs. Excel in Data Mining and Business Analytics: A Herzberg's motivation-hygiene theory perspective," Decision Sciences Journal of Innovative Education, 21, 68–82. https://doi.org/10.1111/dsji.12285.

Srinivasan, R., and Swink, M. (2018), "An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective," Production and Operations Management, 27, 1849–1867. https://doi.org/10.1111/poms.12746.

Stander, J., and Dalla Valle, L. (2017), "On Enthusing Students About Big Data and Social Media Visualization and Analysis Using R, RStudio, and RMarkdown," Journal of Statistics Education, 25, 60–67. https://doi.org/10.1080/10691898.2017.1322474.

"The R Project for Statistical Computing" (2021), The R foundation.

Tucker, M. C., Shaw, S. T., Son, J. Y., and Stigler, J. W. (2022), "Teaching Statistics and Data Analysis with R," Journal of Statistics and Data Science Education, 1–15. https://doi.org/10.1080/26939169.2022.2089410.

U.S. News & World Report (2024), U.S. News & World Report 2024 US Best Jobs Report.

Vance, E. A., Alzen, J. L., and Smith, H. S. (2022), "Creating Shared Understanding in Statistics and Data Science Collaborations," Journal of Statistics and Data Science Education, 30, 54–64. https://doi.org/10.1080/26939169.2022.2035286.

Vargas, E., Chiappe, A., and Durand, J. (2024), "Reshaping education in the era of artificial intelligence: insights from Situated Learning related literature," Journal of Social Studies Education Research, Journal of Social Studies Education Research, 15, 1–28.

Wang, J., and Gu, L. (2018), "Teach MBA data science using R," Issues In Information Systems. https://doi.org/10.48009/2_iis_2018_65-71.

Wirth, R., and Hipp, J. (2000), "CRISP-DM: Towards a standard process model for data mining," Springer-Verlag London, UK.

Why Do Consumers Purchase from a Controversial Online Store? Reasons for Purchasing from Temu

Tiina Kaleton,[1] Tiina Kemppainen,[2] Lauri Frank[1]

[1] University of Jyväskylä, Faculty of Information Technology, Jyväskylä, Finland
tiina.e.kaleton@jyu.fi, lauri.frank@jyu.fi
[2] University of Jyväskylä, School of Business and Economics, Jyväskylä, Finland
tiina.j.kemppainen@jyu.fi

Chinese online retailers have recently received considerable media attention in Western countries, particularly regarding ethical concerns. These stores are often perceived as controversial and having a questionable reputation: while they offer many attractions and low prices, they also raise concerns about product quality and production practices, among other issues. This study examines the reasons for purchasing from online stores perceived as controversial or unethical, focusing on Temu, whose popularity has surged in recent years. The data was collected by interviewing seven Finnish Temu shoppers from the age range of 23-71 years. The findings suggest that controversial online stores' appeal cannot be explained by financial issues alone; rather, it emerges from the interplay of social validation, personal circumstances, experiential enjoyment, and well-designed platform features.

DOI
https://doi.org/10.18690/um.fov.4.2026.28

ISBN
978-961-299-152-4

Keywords:
controversial online stores, online purchasing, behavioral reasoning, Temu, qualitative study

University of Maribor Press

1 Introduction

In recent years, Chinese online retailers have rapidly gained popularity worldwide, while simultaneously attracting increasing public debate concerning issues such as ethics, product quality, and production practices. Temu is a rapidly grown online store platform established in 2022, with extremely low prices that have been aggressively marketed to consumers globally (Backlinko, 2026). Despite being labeled an "evil platform" (Liu & Gómez-Cruz, 2025) and ranked among the most disreputable companies in Finland (Helsingin Sanomat, 2025), Temu continues to attract consumers. Temu generated 1.6–2.3 billion monthly site visits and had more than 140 million monthly active users in the European Union in October 2025 (Backlinko, 2026). The coexistence of widespread criticism and strong consumer adoption highlights the need to understand why consumers continue to purchase from online retailers perceived as controversial. This study develops in-depth understanding of this phenomenon. Specifically, it examines consumers' reasons for purchasing from the Temu online store. Given the exploratory nature of the research, a qualitative approach is adopted, consisting of seven in-depth interviews with Finnish Temu shoppers.

Prior research has examined online stores characterized as controversial or lacking reputable practices. For instance, studies have addressed topics such as online retailers' ethics and repurchase intention (Yuniarti et al., 2022), fake reviews (Barbado et al., 2019), and fraudulent sellers and shill bidding (Fire et al., 2023). One prominent theme is the use of dark patterns (Voigt et al., 2021), which have been identified in online retail environments as mechanisms that may manipulate consumers through marketing strategies, interface design, informational transparency, and pricing (Brignull et al., 2023; Singh et al., 2025). However, irresponsible purchasing may also stem from consumers themselves, influenced by factors such as personal characteristics, limited knowledge, or attitudes (Kemppainen et al., 2021; Mayr et al., 2022; Nyrhinen et al., 2024). Previous research notes that engaging with a controversial marketplace may prompt customers to develop reasons that justify purchasing behaviour which can, in some instances, contradict their values. One way to explain controversial shopping is through the use of neutralization techniques (McGregor, 2008), whereby individuals justify ethically questionable behavior in order to maintain a positive self-concept (Sykes & Matza, 1957). Although Temu has been one of the most discussed controversial

online platforms in recent years, scholarly research on the platform remains limited. For example, Liu and Gómez-Cruz (2025) examined platform cynicism in the Temu application, focusing on user perceptions. Similarly, Roh and Jang (2025) investigated platform adoption in the context of Temu and AliExpress. In addition, Temu has been included in several studies as part of broader analyses of online platforms (Ding et al., 2026; Nielsen & Andersen, 2026).

2 Literature Background

To understand why consumers purchase from controversial online stores, it is necessary to examine the reasoning processes through which such behavior is justified. This study draws on Behavioral Reasoning Theory (BRT), which links beliefs, reasons, motives, intentions, and behaviour in explaining human decision-making (Westaby, 2005). Widely applied in marketing research, BRT emphasizes how individuals justify their actions through both *reasons for* and *reasons against* a particular behaviour (Sahu et al., 2020). According to the theory, beliefs and values shape reasons for or against a behaviour, which in turn influence attitudes or global motives, intentions, and ultimately behaviour (Sahu et al., 2020; Westaby, 2005). BRT's ability to explain reasons for and against certain behaviors has been applied in several studies, including those on remanufactured products (Belbag & Belbag, 2025), sustainable shopping (Habib et al., 2025), and mobile shopping adoption (Gupta & Arora, 2017). Prior research suggests that purchasing decisions are shaped by both enabling and inhibiting factors: while price savings strongly motivate shopping adoption (Gupta & Arora, 2017), financial constraints, social disapproval, personal compromises, and psychological barriers such as low motivation can act as reasons against purchasing (Habib et al., 2025). To conclude, Behavioral Reasoning Theory (BRT) helps explain consumers' reasons both for and against purchasing, which may stem from a wide range of life circumstances, personality-related factors, and situational influences. This study focuses specifically on the *reasons for* purchasing from a controversial online store.

Controversial online stores may contribute to facilitating purchases, yet consumers may also actively choose-or unintentionally gravitate toward-buying from questionable retailers. Companies may employ so-called 'dark online store' elements, where online platforms are designed to benefit the company rather than the user (Brignull et al., 2023). These design patterns may involve aggressive marketing and

emotionally charged cues, while exploiting user unawareness and cognitive biases to influence decision-making (Singh et al., 2025). These factors can lead to compulsive purchasing (Greenberg et al., 2021), impulsive shopping (Nyrhinen et al., 2024; Shahpasandi et al., 2020), or even shopping addiction (Nyrhinen et al., 2023; Zarate et al., 2023). Immersion in online shopping can even result in physical consequences, such as fatigue (Kaleton et al., 2026). Reasons for unethical customer behaviour have also been linked to personal attitudes, social influences, and egoistic motives (Mayr et al., 2022). Customers may exhibit self-indulgent consumption tendencies, resist certain stores or products, or experience trust issues and scepticism toward sustainable shopping (Kemppainen et al., 2021). In sum, reasons for purchasing from questionable retailers may originate from the company, the customer, or a combination of both.

3 Methodology

This study aimed to gain an in-depth understanding of why consumers purchase from online stores perceived as controversial; therefore, a qualitative approach was adopted. Temu was selected as the target online store, as it is one of the most popular online stores in Europe. The data were collected through semi-structured interviews which took place between June 2024 and February 2025. Six participants were recruited from a Finnish Temu-related Facebook group consisting of more than 30,000 members, and one participant was recruited in person. The selection criteria included individuals who had made a purchase from the Temu online store or application within the past three months. Thus, the participants represented "knowledgeable agents" who know what they are doing and can explain what they are doing, intending, and thinking (Gioia et al., 2013). No compensation was provided to the volunteer interviewees. Five interviews were conducted via Microsoft Teams video conferencing, while two were carried out by mobile phone due to technical issues experienced by the interviewees. All interviews were recorded and transcribed using the university's AI-based transcription service, after which the transcripts were carefully reviewed. The quotations presented in this study have been translated from Finnish into English.

The themes discussed with the participants included participant's online shopping habits, shopping experiences, the influence of others, and the visual and structural characteristics of the online store. The interview protocol was designed to elicit

detailed narratives through probing questions, encouraging participants to reflect on their interpretations in an iterative, dialogical format and thereby generating rich data on a sensitive phenomenon. In practice, the interviewer followed an interview guide and asked follow-up questions when appropriate. Issues such as child labor or product safety were addressed only if participants themselves raised them and wished to elaborate. The awerage length of the interviews was 71 minutes, which enabled participants to elaborate on their experiences in a rich and detailed manner. Participant and interview details are provided in Appendix 1. The data were analyzed using Atlas.TI qualitative analysis software. Behavioral reasoning theory by Westaby (2005) was guiding the analysis process by establishing focus on the reasons for Temu purchasing. The first step of analysis was careful familiarization of the collected data by reading it multiple times. In the first stage, reasons-related content was marked in the data, and after that, investigated again to find potential emerging patterns. After multiple rounds of analysis, the main themes describing the reasons for purchasing from a controversial store could be identified. These are explained in detail in the next section.

4 Findings

Our analysis identified five types of reasons that explain why consumers utilize contradictory online stores despite their negative reputation: 1) financial, 2) convenience-related, 3) hedonic, 4) social, and 5) platform-related reasons. These themes are discussed in detail below.

4.1 Financial reasons

Financial reasons were an important motivation for purchasing from Temu. These reasons related to financial considerations, including product prices, affordability, and participants' ability to meet their basic needs. The platform was perceived as offering significantly lower prices compared to domestic online and brick-and-mortar stores, making it highly desirable.

"The price was two or three euros. And if you think about buying the same products from X [a Finnish hypermarket chain], it would be like 20 or 30 euros […] I can understand why people think, 'well, I might as well buy everything from here since it's cheaper.'" [P2]

Overall, affordable pricing emerged as a key enabler of consumption for individuals facing financial constraints, supporting both essential needs and occasional non-essential purchases associated with a better lifestyle. For example, Participant 7 explained: "It allows me to live a decent retiree life at affordable prices."

4.2 Convenience-related reasons

Convenience-related factors were a key motivation for purchasing from Temu, highlighting the ease and low effort of online shopping. The ability to order from home and avoid physical stores was valued, particularly when crowd-related anxiety or physical limitations made in-store shopping burdensome.

"I honestly don't like going to those big supermarkets. It makes me anxious-I'm kind of a weird person like that. […] There's a bit of a difference when compared to shopping on Temu." [P7]

Temu was described as a good substitute for physical stores because it offers a wide range of products, including clothing, everyday items, home décor, and small electronics-all accessible from the comfort of home. While many products were seen as similar to those available in domestic retailers, trust in the platform remained conditional rather than absolute. Participants engaged in careful, product-specific evaluations, particularly avoiding items associated with potential health or safety risks. Purchases were justified through risk-minimizing reasoning, such as emphasizing limited exposure or indirect use. For example, participant 4 explained that dog nail clippers were acceptable to buy, as "the contact time is so short that it would not cause poisoning." These findings suggest that purchasing from Temu is not entirely indiscriminate or impulsive; rather, it involves active boundary-setting and rationalization, allowing consumers to balance convenience and low prices with perceived safety and ethical considerations.

4.3 Hedonic reasons

Temu shopping was also associated with hedonic aspects. Hedonic reasons refer to pleasure and enjoyment derived from the shopping experience, including entertainment, excitement, and visually stimulating environments. Participants described browsing the store as entertaining and engaging, thanks to the visually

stimulating and attention-grabbing interface. The homepage was perceived as inspiring and encouraged exploration of new products.

"It feels like 'Alice in Wonderland' when you go there […] I can spend even two hours just browsing everything they have. And all kinds of things might end up in the shopping cart […] Time flies surprisingly fast. You go in to get something specific, but the homepage offers everything from one end to the other… you start looking at one thing, and then you always drift on to the next." [P4]

"Temu is perhaps more captivating. First, there are more products, and the advertising is much flashier on the site-it's not nearly as restrained as what you are used to." [P3]

The anticipation associated with awaiting deliveries emerged as an affective component of the shopping experience, contributing to feeling of excitement. Temu shopping was likened to gambling, highlighting the uncertainty surrounding product quality and usability as a central element of enjoyment. As illustrated by Participant 3, receiving a usable product at a very low cost was experienced as a "win," reinforcing positive emotions and encouraging continued engagement with the platform.

"It's surprisingly exciting to wait and see what will actually arrive […] It was a bit like gambling—whether the clothes would be wearable or not. And of course, the price is so much lower than anywhere else, so if it turns out to be something you can actually use, it feels a bit like winning at a slot machine." [P3]

4.4 Social reasons

The social environment also played an important role in shaping purchasing decisions. Social factors included recommendations from friends and family, media exposure, and collective shopping experiences, all of which contributed to increased interest in and acceptance of Temu. However, a dual pattern emerged: while social influences often encouraged purchasing, they simultaneously imposed constraints on consumption. Social norms, expectations, and anticipated judgments prompted reflection and justification, at times limiting what or how much participants felt comfortable purchasing.

Recommendations and positive experiences shared by acquaintances encouraged initial engagement with Temu, occasionally transforming purchasing into a collective and socially enjoyable activity, such as group orders with family members. Positive peer experiences were particularly influential in counterbalancing the largely negative media coverage surrounding Chinese online stores. Such endorsements reduced suspicion and normalized purchasing by signaling that buying from Temu was common and acceptable. At the same time, participants questioned claims about unethical practices by comparatively evaluating Temu against other retailers, often concluding that there was insufficient evidence to justify avoiding the platform. Because Temu products were perceived as similar in origin to those sold by socially accepted outlets, purchasing from Temu was ultimately considered acceptable. Interestingly, ongoing media discussion and debate also appeared to spark curiosity about the platform, effectively functioning as "free advertising" for Temu.

"Well, the thing is, it's the same product. Not a similar one-the same product. A Finnish online store has ordered those same compression gloves from Asia… If that's how it is, then I'll order it from Temu." [P7]

"I started looking at something completely different there. And then, when I had heard the news that the clothes were produced cheaply using child labor, I became curious about what they were like." [P3]

Purchases were justified by framing global inequality and unfairness as issues beyond individual control and by emphasizing that all consumption choices involve ethical trade-offs. In addition, buying from controversial retailers was rationalized through appeals to necessity: under financial constraints, purchasing only essential items was considered morally acceptable. However, participants also reported carrying a "cloak of shame" associated with purchasing from Temu, reflecting the stigmatization attached to buying from controversial retailers. Anticipated social disapproval and the risk of moral judgment led some to conceal their purchasing choices and avoid discussing them openly: "I don't dare to mention the place I buy from, because the reaction might be that people start criticizing me" [P2].

4.5 Platform-related reasons

Platform-related factors also strongly attracted consumers to Temu. These included usability features, payment options, delivery practices, and policies that reduced the perceived risk of purchasing. The availability of the platform in the consumer's native language lowered barriers to use and was seen as a competitive advantage. Perceived risk was further mitigated by low prices, which limited potential financial loss, as well as by a 30-day payment term and trusted payment options such as Klarna. Free delivery functioned as a salient incentive, actively shaping purchasing behavior. Participants adjusted order sizes to reach the €25 delivery threshold, sometimes coordinating joint orders with family members. In addition, Temu's satisfaction guarantee-allowing refunds without returning products-was viewed as a significant advantage compared to other online shops, where returns are often cumbersome and costly.

"I have ordered a lot, and I am really satisfied. If there is something wrong with the delivery, I immediately receive a text message saying that I don't need to return the item—they will refund the purchase. I have been very pleased with that." [P1]

The Temu website and application were generally experienced as easy to use, with effective search functions reducing cognitive effort once users became familiar with the platform. In addition, frequent discounts, campaigns, and occasional free or promotional items enhanced the perceived value of shopping.

5 Discussion

This study examined consumers' reasons for purchasing from controversial online stores. The findings indicate that the reasons are multifaceted and can be broadly categorized into social, personal, and platform-related factors. An overview of these findings is presented in Figure 1.

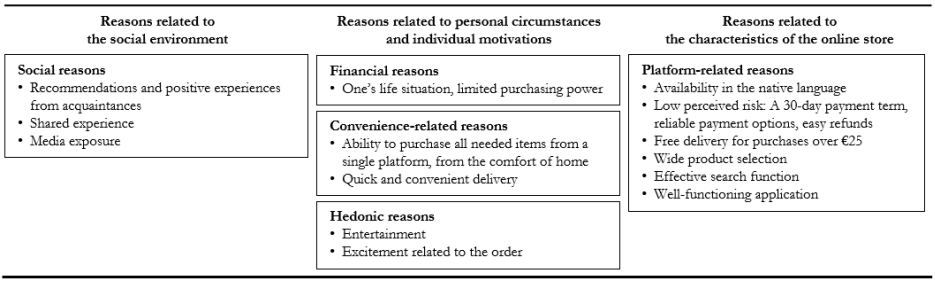

Figure 1: Summary of the findings

Source: Own

According to the findings, the social environment can encourage, but at the same time limit, purchases from controversial stores, reflecting the coexistence of reasons for and reasons against considerations as proposed by BRT (Westaby, 2005). Recommendations and positive experiences from acquaintances, shared shopping experiences, and media exposure encouraged Temu purchasing. These results align with previous research showing that peer influence and social proof can reduce hesitation when engaging retailers (e.g., Chevalier & Mayzlin, 2006). On the other hand, shopping from controversial stores involves considerable reflection, with purchase decisions often requiring justification both internally and in social interactions. This process demonstrates the use of neutralization techniques, through which individuals rationalize ethically questionable behavior in order to maintain a positive self-concept (Sykes & Matza, 1957). For instance, purchasing was commonly legitimized through defence of necessity (McGregor, 2008), as financial constraints were used to justify purchases, and through condemnation of the condemners (McGregor, 2008), whereby criticism was deflected by emphasizing that many others shop from the platform and that identical products are sold by socially accepted domestic retailers at higher prices. Previous research suggests that consumers may face disapproval or criticism if they do not engage in conscious purchasing practices (Habib et al., 2025). Similar dynamics emerged in this study, where criticism and purchase-related shame were evident. However, these social pressures did not ultimately appear to influence or deter purchasing decisions, even though previous research suggests that exposing customers to feelings of guilt can sometimes encourage more responsible purchasing (Makkonen et al., 2023).

Previous research has also demonstrated that irresponsible purchasing is influenced by personal characteristics, limited knowledge, or attitudinal factors (Mayr et al., 2022; Kemppainen et al., 2021; Nyrhinen et al., 2024). The findings of this study support these suggestions. Individual circumstances and personal motivations were central for shopping in a controversial online store. Financial factors, such as limited purchasing power or budget constraints, make Temu's low prices particularly appealing. Similarly, Gupta and Arora (2017) identified potential cost savings as the primary driver of mobile shopping adoption. Convenience-related reasons further reinforce this behavior, as consumers value the ability to shop for all needed items from home with easy delivery. Additionally, hedonic factors-including entertainment and the excitement associated with a delivery-added an experiential dimension to the purchases, suggesting that online shopping can carry emotional rewards.

Finally, the store itself can also act as an attraction. Platform-related characteristics such as availability in the native language, reliable payment options, easy refunds, free delivery thresholds, a wide product selection, and a well-functioning search function enhanced the shopping experience. These findings do not reflect online stores' dark patterns, as defined in previous research (Singh et al., 2025; Voigt et al., 2021), per se. However, they may operate as boundary-blurring mechanisms that reduce perceived risk and ethical friction, thereby facilitating purchasing decisions. In particular, free delivery thresholds and refunds without return encouraged consumers to increase their purchase volume, suggesting that such design elements can assume dark pattern–like characteristics by subtly nudging users to spend more than they initially intended.

Together, the findings suggest that consumers' appeal to purchase from a controversial online store cannot be explained by pricing alone, nor by manipulative platform design. Instead, this attraction appears to emerge from an interplay of social validation, personal circumstances, experiential enjoyment, and well-designed platform features. Overall, the findings enhance understanding of how individuals justify their behavior and how moral ambivalence, everyday constraints, and enjoyment coexist in such purchasing decisions.

As this research is qualitative, the findings are descriptive and grounded in individual participants' experiences. Moreover, the study captures the perspectives of relatively informed and engaged users who appear to be positively disposed toward the platform, which may limit the diversity of viewpoints and underrepresent more critical or disengaged consumers. In addition, the focus of the study was on reasons for purchasing rather than reasons against it, reflecting only one side of the behavioral reasoning process (Westaby, 2005). Consequently, future research would benefit from a more balanced examination of both enabling and inhibiting reasons. In-depth investigation of guilt neutralization processes (Sykes & Matza, 1957) related to controversial purchasing would be a particularly valuable avenue for future research to better understand the underlying consumer behavior. Finally, further studies employing different methods, customer demographics, and cultural contexts are needed to deepen the understanding of the investigated phenomenon.

Acknowledgment

Declaration of AI. Microsoft Copilot (M365 Copilot based on GPT-5) was used for interview citation translation from Finnish to English, and language-editing purposes, improving the clarity of the text. University's own internal AI transcription service was used in interview data transcription, and the quality of the transcriptions was carefully checked afterwards. This study was funded by Business Finland (grant number 6659/31/2023).

References

Backlinko. (2026). Temu revenue and usage statistics for 2026. Backlinko. Retrieved February, 20, 2026 from https://backlinko.com/temu-stats

Barbado, R., Araque, O., & Iglesias, C. A. (2019). A framework for fake review detection in online consumer electronics retailers. *Information Processing & Management*, 56(4), 1234–1244. https://doi.org/10.1016/j.ipm.2019.03.002

Belbag, A. G., & Belbag, S. (2025). Consumer intentions to purchase remanufactured products: A behavioral reasoning theory approach. *Journal of Social Marketing*, 15(1), 1–19. https://doi.org/10.1108/JSOCM-02-2024-0033

Brignull, H., Leiser, M., Santos, C., & Doshi, K. (2023). Deceptive patterns – user inter-faces designed to trick you. deceptive.design. Retrieved February 25, 2026, from https://www.deceptive.design/

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, 43(3), 345-354. https://doi.org/10.1509/jmkr.43.3.345

Ding, J., Wang, Y., & Laufer, D. (2026). The role of social context in live streaming commerce – an interpersonal perspective. *Asia Pacific Journal of Marketing and Logistics*, A-head-of-print.(A-head-of-print.). https://doi.org/10.1108/APJML-06-2025-1207

Fire, M., Puzis, R., Kagan, D., & Elovici, Y. (2023). Large-Scale Shill Bidder Detection in E-commerce. In Proceedings of *the International Database Engineered Applications Symposium Conference*, 79–86. https://doi.org/10.1145/3589462.3589479

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, 16(1), 15–31. https://doi.org/10.1177/1094428112452151

Greenberg, N. R., Zhai, Z. W., Hoff, R. A., Krishnan-Sarin, S., & Potenza, M. N. (2021). Problematic shopping and self-injurious behaviors in adolescents. *Journal of Behavioral Addictions*, 9(4), 1068–1078. https://doi.org/10.1556/2006.2020.00093

Gupta, A., & Arora, N. (2017). Understanding determinants and barriers of mobile shopping adoption using behavioral reasoning theory. *Journal of Retailing and Consumer Services*, 36, 1–7. https://doi.org/10.1016/j.jretconser.2016.12.012

Habib, M. D., Attri, R., Salam, M. A., & Yaqub, M. Z. (2025). Bright and dark sides of green consumerism: An in-depth qualitative investigation in retailing context. *Journal of Retailing and Consumer Services*, 82, 104145. https://doi.org/10.1016/j.jretconser.2024.104145

Helsingin Sanomat. (2025, 3.10.2025). Näitä yrityksiä suomalaiset kammoavat. Helsingin Sanomat Visio. https://www.hs.fi/visio/art-2000011498228.html

Kaleton, T., Holkkola, M., Kemppainen, T., & Frank, L. (2026). What Makes or Breaks an Immersive Online Customer Journey? Digital Native Generation Z Customers' Experiences. In Proceedings of *the 59th Hawaii International Conference on System Sciences*, 2026, 953–962.

Kemppainen, T., Frank, L., Makkonen, M., & Hyvönen, O.-I. (2021). Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers. In *the Proceedings of the 34th Bled eConference Digital Support from Crisis to Progressive Change*, 323–335. https://doi.org/10.18690/978-961-286-485-9.24

Liu, S., & Gómez-Cruz, E. (2025). "Temu is evil, but I still use it": Platform cynicism in platform capitalism. *New Media & Society,* 1–20.

Makkonen, M., Frank, L., & Paananen, T. (2023). The Role of Anticipated Guilt and its Neutralisation in Explaining Responsible Online Shopping. In Proceedings of *the 36th Bled eConference – Digital Economy and Society: The Balancing Act for Digital Innovation in Times of Instability*, 595–614. https://doi.org/10.18690/um.fov.6.2023.38

Mayr, K., Schwendtner, T., Teller, C., & Gittenberger, E. (2022). Unethical customer behaviour: Causes and consequences. *International Journal of Retail & Distribution Management*, 50(13), 200–224. https://doi.org/10.1108/IJRDM-06-2022-0194

McGregor, S.L.T. (2008), Conceptualizing Immoral and Unethical Consumption Using Neutralization Theory. Family and Consumer Sciences Research Journal, 36, 261–276. https://doi.org/10.1177/1077727X07312190

Nielsen, M. D., & Andersen, K. R. (2026). Exploring the dynamics of fashion resale: A critical typology of resale business models in Denmark. *Journal of Fashion Marketing and Management: An International Journal*, 30(3), 593–612. https://doi.org/10.1108/JFMM-03-2025-0146

Nyrhinen, J., Lonka, K., Sirola, A., Ranta, M., & Wilska, T.-A. (2023). Young adults' online shopping addiction: The role of self-regulation and smartphone use. *Int J Consum Stud*, 47, 1871–1884.

Nyrhinen, J., Sirola, A., Koskelainen, T., Munnukka, J., & Wilska, T.-A. (2024). Online antecedents for young consumers' impulse buying behavior. *Computers in Human Behavior,* 153, 108129. https://doi.org/10.1016/j.chb.2023.108129

Roh, M., & Jang, H. (2025). Evolution of adoption determinants in cross-border e-commerce: Comparing AliExpress and Temu. *Asia Pacific Journal of Marketing and Logistics*, A-head-of-print.(A-head-of-print.). https://doi.org/10.1108/APJML-01-2025-0028

Sahu, A. K., Padhy, R. K., & Dhir, A. (2020). Envisioning the Future of Behavioral Decision-Making: A Systematic Literature Review of Behavioral Reasoning Theory. *Australasian Marketing Journal*, 28(4), 145–159. https://doi.org/10.1016/j.ausmj.2020.05.001

Shahpasandi, F., Zarei, A., & Nikabadi, M. S. (2020). Consumers' Impulse Buying Behavior on Instagram: Examining the Influence of Flow Experiences and Hedonic Browsing on Impulse Buying. *Journal of Internet Commerce*, 19(4), 437–465. https://doi.org/10.1080/15332861.2020.1816324

Singh, V., Vishvakarma, N. K., & Kumar, V. (2025). Profit over principles: Unveiling the motivating factors behind dark patterns in e-commerce through the lens of agency theory. *Journal of Enterprise Information Management*, 38(3), 821–848. https://doi.org/10.1108/JEIM-08-2024-0409

Sykes, G. M., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. American Sociological Review, 22(6), 664–670.

Voigt, C., Schlögl, S., & Groth, A. (2021). Dark Patterns in Online Shopping: Of Sneaky Tricks, Perceived Annoyance and Respective Brand Trust. In *8th International Conference, HCIBGO 2021*, 12783, 143–155. https://doi.org/10.1007/978-3-030-77750-0_10

Westaby, J. D. (2005). Behavioral reasoning theory: Identifying new linkages underlying intentions and behavior. *Organizational Behavior and Human Decision Processes*, 98(2), 97–120. https://doi.org/10.1016/j.obhdp.2005.07.003

Yuniarti, F., R. Arief, H., Meydia, H., & Yevis, M. O. (2022). Online retailers' ethics and its effect on repurchase intention: The mediating role of perceived risk. *Cogent Business & Management*, 9(1), 2051691. https://doi.org/10.1080/23311975.2022.2051691

Zarate, D., Fullwood, L., Prokofieva, M., Griffiths, M. D., & Stavropoulos, V. (2023). Problematic Shopping Behavior: An Item Response Theory Examination of the Seven-Item Bergen Shopping Addiction Scale. *International Journal of Mental Health and Addiction*, 21(6), 4108–4126. https://doi.org/10.1007/s11469-022-00844-8

Appendix 1: Overview of Participants and Interviews

	Age	Gender	Living environment	Purchased Temu product(s)	Transaction channel and device	Dura-tion (min)
P1	71	Female	Rural	Jewellery, underwear, socks, clothing, dog toys, men's swimsuit	Application, mobile phone	56
P2	23	Non-binary	Urban	Eyebrow trimmer razors, hair bands, elastic hair bands	Application, mobile phone	82
P3	28	Male	Urban	Clothing, golf gear, socks, baseball caps, sunglasses	Online store, mobile phone	63
P4	54	Female	Urban	Clothing, crafting supplies, wall hook, wall stickers, bracelet beads, wall decoration, metallic toy cars, mug	Application, mobile phone	77
P5	54	Female	Rural	Vacuum spare parts, clothing, sandals, socks, essential oils, makeup, cat toys, make up sponges, solar cell lights, balloons	Application, mobile phone	70
P6	54	Female	Urban	Clothing, stickers, flashlights, pens, decorations, blankets, cat products, color catcher towels, smartwatch	Online store, laptop	63
P7	67	Female	Urban	Clothing, home items, decoration, health products, shoes, lamps, phone, headphones, kitchen products	Online store, laptop	87

Impact of Digital Literacy on Turnover Intention Among Slovenian Healthcare Professionals

Nejc Bernik, Polona Šprajc
University of Maribor, Faculty of Organisational Sciences, Kranj, Slovenia
nejc.bernik1@um.si, polona.sprajc@um.si

Digital literacy has become an essential competence in healthcare systems undergoing rapid digital transformation. As digital technologies and information systems become increasingly integrated into clinical practice, healthcare professionals must possess adequate digital competencies to perform their tasks effectively. At the same time, workforce stability remains a significant challenge for healthcare organisations, making turnover intention an important issue for healthcare management. This study examines the relationship between digital literacy and turnover intention among healthcare professionals working in Slovenian public healthcare organisations. A quantitative research design was employed, using a 5-point Likert-scale survey questionnaire to measure digital literacy and turnover intention. The final sample comprised 129 healthcare professionals employed by public healthcare organisations in Slovenia.

DOI
https://doi.org/10.18690/um.fov.4.2026.29

ISBN
978-961-299-152-4

Keywords:
digital literacy, turnover intention, healthcare professionals, digital technology, public healthcare organisations

University of Maribor Press

1 Introduction

Public healthcare systems are increasingly undergoing digital transformation as digital technologies (DTs) and information systems become integrated into everyday clinical practice. Electronic health records, digital platforms, and interconnected health information systems are changing how healthcare organisations (HCOs) operate and deliver services. In this context, digital literacy (DL) has become an important competence for healthcare professionals (HCPs), enabling them to use DTs effectively and perform clinical and administrative tasks in digitalised work environments.

At the same time, public HCOs face persistent workforce challenges, including staff shortages and increasing turnover intention (TI) among HCPs. Understanding the factors that may influence employee retention has therefore become an important issue for public healthcare management. While previous research has focused mainly on organisational and work-related determinants of TI, the role of DL in this relationship remains less explored.

This study examines the relationship between DL and TI among HCPs working in Slovenian public HCOs. Using a preliminary quantitative research design based on survey data, the study investigates whether higher levels of DL are associated with lower TI. By addressing this relationship, the paper contributes to the broader discussion on digital transformation and workforce sustainability in public healthcare systems.

2 Slovenian public healthcare organisations

In the Slovenian context, public HCOs represent a structured and hierarchical system embedded within a publicly funded, state-regulated healthcare model, primarily governed by the Health Care and Health Insurance Act (*Zakon o Zdravstvenem Varstvu in Zdravstvenem Zavarovanju (ZZVZZ) (PISRS)*, n.d.). The system operates on the principle of universal access, with the Health Insurance Institute of Slovenia acting as the central financier of public HCOs and their services (Domej, 2019). Individual public HCOs operate under the Ministry of Health's coordination and are required to comply with national standards on quality, safety, and accessibility.

In Slovenia, the public healthcare system is described by (*Zdravstveni Sistem v Sloveniji*, n.d.) includes two university medical centres, ten general hospitals, sixteen specialised hospitals, sixty-four community health centres, and five specialised national public health institutions: the Blood Transfusion Centre of Slovenia, the National Institute of Public Health, the National Laboratory of Health, Environment and Food, Slovenia Transplant, and the Youth Health and Climate Resort Rakitna. As those public HCOs represent a critical component of national infrastructure, the performance of HCPs and the execution of their work tasks must be carried out reliably and without errors, as the health and well-being of Slovenian citizens depend on the quality and continuity of these services.

To perform their work tasks effectively, HCPs must communicate, collaborate, and support one another. Effective communication facilitates the timely exchange of information and coordination of work processes within healthcare teams. Moreover, as a growing share of healthcare work relies on DTs, HCPs must possess at least basic DL to perform administrative and clinical tasks efficiently and ensure the accurate use of DTs

3 Digital literacy and turnover intention

DL constitutes an essential competence for working effectively in contemporary digitalised environments (Vodă et al., 2022). It is commonly defined as an individual's capacity to use DTs in a deliberate, critical and effective manner to access, evaluate, create and communicate information, and to participate responsibly in digital environments (Reddy et al., 2020). This competence is particularly important in public HCOs, which have increasingly undergone digital transformation through the integration of advanced information systems and digital solutions (Muhammad Saif ud Din, 2025).

Public HCOs are progressively implementing national and international digital information systems to enhance patient-centred care and improve the accessibility of healthcare services (Awad et al., 2021; Galazzi et al., 2025). At the EU level, digital transformation in healthcare is reflected in the development of eHealth, which refers to the structured use of DTs and data-driven solutions to support healthcare delivery, public health activities, health system governance, and cross-border health services (Carradinha et al., 2026). These initiatives aim to improve the accessibility,

quality, efficiency, and long-term sustainability of public healthcare systems across EU Member States (*Digital Health and Care - Public Health - European Commission*, 2025).

Slovenia is one of the EU Member States that has significantly advanced the implementation of DTs in healthcare over the past decade. Particularly after 2015, when the National Institute of Public Health assumed coordination of the national eHealth project, substantial progress was achieved (*eZdravje - eZdravje*, 2019). With financial support from the EU, several key digital solutions were introduced, including the national digital platform zVEM, which enables Slovenian citizens to access their personal health data online (*zVEM - eZdravje*, 2020), and the Central Registry of Patient Data (CRPD). The latter represents Slovenia's national electronic health data repository, which securely stores and integrates patients' medical records to enable authorised HCPs to access comprehensive clinical information across healthcare providers (*CRPP - eZdravje*, 2020). These systems contribute to improved interoperability, accessibility, and efficiency within the Slovenian public healthcare system.

The difference between zVEM and the CRPD lies primarily in their intended users and functional purposes. While zVEM is designed to enable patients to access and manage their personal health data, the CRPD is an integrated national clinical information system intended exclusively for authorised HCPs. It aggregates and stores structured medical documentation, such as discharge summaries, specialist reports, laboratory results, and imaging records, enabling HCPs to obtain a comprehensive overview of a patient's medical history across different public HCOs (*How Did Slovenia Modernise Its eHealth System?*, 2024). The effective use of the CRPD requires a sufficient level of DL, as HCPs must navigate complex digital interfaces, accurately interpret electronic health data, ensure data security, and integrate digital information into clinical decision-making.

For HCPs to attain and maintain an adequate level of DL, systematic training and continuous professional development must be ensured within HCOs (Shiri et al., 2023). In this context, human resource management (HRM) plays a pivotal role in designing, implementing, and evaluating structured educational programmes that enable HCPs to use information systems, such as the CRPD, effectively and responsibly.

By providing targeted training, ongoing support, and opportunities for skills development, HRM contributes to strengthening digital competence, professional confidence, and operational efficiency. When HCPs are adequately equipped to navigate digital information systems and other job-related DTs, and to integrate electronic health data into clinical practice, they are more likely to perform their duties effectively and experience lower levels of work-related strain. Such organisational support may, in turn, contribute to lower TI by enhancing digital competencies.

4 Methodology

This study employs a quantitative research design to examine whether a relationship exists between DL and TI among Slovenian HCPs. At the time of writing, data collection was still ongoing; however, 135 responses had already been collected. Six respondents reported employment in private HCOs and were therefore excluded from the analysis, as the study focuses exclusively on public HCOs. The final dataset thus included 129 completed questionnaires from HCPs working in public HCOs. Accordingly, the results presented in this paper should be interpreted as preliminary, and the scope of conclusions is limited to the current sample.

The survey consisted of a demographic section and a main section that measured DL and TI using a 5-point Likert scale. DL was assessed using a 10-item scale adapted from Ng (2012), which captures individuals' competencies in using DTs in a critical and effective manner. The scale is structured across three dimensions: technical, cognitive, and socio-emotional, reflecting the multifaceted nature of DL. TI was measured using a three-item scale adapted from Dalessio et al. (1986), employing Likert-type items that capture respondents' intention to leave their current HCO.

Demographic questions included respondents' age, gender, and total years of work experience in public HCOs.

The study addresses the following research question:

RQ: Is there a negative linear relationship between DL and TI among publicly employed HCPs?

This question assumes that DL may represent an important organisational competence that contributes to reducing TI, thereby informing future workforce retention strategies in public HCOs.

5 Results

This section presents the study's empirical findings. First, the demographic characteristics of the respondents are described, including gender, age, and work experience. This is followed by the analysis of the relationship between DL and TI among HCPs employed in Slovenian public HCOs.

The final sample consisted of 129 Slovenian HCPs employed in public HCOs. Most of them were women (N = 110; 85.3%), while men represented 19 respondents (14.7%) of the sample (Table 1)

Table 1: Gender, work experience, age of HCPs

Gender					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	M	19	14,7	14,7	14,7
	W	110	85,3	85,3	100
	Total	129	100	100	
Work years					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1-10	57	44,2	44,2	44,2
	11-20	32	24,8	24,8	69
	21-30	20	15,5	15,5	84,5
	31-40	18	14	14	98,4
	41-45	2	1,6	1,6	100
	Total	129	100	100	
Age groups					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20-30	33	25,6	25,6	25,6
	31-40	47	36,4	36,4	62
	41-50	21	16,3	16,3	78,3
	51-60	26	20,2	20,2	98,4
	61-65	2	1,6	1,6	100
	Total	129	100	100	

Source: own

About age distribution, most respondents were between 31 and 40 years old (N = 47). This group was followed by respondents aged 20–30 years (N = 33), 51–60 years (N = 26), 41–50 years (N = 21), and 61–65 years (N = 2).

Regarding work experience, the largest proportion of respondents had 1–10 years (N = 57). This was followed by respondents with 11–20 years of experience (N = 32), 21–30 years (N = 20), 31–40 years (N = 18), and 41–45 years (N = 2).

The primary objective of the analysis was to examine whether a negative linear relationship exists between DL and TI among Slovenian HCPs. To test this relationship, the variables DL_average and TI_average were calculated as the mean item scores for each construct (Table 2).

A one-tailed Pearson correlation analysis was conducted to examine the linear relationship between the two variables. The hypotheses tested were:

H_0: There is no linear relationship between DL and TI.
H_1: There is a negative linear relationship between DL and TI.

Table 2: Linear correlation analysis

Correlations			
		DL_average	TI_average
DL_average	Pearson Correlation	1	-0,031
	Sig. (1-tailed)		0,036
	N	129	129
TI_average	Pearson Correlation	-0,031	1
	Sig. (1-tailed)	0,036	
	N	129	129

Source: own

The results in Table 2 indicate a weak negative correlation between DL and TI ($r = -0.031$). Although the effect size is negligible, the relationship reached statistical significance ($p = 0.036$). This suggests that the direction of the relationship is consistent with the theoretical expectation that higher levels of DL are associated with lower TI; however, the strength of the association is very limited within the current sample.

Accordingly, while the statistical test formally allows for the rejection of the null hypothesis (H_0) in favour of H_1, this conclusion should be interpreted with caution. The negligible effect size indicates that the relationship has minimal practical or explanatory relevance. Therefore, the results provide only limited empirical support for the hypothesised relationship.

These findings are preliminary and based on a relatively small sample (N = 129), as data collection was still ongoing at the time of analysis. The limited sample size may have affected the stability of the estimates and attenuated the observed effect. It is therefore plausible that the relationship between DL and TI may be more pronounced in larger and more representative samples.

6 Discussion

This study examined the relationship between DL and TI among HCPs working in Slovenian public HCOs. As healthcare systems increasingly undergo digital transformation, the ability of HCPs to use DTs effectively represents an important professional competence (Ferreira & Magalhães, 2026). The findings of this study contribute to the ongoing discussion on how digital competencies, particularly DL, may influence workforce stability within healthcare systems (Levin-Zamir, 2023).

The empirical results indicate a negative relationship between DL and TI, which is consistent with the theoretical assumption that higher levels of digital competence are associated with lower TI. Although the observed effect size is negligible ($r = -0.031$), the relationship was statistically significant ($p < 0.05$). This suggests that the direction of the relationship aligns with theoretical expectations, even though the association's strength remains limited in the current sample.

It is important to emphasise that these findings are preliminary and based on a relatively small sample (N = 129), as data collection was still ongoing at the time of analysis. The limited sample size may have influenced the stability of the estimates and attenuated the observed effect size. It is therefore plausible that the relationship between DL and TI may be more pronounced in larger and more representative samples. Future studies are needed to validate and further examine this relationship under more robust empirical conditions.

Importantly, the findings indicate that DL should not be interpreted as a strong standalone predictor of TI. Rather, DL appears to function as an enabling competence that may exert its influence indirectly through other organisational and psychological mechanisms. In line with contemporary organisational research, it is likely that DL contributes to workforce stability through mediating variables such as job satisfaction, perceived organisational support, work-related stress, or employee

engagement. This highlights the need to conceptualise DL within more complex explanatory frameworks rather than as an isolated determinant.

Despite the modest effect size, the study reinforces the strategic relevance of DL in digitally transforming healthcare environments. As DTs become increasingly integrated into clinical practice, HCPs must possess adequate digital competencies to perform their work effectively and confidently. Organisational practices that support the development of DL, such as structured training programmes, continuous professional development, and supportive HRM policies, may therefore contribute to improved employee retention.

The findings offer a valuable foundation for future research and emphasise the importance of strengthening DL as part of broader strategies to ensure the long-term sustainability and resilience of public HCOs in Slovenia and other EU Member States with similar public healthcare systems.

References

Awad, A., Trenfield, S. J., Pollard, T. D., Ong, J. J., Elbadawi, M., McCoubrey, L. E., Goyanes, A., Gaisford, S., & Basit, A. W. (2021). Connected healthcare: Improving patient care using digital health technologies. *Advanced Drug Delivery Reviews*, *178*, 113958. https://doi.org/10.1016/j.addr.2021.113958

Carradinha, M., Mendes, V. I. S., Moura, R. P., Silva, N. P., Rocha, L., Gonçalves, J., Antunes, I., Schiza, E., Pattichis, C. S., Zanini, A., Pajić, V., & Pinto, C. S. (2026). Telemedicine and the European Health Data Space: A new paradigm for healthcare in the EU. *Frontiers in Digital Health*, *7*, 1713758. https://doi.org/10.3389/fdgth.2025.1713758

CRPP - eZdravje. (2020, March 12). https://ezdrav.si/resitve/crpp/

Dalessio, A., Silverman, W. H., & Schuck, J. R. (1986). Paths to Turnover: A Re-analysis and Review of Existing Data on the Mobley, Horner, and Hollingsworth Turnover Model. *Human Relations*, *39*(3), 245–263. https://doi.org/10.1177/001872678603900305

Digital health and care—Public Health—European Commission. (2025, October 21). https://health.ec.europa.eu/ehealth-digital-health-and-care/digital-health-and-care_en

Domej, J. (2019, May 22). Recourse claims of the Health Insurance Institute of Slovenia (ZZZS) and Pension and Disability Insurance Institute of Slovenia (ZPIZ) against Employers. *Jadek & Pensa*. https://jadek-pensa.si/en/recourse-claims-of-the-health-insurance-institute-of-slovenia-zzzs-and-pension-and-disability-insurance-institute-of-slovenia-zpiz-against-employers/

eZdravje—eZdravje. (2019, December 12). https://ezdrav.si/ezdravje/

Ferreira, J. P., & Magalhães, T. (2026). Instruments to assess the digital health competencies of healthcare professionals: A scoping review. *Frontiers in Public Health*, *13*, 1726452. https://doi.org/10.3389/fpubh.2025.1726452

Galazzi, A., Fonda, F., Chiappinotto, S., Justi, L., Sønderskov Frydensberg, M., Lehmann Boesen, R., Macur, M., de San Pedro, M., Reixach Espaulella, E., & Palese, A. (2025). Recommendations to promote the digital healthcare transformation in the clinical practice: Findings from an

international consensus development method. *BMC Health Services Research*, *25*, 929. https://doi.org/10.1186/s12913-025-13079-9

How did Slovenia modernise its eHealth system? - Better. (2024, January 31). https://www.better.care/blog-en/how-did-slovenia-modernise-its-ehealth-system/

Levin-Zamir, D. (2023). Digital health literacy and health technology in health systems and beyond: The importance of measurement, planned action, and policy for readiness and sustainability1. *Information Services and Use*, *43*(2), 143–150. https://doi.org/10.3233/ISU-230192

Muhammad Saif ud Din, S. (2025). Global Digital Transformation in Healthcare: Impact of Connected Healthcare Systems on Patient Safety. *Global Journal on Quality and Safety in Healthcare*, *8*(2), 60–62. https://doi.org/10.36401/JQSH-25-X1

Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, *59*(3), 1065–1078. https://doi.org/10.1016/j.compedu.2012.04.016

Reddy, P., Sharma, B., & Chaudhary, K. (2020). Digital Literacy: A Review of Literature. *International Journal of Technoethics (IJT)*, *11*(2), 65–94. https://doi.org/10.4018/IJT.20200701.oa1

Shiri, R., El-Metwally, A., Sallinen, M., Pöyry, M., Härmä, M., & Toppinen-Tanner, S. (2023). The Role of Continuing Professional Training or Development in Maintaining Current Employment: A Systematic Review. *Healthcare*, *11*(21), 2900. https://doi.org/10.3390/healthcare11212900

Vodă, A. I., Cautisanu, C., Grădinaru, C., Tănăsescu, C., & De Moraes, G. H. S. M. (2022). Exploring Digital Literacy Skills in Social Sciences and Humanities Students. *Sustainability*, *14*(5), 2483. https://doi.org/10.3390/su14052483

Zakon o zdravstvenem varstvu in zdravstvenem zavarovanju (ZZVZZ) (PISRS). (n.d.). Retrieved 18 July 2025, from https://pisrs.si/pregledPredpisa?id=ZAKO213

Zdravstveni sistem v Sloveniji. (n.d.). Retrieved 10 March 2026, from https://www.gov.si/assets/ministrstva/MZ/DOKUMENTI/NOVICE/Zdravstveni-sistem-v-Sloveniji-januar-2023.pdf

zVEM - eZdravje. (2020, March 18). https://ezdrav.si/resitve/zvem/

Inertial Motion Capture and Digital Human Modeling for Ergonomic Risk Assessment of Accordion Performance

Tilen Medved,[1] Branka Jarc Kovačič,[2]
Janez Kalan,[2] Zvone Balantič[1]

[1] University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
tilen.medved2@um.si, zvone.balantic@um.si
[2] Higher Vocational College, School Centre Kranj, Kranj, Slovenia
branka.jarc.kovacic@sckr.si, janez.kalan@sckr.si

Reliable ergonomic risk assessment requires methods that capture short, asymmetric, and repetitive postures. We present a wearable, model-based workflow to evaluate posture risk during accordion playing. One experienced right-handed player performed a 10-minute standing piece wearing inertial sensors. Motion data were mapped to a scaled digital human model, the accordion mass was added at the shoulders, and postures were scored using ergonomic methods. Joint-angle exposure and L4/L5 compression, shear, and moments were also measured. The neck and trunk remained mostly near neutral, while the upper limbs showed sustained non-neutral loading, including deep right-elbow flexion and continuous wrist flexion and deviation. Peak L4/L5 compression was around 1750 N, with low shear. Results support strap adjustments, technique coaching, and practice planning to reduce upper-limb risk. This case study shows how IMU data and digital human modeling can provide fast, repeatable risk scoring for performance health.

DOI
https://doi.org/10.18690/um.fov.4.2026.30

ISBN
978-961-299-152-4

Keywords:
inertial measurement units, digital human modeling, ergonomic risk assessment, musculoskeletal disorders, musician ergonomics, motion capture

University of Maribor Press

1 Introduction

Ergonomic risk assessment is a cornerstone of occupational health and safety, directly linked to preventing work-related injuries and improving health (Balantič et al., 2016). Traditional ergonomic evaluations often rely on checklists and observers' subjective judgment, which may miss transient or subtle risky postures. To address this, modern approaches increasingly combine the physical, biological, and technological aspects of work environments into a holistic assessment framework (Castillo et al., 2022; Carnazzo et al., 2024). These approaches also leverage advanced technologies (such as IoT and cloud computing) to improve processes and reduce worker strain. Smart ergonomics that use IoT sensors and cloud tools can stream motion data, spot risky patterns, and prompt early action (Sangeethalakshmi et al., 2023; Bances et al., 2024). Integrating these domains requires objective methods for data capture and analysis. The same need for objective, time-based risk monitoring applies to performance tasks, where exposure builds over long periods of practice and repeated movement.

Best practices for studying ergonomics in the workplace include a comprehensive approach that considers various aspects of the work environment and employees' needs (Balantič et al., 2024). Modern wearable motion sensors (inertial measurement units/IMUs) can record a person's movements in three-dimensional space and also add the fourth dimension – time – to the ergonomic analysis. These sensors capture continuous joint angles and body segment orientations throughout a task, and their data can be imported into digital human modeling software to recreate and evaluate the activity virtually. Using this approach, standardized ergonomic methods can be applied to motion data to objectively quantify risk levels. The high consistency and frequency of sensor measurements improve the reliability of risk assessments and allow much faster analysis than manual observation (Lamooki et al., 2022; Akhmad et al., 2020). Therefore, advanced analytical tools can provide ergonomics experts with immediate support in analyzing collected data to reduce risks in the workplaces under consideration (Greco et al., 2020; Kotowski & Gibson, 2023; Hilmi et al., 2024; Rybnikár et al., 2022; Yunus et al., 2021).

This paper presents a case study that applies inertial motion tracking and 4D analysis (full-body 3D movement data collected over time) to the task of playing a piano accordion. This activity involves complex upper-body movements and static load

from a heavy instrument. In the context of public health, it is important to note that musicians are among the most vulnerable professional groups to musculoskeletal disorders, with prevalence rates comparable to those of healthcare and industrial workers (Stanhope et al., 2022; Kok et al., 2018; Kok et al., 2013). While much ergonomic research focuses on industrial or office tasks, musicians can also experience work-related musculoskeletal disorders due to awkward postures and repetitive practice. Prior studies have found that young accordion players are prone to shoulder and wrist pain and to altered spinal postures (forward head position and increased lumbar lordosis) compared with non-musicians (Aparicio et al., 2016). These issues underscore the need for ergonomic evaluation in musical performance. High rates of musculoskeletal pain have also been reported among pianists and violinists, especially in the neck, shoulders, and wrists (Ozdemir et al., 2019; Lima et al., 2020). Kinect-based motion analysis has shown potential for evaluating musicians' postures objectively (Buisseret et al., 2018). However, systematic reviews (e.g., Rousseau et al., 2023) emphasize that existing studies are fragmented, typically focusing on one instrument group at a time, and that there is a lack of longitudinal and instrument-specific research. Importantly, no systematic ergonomic assessments have yet been conducted for the accordion, despite its unique challenges: instrument mass (10–15 kg), asymmetric loading from shoulder straps, and repetitive left-hand movements operating the bellows and bass keys.

The main contribution of this study is a practical workflow that combines inertial motion capture with digital human modeling for ergonomic evaluation in musicians. While musculoskeletal risk in instrumental performance has been studied previously, integration of IMU-based kinematics, biomechanical modeling, and ergonomic scoring remains limited. This case study demonstrates how motion data can be translated into joint loading estimates and time-based posture risk scores for musician health.

1.1 Research problem

Conventional ergonomic methods struggle to capture the dynamic, asymmetric, and repetitive nature of playing a musical instrument. Although systematic reviews (e.g., Rousseau et al., 2023) have highlighted common musculoskeletal risks among musicians in general, they also note that most studies are fragmented, instrument-specific, and rarely longitudinal. Importantly, no systematic ergonomic assessments

have been conducted for accordion playing, despite its unique load profile. The challenge is to improve risk assessment by combining inertial motion tracking with digital human models, moving beyond subjective observation and static postural snapshots. This dual focus addresses both a methodological gap (the limitations of conventional observational approaches) and a content-related gap (the lack of instrument-specific studies on accordion players). Therefore, this study addresses the following research question: How can inertial motion capture combined with digital human modeling identify time-dependent ergonomic risks during standing accordion performance, and which body regions exhibit the highest exposure? To answer this research question, the case study had four aims:

1. Measuring dynamic accordion playing postures.
2. Linking motion data to ergonomic criteria.
3. Identifying critical exposure regions.
4. Supporting preventive interventions.

2 Methods

This case study involved an experienced male accordion player performing a typical musical piece while wearing a full-body inertial motion capture suit (Figure 1). The participant was 28 years old, weighed 95 kg, was right-handed, and had over 15 years of playing experience. The participant (who is also a co-author) gave written consent for data collection and publication. A set of 17 wireless Xsens IMUs were placed on key body segments to capture the subject's movements in real time. The participant reported that the sensors did not restrict playing movements during the 10-minute task, although the suit and sensor straps may still affect comfort during longer recordings. The case study design (single participant) is a pilot-level investigation that provides detailed insights but limits generalizability.

The sensors were calibrated according to the manufacturer's guidelines (Xsens Technologies, 2021) to establish a baseline neutral pose before recording. During the experiment, the player stood and performed a 10-minute set of traditional pieces (waltz and polka) without breaks. At the same time, the system streamed motion data at 60 Hz, allowing the inertial system to log the dynamic joint angles and accelerations throughout the activity. The standing posture was deliberately selected,

as it places the full weight of the instrument on the shoulder girdle, thereby accentuating upper-limb loading compared with seated performance.

Figure 1: The accordion player wearing sensors before and during the experiment
Source: Own

Before simulation, detailed anthropometric data were collected from the participant, including body height, arm span, limb lengths, and segment circumferences. These values were used to scale the digital human model so that joint centers, limb proportions, and mass distribution matched the player's actual body. Accurate anthropometric input is essential, since it directly affects joint angle calculations and biomechanical load estimates, allowing the model to reflect the ergonomic stresses experienced by this individual. For clarity, only the most relevant anthropometric measures are reported in Table 1.

Table 1: Anthropometric measures of the accordion player

Body Dimension	Measurement
Body Height	188 cm
Shoulder Height	160 cm
Shoulder Width	47 cm
Elbow Span	95 cm
Arm Span	188 cm
Hip Height	110 cm
Knee Height	56 cm

Motion capture data from the Xsens system were processed in MVN Analyze and imported into Siemens Tecnomatix Process Simulate for ergonomic analysis (Siemens, 2024). The virtual avatar (Figure 2) was scaled to the subject's anthropometry and placed in a simulated standing posture. The 13-kg accordion was modeled as an external load attached to the torso via shoulder straps to match the real weight distribution. IMUs provided continuous kinematic data, while digital human modeling enabled standardized ergonomic assessment and biomechanical load estimation. This workflow allowed simultaneous evaluation of dynamic postures and internal loading. Ergonomic modules in Tecnomatix were then applied using OWAS (Ovako Working Posture Analysis System) and RULA (Rapid Upper Limb Assessment), along with tools for joint-angle exposure and lower-back load calculation. OWAS classifies whole-body postures into 4 action categories, where higher numbers indicate a greater need for corrective action. RULA scores upper-limb, neck, trunk, and wrist postures on a scale from 1 to 7, where higher scores indicate higher ergonomic risk. The analysis covered five outcome domains: OWAS categories, RULA scores (left/right), time in neutral/caution/high-risk postures (neck, trunk, shoulders, elbows, wrists), L4/L5 compressive and shear forces, and L4/L5 trunk moments.

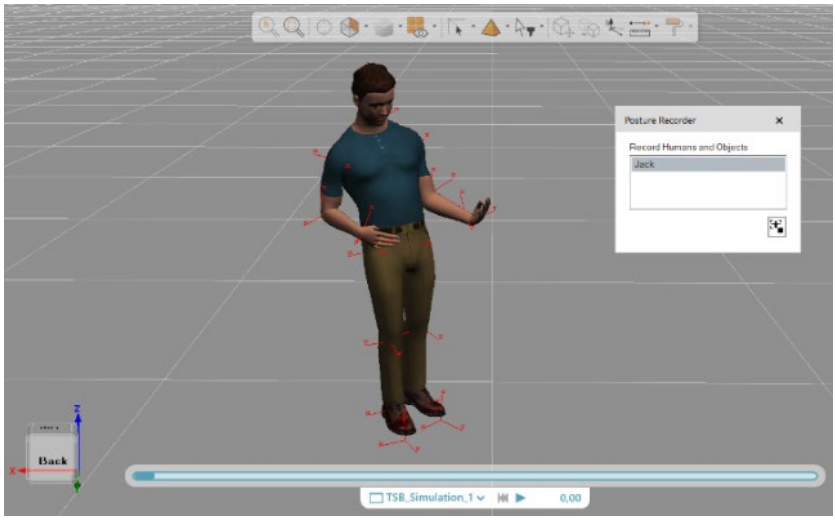

Figure 2: Accordion player's model inside the Tecnomatix Process Simulate environment
Source: Own

3 Results

Using an inertial-sensing and simulation approach, we obtained a detailed ergonomic risk profile for accordion playing. The analysis combined high-level scores (OWAS categories and RULA results) with biomechanical metrics, including joint-angle distributions and L4/L5 forces and moments. Posture classification identified moderate-risk postures but no sustained extreme or immediately dangerous exposures. The accordion's mass and the player's arm positions were the main contributors to elevated risk scores. The shoulders, elbows, and wrists emerged as the most exposed body regions, while the trunk remained largely neutral.

Figure 3 shows the OWAS results for the recorded accordion performance. It indicates the proportion of time spent in each OWAS action category during the task. Most of the players' postures fell into AC 2 (yellow). Category 2 indicates a posture that may cause harm to the musculoskeletal system; corrective actions are required in the near future. This moderate-risk classification is mainly due to the slightly bent trunk posture. No high-risk OWAS categories were observed, indicating the player avoided extreme bending or twisting while playing. However, the prevalence of continuous AC 2 suggests that posture variability was low, which aligns with the sustained asymmetric loading observed in the upper limbs.

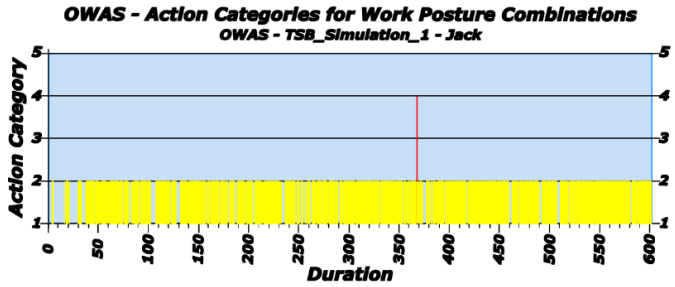

Figure 3: OWAS action categories during the 10-minute accordion performance
Source: Own

The RULA evaluation of the accordion-playing task yielded final scores of 6 on the left and 5 on the right, indicating a high risk that calls for intervention. This score was driven primarily by the arm and wrist positions required to play the instrument's keyboard and operate the bellows. In particular, the left wrist was held at extreme

angles, and the left shoulder was raised for prolonged periods, which contributed to the high RULA score. According to the RULA action levels, scores of 5 and 6 mean that further investigation is required and changes should be implemented soon to reduce the risk of injury. These findings underline that upper-limb postures are the main risk drivers, while the trunk remained largely neutral.

The next analysis evaluated ergonomic metrics as the percentage of time body segments were held within defined posture ranges and projected these exposures to a simulated 4-hour playing duration. A total of 26 segments were tracked and classified into three risk levels (green, yellow, red). Figure 4 shows the twelve segments with the highest exposure. Neck and trunk postures remained mostly neutral, confirming that trunk loading was limited. In contrast, the upper limbs showed sustained non-neutral positions. The right shoulder remained in extension for most of the task, while the left shoulder showed prolonged internal rotation. Elbow postures were asymmetric, with the right elbow frequently in high flexion. The most extreme exposure occurred in the wrists, which stayed in large flexion and deviation angles for most of the recording.

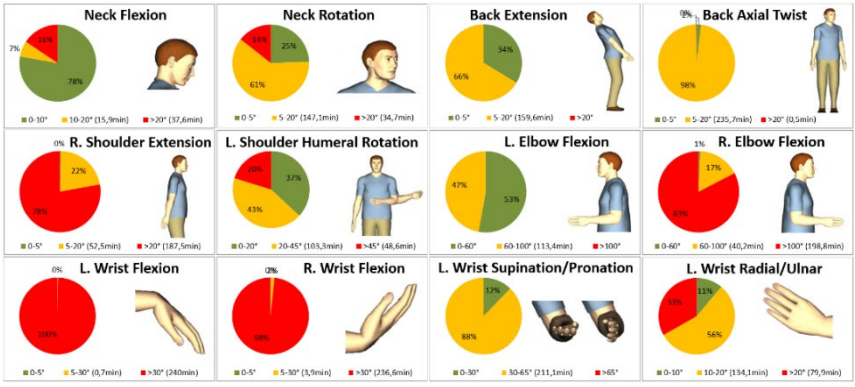

Figure 4: Time share (%) spent in neutral, caution, and high-risk joint-angle ranges for the most exposed body regions

Source: Own

We continued with lower back loading, as indicated by L4/L5 compressive and shear forces throughout the performance. The red curve in Figure 5 shows the compressive force on the lumbar spine, peaking at approximately 1750 N. This peak is below the NIOSH-recommended maximum of 3400 N for safe lumbar

compression (Albers et al., 2005), indicating that the vertical load from the accordion remained within acceptable limits. The blue and green curves show anterior-posterior and lateral shear forces, respectively. These shear forces were minimal, staying below 50 N at all times – far below the 700 N threshold often cited as the threshold for concerning shear loads (Afshari et al., 2018). Measurements of L4/L5 compressive and shear forces, therefore, remained within established safety thresholds, confirming that trunk loading was moderate compared to the persistent postural strain in the upper limbs.

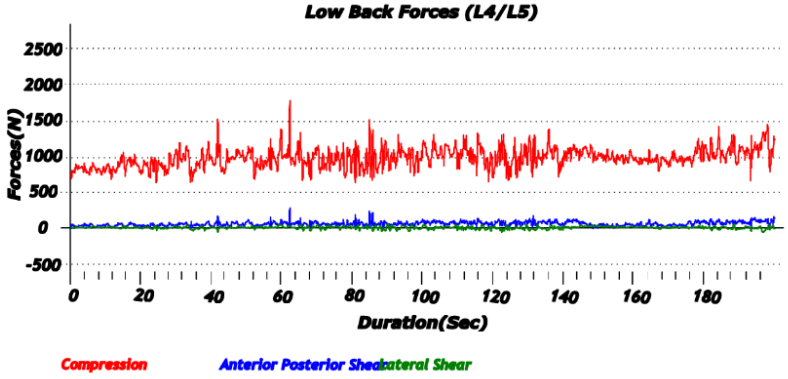

Figure 5: L4/L5 compressive and shear forces during accordion performance
Source: Own

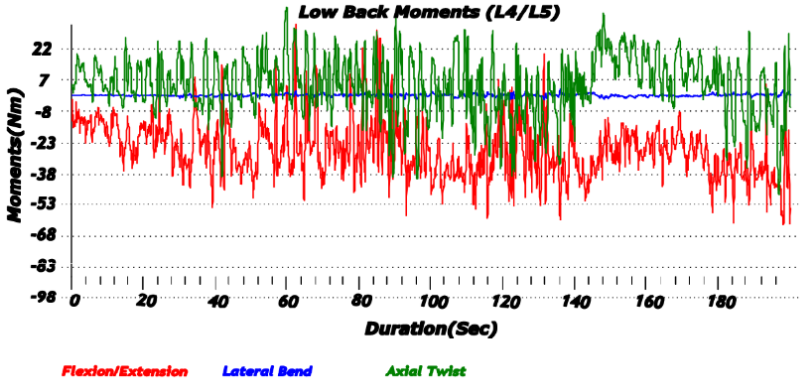

Figure 6: L4/L5 moments in flexion/extension, lateral bending, and axial twisting during accordion performance
Source: Own

Figure 6 shows L4/L5 moments in three planes: flexion/extension, lateral bending, and axial twisting. Flexion/extension was the dominant component and increased when the player adjusted under the accordion's weight. Lateral bending moments remained near zero, while axial twisting showed repeated peaks of about 30 Nm during torso rotation and bellows movement. Although moderate, these repeated twisting moments may contribute to lower-back fatigue over time.

4 Discussion

This case study evaluated ergonomic risk during standing accordion performance using inertial motion tracking and digital human modeling. Wearable sensors, software, and algorithms now support movement data collection and analysis across fields (Lind et al., 2020; Hilmi et al., 2024; Bances et al., 2024), offering objective insight for better ergonomic decisions and work processes (Lind, 2024). The main contribution is a practical workflow that combines inertial sensors with digital human modeling to produce repeatable posture scores and spinal load estimates for musician ergonomics.

The results show that inertial tracking with 4D simulation can reveal risks in complex, dynamic tasks that visual observation may miss. Sustained wrist flexion and deviation, high elbow flexion, shoulder loading, and repeated trunk moments were identified, matching reports of frequent upper-limb and shoulder complaints in musicians (Aparicio et al., 2016; Ozdemir et al., 2019; Lima et al., 2020). Time spent in moderate- and high-risk postures supports preventive actions, such as strap redesign, instrument support changes, technique coaching, and hip-strap systems to reduce shoulder and lower back load. Combining these data with wider ergonomic indices, including cognitive and environmental demands, may improve risk ranking and prevention (Rezvanizadeh et al., 2023). The same workflow may also support sports or military training, although OWAS and RULA should be used cautiously in high-intensity tasks.

The assessment showed that trunk and upper-limb demands differed. Most trunk-related postures fell into the OWAS action category 2, showing moderate risk and the need for future correction. In contrast, RULA scores of 5 on the right and 6 on the left indicate the need for early intervention. Thus, overall risk was moderate, but the wrists, elbows, and shoulders showed persistent high-risk exposure, while trunk

loading remained secondary. This agrees with reports of wrist and shoulder problems among musicians (Kok et al., 2016) and supports targeted changes such as strap adjustment, technique modification, and planned rest.

Ergonomic research on musicians remains fragmented, often focusing on violin, piano, and wind instruments, while other instruments receive less attention (Rousseau et al., 2023). This study helps address that gap by examining accordion performance. Unlike violin studies, where neck load is often dominant (Kok et al., 2016), and piano studies, where wrist asymmetry is common (Lima et al., 2020), accordion playing mainly stressed the shoulders, elbows, and wrists, with moderate trunk loading.

Earlier work on young accordion players found marked changes in head and lumbar posture (Aparicio et al., 2016), but trunk posture in this study remained mostly neutral. This may reflect the player's experience or posture control, supporting the value of training. The standing setup is also a limitation: the 13-kg instrument load acted through the shoulder straps, likely increasing upper-limb strain. Seated play could shift some of the load to the thighs and lower back, altering the risk pattern.

The L4/L5 analysis showed repeated flexion–extension and axial rotation moments, which require trunk muscle co-contraction and increase compression beyond instrument weight. Sustained compression with repeated asymmetric moments can contribute to disc loading and degeneration (Cui et al., 2022). Although compression values were moderate, dynamic trunk moments suggest cumulative spinal loading, which may help explain chronic low-back pain in musicians over long practice or performance periods (Hasegawa et al., 2018). In this case, however, upper-limb exposure remained the main risk.

Methodologically, combining IMUs with digital human modeling yielded a more objective and comprehensive assessment than visual observation alone. IMUs provide joint-angle data suitable for ergonomic tools (Rybnikár et al., 2022; Yunus et al., 2021), reducing subjectivity. The 4D analysis also captured time-dependent risk, including wrist flexion duration and peak spinal load timing, which snapshot-based methods may miss. The workflow also supports fast feedback and continuous improvement. Motion data were automatically recorded, and OWAS, RULA, ergonomic metrics, and L4/L5 loads were calculated quickly. This can help detect

unsafe patterns and guide changes to equipment or technique, as shown in earlier work on real-time or near real-time feedback (Santos et al., 2024; Lamooki et al., 2022; Akhmad et al., 2020).

For broader use, the workflow should serve as a decision-support tool rather than a diagnostic tool. Simple outputs, such as time spent in high-risk wrist postures, peak spinal load, and RULA/OWAS action levels, could guide teachers, performers, ergonomists, and health staff in choosing strap or technique changes, rest plans, or clinical referral. Broader use would require faster setup, repeatable sensor placement, secure storage of movement data, and clear consent rules, especially when data connect to health records or are shared in schools, clinics, or workplaces.

5 Conclusion

This study shows that combining inertial sensors with digital human modeling enables objective and repeatable ergonomic risk assessment during accordion performance and other complex tasks. The proposed workflow allows accurate tracking of movement, calculation of joint exposure, and estimation of spinal loading, providing reliable data for evaluating performance health. Such sensor-based methods support precise analysis of repetitive and asymmetric tasks and can improve ergonomics prevention strategies (Vox et al., 2021). With the ongoing development of wearable sensors, digital simulation, and machine-learning methods, these tools are expected to provide more detailed and personalized ergonomic evaluations in the future (Hilmi et al., 2024).

The present case study involved one experienced player and a single 10-minute standing performance, so the results cannot be generalized to all accordion players. The main contribution is the demonstration of a feasible IMU–digital human modeling workflow that can be applied to musician ergonomics and similar tasks involving sustained and asymmetric loading.

Future research should include larger samples, different playing conditions, and longer recordings. Combining motion tracking with physiological measures such as EMG, heart rate, and perceived fatigue would allow a more complete assessment of ergonomic exposure and support the development of evidence-based

recommendations for musicians and other professions with similar physical demands.

References

Afshari, D., Latifi, S. M., Kord, S., & Nourollahi-Darabad, M. (2018). Assessment of low back disorders risk based on allowable weight limits for manual lifting in Iran. *Industrial Health*, 56(4), 327–335. https://doi.org/10.2486/indhealth.2017-0174

Akhmad, S., Arendra, A., Findiastuti, W., Lumintu, I., & Pramudita, Y. D. (2020). Wearable IMU wireless sensors network for smart instrument of ergonomic risk assessment. In *2020 6th Information Technology International Seminar (ITIS)* (pp. 213–218). IEEE. https://doi.org/10.1109/ITIS50118.2020.9321084

Albers, J. T., Hudock, S. D., & Kong, Y. K. (2005). HETA# 2003-0146-2976: Genesis Steel Services, Inc.; Baltimore, Maryland. https://doi.org/10.26616/NIOSHHETA200301462976

Aparicio, L., Lã, F. M., & Silva, A. G. (2016). Pain and posture of children and adolescents who learn the accordion as compared with non-musician students. *Medical Problems of Performing Artists*, 31(4), 187–192. https://doi.org/10.21091/mppa.2016.4034

Balantič, Z., Jarc Kovačič, B., & Medved, T. (2024). Model obladovanja tveganj za utrjevanje vrednot ergonomskih načel v proaktivni ergonomiji [Risk management model for strengthening the values of ergonomic principles in proactive ergonomics]. In B. Urh, M. Maletič, A. Kohont, & M. Bajić (Eds.), *Raziskovalni trendi in trajnostne rešitve v inženiringu poslovnih sistemov* [Research trends and sustainable solutions in business systems engineering] (pp. 85–109). Univerza v Mariboru, Univerzitetna založba.

Balantič, Z., Polajnar, A., & Jevšnik, S. (2016). *Ergonomija v teoriji in praksi*. Nacionalni inštitut za javno zdravje.

Bances, E., Schneider, U., Siegert, J., & Bauernhansl, T. (2024). Enhancing ergonomics in construction industry environments: A digital solution with scalable event-driven architecture. *Human Factors and Systems Interaction*, 154, 103-113. https://doi.org/10.54941/ahfe1005358

Buisseret, F., Dierick, F., Hamzaoui, O., & Jojczyk, L. (2018). Ergonomic risk assessment of developing musculoskeletal disorders in workers with the Microsoft Kinect: TRACK TMS. *IRBM*, 39(6), 436–439. https://doi.org/10.1016/j.irbm.2018.10.003

Carnazzo, C., Spada, S., Lamacchia, S., Manuri, F., Sanna, A., & Cavatorta, M. P. (2024). Virtual reality in ergonomics by wearable devices: Experiences from the automotive sector. *Journal of Workplace Learning*, 36(7), 621–635. https://doi.org/10.1108/JWL-03-2024-0064

Castillo, J. D., Marquez, B. Y., Realyvazquez, A., & Alanis, A. (2022, September). RULA (Rapid Upper Limb Assessment) Methodology for Ergonomics Using Inertial Sensors. In *International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability* (pp. 177-190). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-20316-9_14

Cui, Y., Shen, H., Chen, Y., Zhang, W., Zhu, J., Duan, Z., Liao, Z., & Weiqiang, L. (2022). Study on the process of intervertebral disc disease by the theory of continuum damage mechanics. *Clinical Biomechanics*, 98, 105738. https://doi.org/10.1016/j.clinbiomech.2022.105738

Greco, A., Caterino, M., Fera, M., & Gerbino, S. (2020). Digital twin for monitoring ergonomics during manufacturing production. *Applied Sciences*, 10(21), 7758. https://doi.org/10.3390/app10217758

Hasegawa, T., Katsuhira, J., Oka, H., Fujii, T., & Matsudaira, K. (2018). Association of low back load with low back pain during static standing. *PLOS ONE*, 13(12), e0208877. https://doi.org/10.1371/journal.pone.0208877

Hilmi, A. H., Hamid, A. R. A., & Ibrahim, W. A. R. A. W. (2024). Recent advancements in ergonomic risk assessment: integration of artificial intelligence, wearable technology, and industry-specific approaches. *Malaysian Journal of Ergonomics* (MJEr), 6, 65-75. https://doi.org/10.58915/mjer.v6.2024.1309

Kok, L. M., Vliet Vlieland, T. P. M., Fiocco, M., & Nelissen, R. G. H. H. (2013). A comparative study on the prevalence of musculoskeletal complaints among musicians and non-musicians. *BMC Musculoskeletal Disorders*, 14, Article 9. https://doi.org/10.1186/1471-2474-14-9

Kok, L. M., Huisstede, B. M., Voorn, V. M., Schoones, J. W., & Nelissen, R. G. (2016). The occurrence of musculoskeletal complaints among professional musicians: A systematic review. *International Archives of Occupational and Environmental Health*, 89(3), 373–396. https://doi.org/10.1007/s00420-015-1090-6

Kok, L. M., Groenewegen, K. A., Huisstede, B. M., Nelissen, R. G., Rietveld, A. B. M., & Haitjema, S. (2018). The high prevalence of playing-related musculoskeletal disorders (PRMDs) and its associated factors in amateur musicians playing in student orchestras: A cross-sectional study. *PLOS ONE*, 13(2), e0191772. https://doi.org/10.1371/journal.pone.0191772

Kotowski, S. E., & Gibson, S. L. (2023*). Ergonomic assessment toolkit*. AIHA.

Lamooki, S. R., Hajifar, S., Kang, J., Sun, H., Megahed, F. M., & Cavuoto, L. A. (2022). A data analytic end-to-end framework for the automated quantification of ergonomic risk factors across multiple tasks using a single wearable sensor. *Applied Ergonomics*, 102, 103732. https://doi.org/10.1016/j.apergo.2022.103732

Lima, C., Roriz, A., Leite, A., Colim, A., & Carneiro, P. (2020). Exposure to musculoskeletal risk of piano teachers. In *Occupational and environmental safety and health ii* (pp. 419-426). Cham: Springer International Publishing.

Lind, C. M. (2024). A rapid review on the effectiveness and use of wearable biofeedback motion capture systems in ergonomics to mitigate adverse postures and movements of the upper body. *Sensors*, 24(11), 3345. https://doi.org/10.3390/s24113345

Lind, C. M., Diaz-Olivares, J. A., Lindecrantz, K., & Eklund, J. (2020). A wearable sensor system for physical ergonomics interventions using haptic feedback. *Sensors*, 20(21), 6010. https://doi.org/10.3390/s20216010

Ozdemir, F., Tutus, N., Akgun, S. O., & Kilcik, M. H. (2019). Evaluation of work-related musculoskeletal disorders and ergonomic risk levels among instrumentalist musicians. *Annals of Medical Research*, 26(11), 2630–2634. https://doi.org/10.5455/annalsmedres.2019.08.454

Rezvanizadeh, M., Mohammad-Ghasemi, M., Soltanzadeh, A., & Sadeghi-Yarandi, M. (2023). Development of a novel ergonomic index assessment in the workplace based on physical, cognitive, and environmental components. *Work*, 75(3), 1071–1086.

Rousseau, C., Taha, L., Barton, G., Garden, P., & Baltzopoulos, V. (2023). Assessing posture while playing in musicians: A systematic review. *Applied Ergonomics*, 106, 103883. https://doi.org/10.1016/j.apergo.2022.103883

Rybnikár, F., Kačerová, I., Hořejší, P., & Šimon, M. (2022). Ergonomics evaluation using motion capture technology—Literature review. *Applied Sciences*, 13(1), 162. https://doi.org/10.3390/app13010162

Sangeethalakshmi, K., Lakshmi, V. V., Malathi, N., & Velmurugan, S. (2023). Smart ergonomic practices with IoT and cloud computing for injury prevention and human motion analysis. In *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)* (Vol. 1, pp. 1–6). IEEE. https://doi.org/10.1109/ICAIIHI57871.2023.10489748

Santos, C., Gabriel, A. T., Quaresma, C., & Nunes, I. L. (2024). The use of wearable sensors for ergonomic risk assessment of surgical procedures: A literature review. *Human Factors and Systems Interaction*, 154, 36. https://doi.org/10.54941/ahfe1005376

Siemens. (2024). Tecnomatix eMS documentation. Siemens Documentation Center. Retrieved January 11, 2026, from https://docs.sw.siemens.com/en-US/doc/288782031/PL20231017544579710.tecnomatix_eMS.xid1015765/xid1657141

Stanhope, J., Cook, A., Pisaniello, D., & Weinstein, P. (2022). Musculoskeletal symptoms in professional musicians: Do self-employed and employer-employed musicians differ? *Annals of Work Exposures and Health*, 66(8), 1056–1069. https://doi.org/10.1093/annweh/wxac046

Vox, J. P., Weber, A., Wolf, K. I., Izdebski, K., Schüler, T., König, P., Wallhoff, F., & Friemert, D. (2021). An evaluation of motion trackers with virtual reality sensor technology in comparison to a marker-based motion capture system based on joint angles for ergonomic risk assessment. *Sensors*, 21(9), 3145. https://doi.org/10.3390/s21093145

Xsens Technologies. (2021). MVN user manual. Retrieved September 2, 2025, from https://www.xsens.com/hubfs/Downloads/usermanual/MVN_User_Manual.pdf

Yunus, M. N. H., Jaafar, M., Mohamed, A., Azraai, N., & Hossain, M. (2021). Implementation of kinetic and kinematic variables in ergonomic risk assessment using motion capture simulation: A review. *International Journal of Environmental Research and Public Health*, 18(16), 8342. https://doi.org/10.3390/ijerph18168342

Bridging the Theory-to-Design Gap in Digital Wellness: A Design Science Approach to Motivational and Compassionate System Design

Ayesha Thanthrige, Nilmini Wickramasinghe
La Trobe University, School of Computing, Engineering and Mathematical Sciences, Victoria, Australia
thanthrige87@gmail.com, n.wickramasinghe@latrobe.edu.au

Digital wellness applications hold promise for lifestyle medicine, yet sustained engagement remains challenging because motivational and compassionate design principles are rarely operationalised as system mechanisms. This theory-to-design gap represents the core problem addressed in this study. Adopting Design Science Research Methodology, we develop MiCARE, a progressive web application supporting nutrition and physical activity engagement among young adults in preventive wellness contexts. Drawing on Self-Determination Theory and the CARE framework, we derive three design requirements, autonomy-supportive personalisation, competence-oriented coaching and feedback, and empathic interaction support and instantiate them as operational features. MiCARE was iteratively co-designed with young adults and clinicians. A formative usability evaluation (n=8) demonstrates strong usability (SUS=82.5), high task completion (97.5%), and high perceived empathy (92.5/100), providing preliminary evidence of the enactment of the design mechanisms. The study contributes a transparent DR–DP–DF translation approach and an empirically informed prototype demonstrating initial feasibility within a preventive, non-clinical wellness context.

DOI
https://doi.org/10.18690/um.fov.4.2026.31

ISBN
978-961-299-152-4

Keywords:
digital wellness, lifestyle medicine, design science research methodology, digital health engagement, empathic interaction

University of Maribor Press

1 Introduction

As defined by the American College of Lifestyle Medicine, lifestyle medicine prevents and manages chronic disease through six pillars: nutrition, physical activity, sleep, stress management, social connection, and substance avoidance (ACLM, 2026; Lippman et al. 2024). Nutrition and physical activity are central to prevention, for example, demonstrating reductions in type 2 diabetes incidence among individuals with prediabetes (Demissie et al. 2025). Unlike pharmacological approaches that primarily manage symptoms, lifestyle-oriented approaches emphasise supporting awareness of healthier lifestyles and engagement with wellness practices (Bodai et al. 2018).

Digital wellness technologies offer a promising channel for delivering such interventions at scale (Hazra-Ganju et al. 2023). However, sustained engagement remains a persistent challenge (Eiselt et al. 2025). Many applications successfully support initial uptake but struggle to retain users over time, limiting their long-term preventive value (Szinay et al. 2020). Evidence suggests this is not merely a technical failure but a design failure (Tzimourta 2025). Many systems privilege monitoring and tracking while offering insufficient motivational support, limited empathetic interactions, and limited user assistance (Dostert and Müller 2021).

Despite frequent citation of motivational and behavioural theories in digital health research, there is a persistent failure to operationalise them within artefact design (Heinsch et al. 2021; Zhang et al. 2021). Digital wellness systems rarely integrate both motivational support and compassionate forms of interaction in ways that are visible in system functionality (van Lotringen et al. 2023). This constitutes a theory-to-design gap: a critical middle layer between conceptual frameworks and their translation into concrete, testable system mechanisms. We therefore ask: How can motivational and compassionate care principles be translated into operational system mechanisms within digital wellness applications? We address this question through Design Science Research Methodology (DSRM) and develop MiCARE, a progressive web application prototype supporting nutrition- and physical activity-related wellness engagement among young adults in preventive contexts. Drawing on Self-Determination Theory (SDT) (Ryan and Deci 2000) and the CARE framework (Compassion, Assistance, Respect, and Empathy) (Sinclair et al. 2017, 2018), we derive design requirements (DR), translate them into design principles

(DP) and design features (DF), and instantiate and evaluate them in a functioning prototype. This study makes two primary contributions. First, it advances design science knowledge by providing a transparent and reproducible DR–DP–DF translation approach that links motivational and compassionate theory to concrete system mechanisms. Second, it demonstrates the feasibility of this approach through the MiCARE prototype and formative usability evaluation, providing preliminary evidence of usability and perceived effectiveness.

2 Related Work and Theoretical Foundations

2.1 Digital Wellness Engagement as a Design Challenge

The effectiveness of digital wellness interventions depends not only on access or functionality, but on users' willingness to remain engaged over time (Thanthrige and Wickramasinghe 2025). Existing evidence shows that disengagement remains a major challenge in app-based health interventions, particularly when systems rely on self-monitoring and routine data entry without adequately supporting users' motivation, context, or lived experience (Bailey-Davis et al. 2021; Lim et al. 2022).

Prior studies indicate that declining engagement is shaped by repetitive interaction patterns, low personal relevance, insufficient personalisation, and limited emotional support during setbacks or lapses (LeSeure et al. 2024; Oh et al. 2021). For example, culturally tailored features such as Portuguese-specific carbohydrate tracking and simplified no-login interfaces may improve relevance and usability, yet apps can remain limited by incomplete feedback, restricted language support, or weak real-time interaction (LeSeure et al. 2024; Xue et al. 2025). Similarly, chatbot-based interventions have shown promise when they incorporate personalised goal setting, relational strategies, or conversational support, but engagement reporting and sustained use remain inconsistent across studies (Nadarzynski et al. 2024; Oh et al. 2021; Wu et al. 2024; Zhang et al. 2020). Prospective evidence further confirms that meaningful feature engagement not mere app access is associated with better health outcomes (Lim et al. 2022), underscoring that the design of system mechanisms, not just their presence, is what matters. From an IS perspective, this is a design challenge. The key question is how to embed motivational, supportive, and context-sensitive mechanisms into system functionality rather than treating them as abstract aspirations. Within IS, design science research has been increasingly applied to

address such challenges. It does so by developing theory-grounded artefacts that embed motivational and relational mechanisms into system functionality to support sustained user engagement (Hevner et al. 2004; Baskerville et al. 2018).

2.2 Theoretical Foundations: SDT and CARE Framework

This study draws on two complementary theoretical sources selected because they directly address the two recurrent design deficiencies identified above: insufficient motivational support and limited relational sensitivity. SDT explains sustained motivation through the fulfilment of three basic psychological needs: autonomy, competence, and relatedness (Ryan and Deci 2000). Autonomy concerns a sense of volition and choice; competence reflects feelings of effectiveness and progress; and relatedness refers to feeling supported and understood (Martela and Riekki 2018). These constructs are highly relevant to digital wellness contexts, where users must often sustain effort without direct human supervision (Amirabdolahian et al. 2025). While SDT identifies autonomy, competence, and relatedness as distinct needs, this study operationalises relatedness through the CARE framework rather than as a standalone design requirement. This decision reflects a deliberate theoretical choice: in a non-clinical digital wellness context, relatedness is primarily experienced through the quality of interaction rather than as a structural system property. The CARE framework, comprising compassion, assistance, respect, and empathy, highlights the communicative behaviours through which a system can fulfil the relatedness need, making it the more actionable operationalisation for artefact design. CARE captures how relatedness is experienced in interaction through empathy, compassion, and respect thereby translating relational constructs into actionable design mechanisms.

Complementing SDT, the CARE framework comprising compassion, assistance, respect, and empathy highlights the importance of compassionate, respectful, and responsive interaction in healthcare settings (Sinclair et al. 2017; Sinclair et al. 2018). The CARE framework is particularly relevant for digital health systems because it operationalises compassionate communication behaviours that can be embedded in conversational interfaces. SDT and CARE are theoretically complementary rather than competing: SDT explains the motivational conditions required for sustained engagement, while CARE specifies the relational quality through which those conditions can be communicated in system interactions. Within IS research, empathic and trust-building interaction has been identified as a critical antecedent

of user acceptance and continued engagement in conversational agent systems (Ayers et al. 2023; Oh et al. 2021), positioning CARE as directly applicable to IS artefact design. Together, they provide a comprehensive basis for designing wellness systems that support user experience.

2.3 Translating Theory into Digital Wellness Artefacts

Although motivational and CARE concepts are widely recognised in health and behavioural research, their translation into digital wellness artefacts remains limited (Anser et al. 2025). Existing applications often reference behaviour change concepts at a high level but provide little transparency regarding how such concepts influence functional design decisions (Del-Valle-Soto et al. 2024). This limitation reflects a persistent theory-to-design gap, a critical middle layer problem in digital health systems (Persson and Rydenfält 2021). For design science research, this gap is especially important because the value of an artefact lies not only in whether it works, but also in how its design can be justified and communicated (Wiegmann et al. 2023). A structured approach is therefore required to connect theoretical constructs with DR, DP, and DF. In this study, SDT and CARE are treated as design knowledge sources and translated into a systematic DR–DP–DF structure guiding the development of MiCARE. Figure 1 illustrates the theory-driven design logic used in this study, linking lifestyle medicine pillars with the theoretical foundations, DRs, and implemented system features that lead to expected user outcomes.

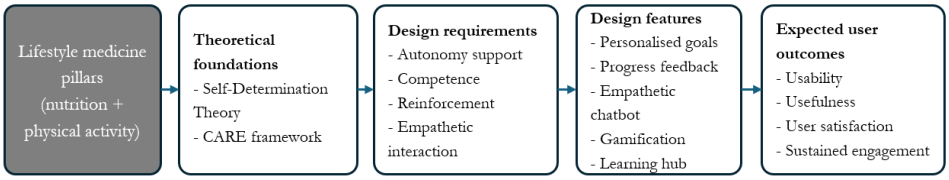

Figure 1: Theory-Driven Framework for Motivational Digital Wellness Design
Source: Own

In addition, prior research highlights that sustained engagement is strongly influenced by perceived outcomes (e.g., improved energy, wellbeing, or lifestyle confidence) and by the need for continued novelty and value in interactions. While outcome measurement is beyond the scope of the present formative study, the

proposed framework is intentionally extensible to incorporate outcome-visibility and adaptive engagement mechanisms in future iterations.

Importantly, many digital health interventions are vulnerable to the "eHealth law of attrition" (Eysenbach, 2005), in which a large proportion of users discontinue engagement over time despite promising initial uptake. While a pleasant interface and positive short-term user experience help mitigate early dropout, sustained use requires ongoing perceived value and newness in each interaction elements that go beyond the three core requirements addressed in this study.

3 Methodology

This study adopts DSRM to address the research question (Hevner et al. 2004; Peffers et al. 2008). DSRM is appropriate because the study seeks not merely to understand an engagement problem but to design, develop, and evaluate a novel artefact that addresses it an aim that aligns directly with DSRM's dual goals of problem-solving utility and design knowledge contribution (Baskerville et al. 2018). The research process followed five established DSRM stages: problem identification, definition of solution objectives, artefact design and development, evaluation, and communication of research contributions, illustrated in Figure 2. Ethics approval was obtained from the La Trobe University Human Research Ethics Committee (Project No. HEC24507).

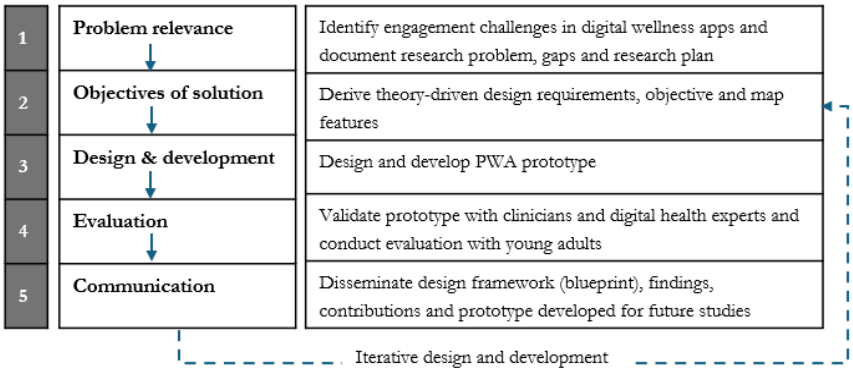

Figure 2: DSRM Research Process for MiCARE

Source: Own

In addition to the DSRM logic, the study adopted co-design sessions during requirements exploration and prototype refinement, ensuring theoretical insights were grounded in empirically collected stakeholder perspectives rather than translated into design decisions in isolation (Vargas et al. 2025). Two stakeholder groups informed the design across iterative rounds. Six domain experts consisting of a senior clinician, a digital health academic, a computer science academic, and three research fellows provided structured feedback on educational content accuracy, chatbot response tone, cultural alignment, gamification strategies, and the boundary between supportive wellness guidance and inappropriate clinical advice. Twenty young adults aged 18–34 participated in co-design sessions, contributing UI preferences, navigation patterns, interaction style expectations, and motivational support needs through structured surveys assessing clarity, cognitive load, personalisation, and perceived usefulness. Survey responses and expert feedback were thematically coded against the SDT constructs of autonomy, competence, and relatedness, as well as the CARE dimensions, to identify convergent design priorities. Codes were then grouped into emergent design themes, which directly informed the derivation of the three DRs. Together, these 26 stakeholders grounded the DR–DP–DF structure in both clinical rigour and user-centred evidence prior to prototype development. Also, no clinical or sensitive health data were collected, in line with ethics approval (HEC24507).

4 Objectives of Solution and DRs (DSRM Phase II)

Having established the theory-to-design gap as the core problem, DSRM Phase II defines the solution objective and derives DRs. The objective is to develop a digital wellness artefact that operationalises motivational and compassionate principles to support engagement with wellness information and lifestyle-related digital content.

SDT and the CARE framework inform three DRs. SDT posits that sustained engagement depends on meeting the psychological needs of autonomy, competence, and relatedness (Ryan and Deci 2000); CARE posits that the manner of interaction compassionate, respectful, empathetic shapes how those needs are perceived as met (Sinclair et al. 2018). Applying these jointly in a digital system requires embedding both motivational scaffolding and relational communication tone into artefact mechanisms. Three DRs follow: autonomy-supportive personalisation (DR1),

competence-oriented coaching and feedback (DR2), and empathic interaction support (DR3).

Autonomy-supportive personalisation (DR1): This requirement addresses SDT's emphasis on autonomy. Digital wellness systems should support user-defined goals rather than imposing prescriptive recommendations (Babu and Joseph 2025). Evidence indicates that digital wellness programs allowing user choice can strengthen participation and perceived engagement, particularly in diabetes and prediabetes management (Bailey-Davis et al., 2021; LeSeure et al., 2024; Lim et al., 2022). Consequently, the corresponding DP prioritises personalised goal setting and culturally tailored options, which have been shown to enhance engagement in digital health tools (LeSeure et al., 2024; Oh et al., 2021). MiCARE operationalises this through customisable nutrition and physical activity goals with adjustable targets. By providing flexibility and user choice, the system aims to reinforce intrinsic motivation and reduce perceived external pressure (Li et al. 2025b).

Competence-oriented coaching and feedback (DR2): Research shows that competence is strengthened not only through feedback but through guided capability development helping users understand what to do, why it matters, and how to improve over time. While progress indicators and gamification support motivation, they are insufficient alone. Therefore, the corresponding DP includes structured guidance, micro-coaching, and actionable behavioural support, in addition to visual feedback. MiCARE operationalises this through progress dashboards, gamified elements, and embedded educational guidance delivered via the learning hub and chatbot, focused on general wellness information rather than clinical advice.

Empathic interaction support (DR3): This requirement focuses on fostering empathetic, relational communication. Empathy is operationalised through supportive tone, non-judgmental language, and context-sensitive responses, informed by the CARE framework. Importantly, clinician review contributes to content reliability and safety rather than empathy itself. MiCARE achieves this through a bounded, clinician-reviewed conversational chatbot providing general wellness guidance and supportive interaction, without offering clinical diagnosis or treatment, ensuring alignment with ethical constraints.

Table 1: DRs Derived from Theoretical Foundations

Design Requirement	Theoretical Basis	Design Principle	System Feature
Autonomy support (DR1)	SDT	Enable users to define personal wellness goals (DP1)	Personalised goal setting (DF1)
Competence-oriented coaching and feedback (DR2)	SDT	Provide feedback and structured guidance (DP2)	Progress dashboard, activity tracking, gamification elements and embedded educational guidance (DF2)
Empathic interaction support (DR3)	CARE Framework	Support empathic communication (DP3)	Clinician-reviewed conversational chatbot and evidence-based learning hub (DF3)

In addition to these functional requirements, the artefact was developed with consideration of several non-functional aspects, including usability, accessibility, privacy, and practical deployment within digital wellness contexts. Table 2 outlines the key functional features implemented in the MiCARE prototype and their intended outcomes.

Table 2: Key Functional Features of the MiCARE System

Feature	Description	Intended Outcome
Personalised goal setting (DF1)	Allows users to define personalised nutrition and physical activity goals aligned with their preferences and lifestyle context.	Support user autonomy
Progress dashboard, activity tracking, gamification elements (DF2)	Includes a progress dashboard, activity tracking, and gamified elements that provide visual feedback on progress within the application interface and milestone achievements.	Reinforce competence through feedback and structured guidance
Clinician-reviewed conversational chatbot and evidence-based learning hub (DF3)	Clinically verified conversational chatbot and evidence-based learning hub (e.g. based on Australian Dietary Guidelines, Australia's Physical Activity and Sedentary Behaviour Guidelines, National Health and Medical Research Council (NHMRC), Guidelines and Diabetes Australia) providing supportive messages, guidance, and evidence-based wellness information.	Provide empathic support and encourage continued engagement

Together, the three DRs, DPs, and DFs form a structured DR–DP–DF framework that clearly links theory to system design the core design knowledge contribution of this study.

5 Design and Development of the Prototype (DSRM Phase III)

Following derivation of DRs, DSRM Phase III involved design and development of the MiCARE prototype as a browser-based progressive web application. The architecture comprises four primary layers presentation, application, logic, and data storage structured to support modular development while ensuring each layer directly instantiates one or more of the three DRs.The presentation layer provides a user interface designed for clarity, accessibility, and ease of use, informed by Web Content Accessibility Guidelines (WCAG) guidelines. Visual dashboards present activity progress and evidence-based educational content in an intuitive format. The application layer contains the primary interaction modules operationalising the DRs: goal personalisation (DF1), progress tracking and gamification (DF2), conversational support (DF3), configurable reminders, and educational content delivery. The reminder function was designed to be user-configurable rather than prescriptive, reinforcing autonomy rather than compliance (DR1). The logic layer processes user activity data, manages dynamic system responses, and bounds the chatbot to preventive wellness education. Clinician-reviewed response intents and escalation wording were applied to ensure responsible design preventing inappropriate clinical guidance. Empathic interaction was technically operationalised through clinician-reviewed Dialogflow ES intents with a 0.3 confidence threshold triggering fallback responses, and escalation wording designed to prevent clinical overreach while maintaining a supportive conversational tone. The data layer stores user profiles, activity records, and interaction logs supporting both system functionality and evaluation. The novelty of MiCARE lies not in any individual feature but in the systematic, coherent integration of motivational, empathic, and responsible design mechanisms across the full system architecture demonstrating how the theory-to-design gap can be bridged in practice.

6 Evaluation of Prototype (DSRM Phase IV)

In line with DSRM, we conducted a formative usability evaluation to assess whether the three design mechanisms autonomy support, competence-oriented coaching and feedback, and empathic interaction were usably enacted in the MiCARE prototype (Hevner et al. 2004). Formative evaluation at this stage is consistent with DSRM, which emphasises iterative assessment of artefact utility prior to large-scale deployment (Peffers et al. 2008). The evaluation assessed three dimensions standard

in interactive health system assessment: effectiveness (task completion and error rates), efficiency (task duration, clicks, and chatbot utterances), and satisfaction (System Usability Scale (SUS) score and perceived empathy) (Nielsen and Landauer, 1993). Eight young adults aged 18–34 years (mean age 26.9, SD=4.2; 50% female) were recruited using purposive sampling to ensure diversity across cultural background, gender, and geographic location within Victoria, Australia. Purposive sampling is appropriate for formative usability evaluation where the aim is to identify interface issues rather than achieve statistical representativeness (Nielsen and Landauer, 1993). Eligible participants owned a smartphone or computer and expressed interest in improving diet or physical activity habits. Individuals with diagnosed diabetes or requiring clinical dietary advice were excluded to maintain alignment with the preventive wellness scope of the system and approved ethical guidelines. Eight participants are consistent with established usability research demonstrating that this sample size identifies the majority of interface issues in early-stage evaluation (Nielsen and Landauer, 1993).

Testing was conducted remotely via one-on-one video conferencing sessions to ensure accessibility across metropolitan and regional Victoria. Each session followed a four-component structured protocol: (1) a background questionnaire capturing demographics and prior wellness app experience; (2) ten representative task-based interactions, whereby participants provide thoughts and expectations (3) the SUS questionnaire; and (4) a brief semi-structured interview. Sessions lasted 60–75 minutes and were audio-recorded with participant consent. The lead researcher facilitated all sessions using a standardised guide to ensure consistency. Chatbot non-recognition events were defined as user utterances not matched to any predefined conversational intent, triggering a fallback response. The SUS is a validated 10-item instrument widely used in digital health usability research, with scores above 80 classified as excellent (Bangor et al. 2008). Perceived empathy was assessed using a 0–100 scale, consistent with prior conversational agent research (Ayers et al. 2023).

Quantitative data, task completion rates, error counts, SUS scores, chatbot utterance counts, and empathy ratings were summarised using descriptive statistics (means, standard deviations, counts, and percentages). Qualitative data from task based interactions and semi-structured interviews were analysed using thematic analysis (Braun and Clarke 2006), in which codes were iteratively developed and organised into themes and subthemes through repeated engagement with the data.

Triangulation across quantitative performance metrics and qualitative findings supported a comprehensive and internally consistent interpretation of user experience (Creswell and Plano Clark 2018). Findings are reported according to the three design mechanisms to maintain direct traceability to the DR–DP–DF framework.

Autonomy support (DR1/DF1) was positively reflected in qualitative feedback and interaction patterns. Participants valued goal configurability and described the interaction as flexible due to quick-access options and self-paced navigation. Two minor navigation slips occurred when switching between modules, suggesting the need for clearer first-use guidance and a persistent navigation menu identified as a targeted refinement for the next iteration. Competence reinforcement (DR2/DF2) was evidenced by efficient task performance and usability ratings. Participants successfully completed 78 of 80 tasks (97.5%) without facilitator intervention. MiCARE achieved a mean SUS score of 82.5 (SD=11.3), classified as excellent usability. Participants highlighted the visibility of progress indicators and interface simplicity as supporting confidence and competence during task completion. Empathic interaction (DR3/DF3) was strongly endorsed. Across 105 chatbot utterances, 10 (9.5%) triggered recoverable fallback prompts, consistent with the bounded, rule-based design of the conversational system. Perceived empathy was high (mean 92.5/100), with participants frequently describing the chatbot as 'supportive' and 'non-judgmental'. Participants suggested that fallback prompts could provide clearer contextual guidance when queries were ambiguous with targeted improvement for subsequent iterations. Overall, the evaluation indicates that all three design mechanisms were positively perceived: the DR–DP–DF instantiation showed promising results across effectiveness, efficiency, and satisfaction dimensions. These results provide preliminary, formative evidence that the DR–DP–DF mechanisms are usable and meaningfully perceived by users, but do not establish long-term engagement, behavioural change, or clinical outcomes.

7 Discussion

This study examined how motivational and compassionate principles can be operationalised in digital wellness systems. The central contribution is a structured DR–DP–DF framework that bridges the theory-to-design gap identified in prior digital health research: three theoretical constructs (SDT autonomy, SDT

competence, and CARE empathy as operationalisation of SDT relatedness) were systematically translated into three DRs, three DPs, and three implemented features, each evaluated formatively. Findings confirm that the mechanisms were positively perceived across all three dimensions. The evaluation results indicate that the instantiated design mechanisms derived through the DR–DP–DF structure were usably enacted in the prototype, suggesting that the framework provides a workable mechanism for translating theoretical constructs into operational system design.

The contributions of this study are twofold. Theoretical contribution: Although motivational and behavioural theories are frequently referenced in digital health research, prior studies rarely demonstrate how such theories are translated into concrete artefact design (Heinsch et al. 2021). By operationalising SDT and the CARE framework through a structured DR–DP–DF model, this study provides a transparent mechanism for translating theoretical constructs into implementable system functionality. The DR–DP–DF structure itself comprises design knowledge that extends the IS design science literature: it offers an explicit, auditable middle layer that prior digital health artefacts lack. The integration of SDT and CARE is theoretically novel: SDT addresses motivational conditions while CARE addresses the relational quality of interactions, and their combined operationalisation has not previously been demonstrated in a functioning artefact. Practical contribution: The MiCARE prototype demonstrates how digital wellness platforms can move beyond simple monitoring and tracking toward supportive, engaging, and human-centred experiences. The DR–DP–DF framework provides a transferable design template for developers and healthcare organisations building responsible digital wellness interventions. The clinician-reviewed, bounded chatbot design offers a practical model for responsible conversational design that avoids inappropriate clinical advice.

This study has three limitations. First, the evaluation was limited to short-term formative usability testing rather than long-term real-world deployment. Second, the usability evaluation sample was small and drawn from a single geographic region, limiting generalisability to broader populations and cultural contexts. Third, the rule-based chatbot faced constraints common in bounded conversational systems when responding to ambiguous inputs. Importantly, sustained engagement remains an open challenge. Prior research highlights the "eHealth law of attrition" (Eysenbach, 2005), where user engagement declines over time despite initial usability. While

MiCARE demonstrates strong initial interaction quality, long-term adoption will depend on continued novelty, contextual relevance, and perceived value. Future work will incorporate adaptive and outcome-oriented features within appropriate ethical approvals. Future work may also explore structured requirement–feature mapping approaches such as Quality Function Deployment (QFD) to expand the DR–DP–DF framework while maintaining traceability between requirements and system features.

8 Conclusion

This study demonstrates how motivational and compassionate design principles can be translated into operational system mechanisms, addressing the theory-to-design gap in digital health. Using DSRM, MiCARE operationalises SDT and the CARE framework into an explicit DR–DP–DF structure. Formative evaluation provides preliminary evidence that autonomy support, competence-oriented coaching, and empathic interaction were positively perceived by users. Findings remain early-stage and usability-focused; long-term engagement and behavioural outcomes require further validation. The study contributes a design-oriented framework and an initial artefact demonstrating feasibility within a preventive, non-clinical wellness context.

Acknowledgment

The authors thank Optus for research support.

End notes

This study was conducted as part of a PhD research project at La Trobe University.

References

Amirabdolahian, S., Pare, G., and Tams, S. 2025. "Digital Wellness Programs in the Workplace: Meta-Review," *J Med Internet Res* (27), p. e70982.

Anser, M. K., Nabi, A. A., Ahmad, I., Abro, M. M. Q., and Zaman, K. 2025. "Advancing Mental Health Care: A Comprehensive Review of Digital Tools and Technologies for Enhancing Diagnosis, Treatment, and Wellness," *Health Care Sci* (4:3), pp. 163-178.

Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., and Smith, D. M. 2023. "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum," *JAMA internal medicine* (183:6), pp. 589-596.

Babu, A., and Joseph, A. P. 2025. "Digital Wellness or Digital Dependency? A Critical Examination of Mental Health Apps and Their Implications," *Front Psychiatry* (16), p. 1581779.

Bailey-Davis, L., Wood, G. C., Cook, A., Cunningham, K., Jamieson, S., Mowery, J., Naylor, A., Rolston, D. D., Seiler, C., and Still, C. D. 2021. "Communicating Personalized Risk of Diabetes and Offering Weight Reduction Program Choice: Recruitment, Participation, and Outcomes," *Patient education and counseling* (104:5), pp. 1193-1199.

Bangor, A., Kortum, P. T., and Miller, J. T. 2008. "An Empirical Evaluation of the System Usability Scale," International Journal of Human-Computer Interaction (24:6), pp. 574–594.

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., and Rossi, M. 2018. "Design Science Research Contributions: Finding a Balance between Artifact and Theory," Journal of the Association for Information Systems (19:5), pp. 358–376.

Braun, V., and Clarke, V. 2006. "Using Thematic Analysis in Psychology," Qualitative Research in Psychology (3:2), pp. 77–101.

Bodai, B. I., Nakata, T. E., Wong, W. T., Clark, D. R., Lawenda, S., Tsou, C., Liu, R., Shiue, L., Cooper, N., Rehbein, M., Ha, B. P., McKeirnan, A., Misquitta, R., Vij, P., Klonecke, A., Mejia, C. S., Dionysian, E., Hashmi, S., Greger, M., Stoll, S., and Campbell, T. M. 2018. "Lifestyle Medicine: A Brief Review of Its Dramatic Impact on Health and Survival," *Perm J* (22), pp. 17-025.

Castro, V. d., Martín-Peña, M. L., Martínez, E. M., and Salgado, M. 2025. "Combining Action Research with Design Science as a Qualitative Research Methodology. An Application to Service (Operations) Management Research," *International Journal of Qualitative Methods* (24), p. 16094069241312018.

Creswell, J. W., and Plano Clark, V. L. 2018. Designing and Conducting Mixed Methods Research (3rd ed.). Thousand Oaks, CA: Sage.

Del-Valle-Soto, C., López-Pimentel, J. C., Vázquez-Castillo, J., Nolazco-Flores, J. A., Velázquez, R., Varela-Aldás, J., and Visconti, P. 2024. "A Comprehensive Review of Behavior Change Techniques in Wearables and Iot: Implications for Health and Well-Being," *Sensors (Basel)* (24:8).

Demissie, G. D., Birungi, J., Shrestha, A., Haregu, T., Thirunavukkarasu, S., and Oldenburg, B. 2025. "The Effectiveness of Lifestyle Interventions in Reducing Cardiovascular Risk and Risk Factors in People with Prediabetes: A Systematic Review and Meta-Analysis," *Nutrition, Metabolism and Cardiovascular Diseases* (35:10), p. 104130.

Dostert, J., and Müller, R. 2021. "Motivational Assistance System Design for Industrial Production: From Motivation Theories to Design Strategies," *Cognition, Technology & Work* (23:3), pp. 507-535.

Eiselt, A. K., Kirkendall, S., Xiong, E., Langner, D., and Goldfarb, M. 2025. "Achieving Clinically Meaningful Outcomes in Digital Health: A Six-Step, Cyclical Precision Engagement Framework (Engage)," *Front Digit Health* (7), p. 1713334.

Eysenbach, G. (2005). The law of attrition. Journal of Medical Internet Research, 7(1), e11. https://doi.org/10.2196/jmir.7.1.e11

Gregor, S., and Zwikael, O. 2024. "Design Science Research and the Co-Creation of Project Management Knowledge," *International Journal of Project Management* (42:3), p. 102584.

Hazra-Ganju, A., Dlima, S. D., Menezes, S. R., Ganju, A., and Mer, A. 2023. "An Omni-Channel, Outcomes-Focused Approach to Scale Digital Health Interventions in Resource-Limited Populations: A Case Study," *Front Digit Health* (5), p. 1007687.

Heinsch, M., Wyllie, J., Carlson, J., Wells, H., Tickner, C., and Kay-Lambkin, F. 2021. "Theories Informing Ehealth Implementation: Systematic Review and Typology Classification," *J Med Internet Res* (23:5), p. e18500.

Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS quarterly* (28:1), pp. 75-105.

Hussain, W., and Grundy, J. 2025. "Advice for Diabetes Self-Management by Chatgpt Models: Challenges and Recommendations,").

LeSeure, P., Chin, E., and Zhang, S. 2024. "A Culturally Sensitive Mobile App (Diafriend) to Improve Self-Care in Patients with Type 2 Diabetes: Development Study," *JMIR diabetes* (9), p. e63393.

Li, M., Liu, S., Yu, B., Li, N., Lyu, A., Yang, H., He, H., Zhang, N., Ma, J., Sun, M., Du, H., and Gao, R. 2025a. "Assessing the Effectiveness of Digital Health Behavior Strategies on Type 2 Diabetes Management: Systematic Review and Network Meta-Analysis," *J Med Internet Res* (27), p. e63209.

Li, X., Pei, X., and Zhao, J. 2025b. "Intrinsic Motivation and Self-Efficacy as Pathways to Innovative Teaching: A Mixed-Methods Study of Faculty in Chinese Higher Education," *BMC Psychol* (13:1), p. 859.

Lim, S. L., Tay, M. H. J., Ong, K. W., Johal, J., Yap, Q. V., Chan, Y. H., Yeo, G. K. N., Khoo, C. M., and Yaxley, A. 2022. "Association between Mobile Health App Engagement and Weight Loss and Glycemic Control in Adults with Type 2 Diabetes and Prediabetes (D'lite Study): Prospective Cohort Study," *JMIR diabetes* (7:3), p. e35039.

Lippman, D., Stump, M., Veazey, E., Guimarães, S. T., Rosenfeld, R., Kelly, J. H., Ornish, D., and Katz, D. L. 2024. "Foundations of Lifestyle Medicine and Its Evolution," *Mayo Clin Proc Innov Qual Outcomes* (8:1), pp. 97-111.

Martela, F., and Riekki, T. J. J. 2018. "Autonomy, Competence, Relatedness, and Beneficence: A Multicultural Comparison of the Four Pathways to Meaningful Work," *Front Psychol* (9), p. 1157.

Nadarzynski, T., Knights, N., Husbands, D., Graham, C. A., Llewellyn, C. D., Buchanan, T., Montgomery, I., Ridge, D., and Silva, J. N. A. 2024. "Achieving Health Equity through Conversational Ai: A Roadmap for Design and Implementation of Inclusive Chatbots in Healthcare," PLOS digital health (3:5), p. e0000492.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (pp. 206–213). New York: ACM. https://doi.org/10.1145/169059.169166

Oh, Y. J., Zhang, J., Fang, M.-L., and Fukuoka, Y. 2021. "A Systematic Review of Artificial Intelligence Chatbots for Promoting Physical Activity, Healthy Diet, and Weight Loss," The international journal of behavioral nutrition and physical activity (18:1), pp. 160-160.

Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. 2008. "A Design Science Research Methodology for Information Systems Research," *Journal of management information systems* (24:3), p. 45.

Persson, J., and Rydenfält, C. 2021. "Why Are Digital Health Care Systems Still Poorly Designed, and Why Is Health Care Practice Not Asking for More? Three Paths toward a Sustainable Digital Work Environment," *J Med Internet Res* (23:6), p. e26694.

Ryan, R. M., and Deci, E. L. 2000. "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," *The American psychologist* (55:1), pp. 68-78.

Sinclair, S., Beamer, K., Hack, T. F., McClement, S., Raffin Bouchal, S., Chochinov, H. M., and Hagen, N. A. 2017. "Sympathy, Empathy, and Compassion: A Grounded Theory Study of Palliative Care Patients' Understandings, Experiences, and Preferences," *Palliative medicine* (31:5), pp. 437-447.

Sinclair, S., Hack, T. F., Raffin-Bouchal, S., McClement, S., Stajduhar, K., Singh, P., Hagen, N. A., Sinnarajah, A., and Chochinov, H. M. 2018. "What Are Healthcare Providers' Understandings and Experiences of Compassion? The Healthcare Compassion Model: A Grounded Theory Study of Healthcare Providers in Canada," *BMJ open* (8:3), pp. e019701-e019701.

Szinay, D., Jones, A., Chadborn, T., Brown, J., and Naughton, F. 2020. "Influences on the Uptake of and Engagement with Health and Well-Being Smartphone Apps: Systematic Review," *J Med Internet Res* (22:5), p. e17572.

Thanthrige A, Wickramasinghe N. Digital Health Solutions for Type 2 Diabetes and Prediabetes: Systematic Review of Engagement Barriers, Facilitators, and Outcomes. JMIR Diabetes. 2026 Mar 12;11:e80582. doi: 10.2196/80582. PMID: 41818488; PMCID: PMC12981377.

Tzimourta, K. D. 2025. "Human-Centered Design and Development in Digital Health: Approaches, Challenges, and Emerging Trends," *Cureus* (17:6), p. e85897.

van Lotringen, C., Lusi, B., Westerhof, G. J., Ludden, G. D. S., Kip, H., Kelders, S. M., and Noordzij, M. L. 2023. "The Role of Compassionate Technology in Blended and Digital Mental Health Interventions: Systematic Scoping Review," *JMIR Ment Health* (10), p. e42403.

Vargas, C., Zorbas, C., Longworth, G. R., Ugalde, A., Needham, C., Sunil, A., Venegas Hargous, C., Bennett, R., Forrester-Bowling, T., Cardoso Richter, A. P., and Allender, S. 2025. "Exploring Co-Design: A Systematic Review of Concepts, Processes, Models, and Frameworks Used in Public Health Research," *J Public Health (Oxf)* (47:4), pp. e616-e639.

Wiegmann, P. M., Talmar, M., and de Nijs, S. B. 2023. "Forging a Sharper Blade: A Design Science Research Approach for Transition Studies," *Environmental Innovation and Societal Transitions* (48), p. 100760.

Wu, Y., Zhang, J., Ge, P., Duan, T., Zhou, J., Wu, Y., Zhang, Y., Liu, S., Liu, X., Wan, E., and Sun, X. 2024. "Application of Chatbots to Help Patients Self-Manage Diabetes: Systematic Review and Meta-Analysis," *Journal of medical Internet research* (26:10397), p. e60380.

Xue, H., Zhang, L., Shi, Y., Zhang, H., Zhang, C., Liu, Y., Tan, W., and Liu, Y. 2025. "The Effectiveness of Digital Health Intervention on Glycemic Control and Physical Activity in Patients with Type 2 Diabetes: A Systematic Review and Meta-Analysis," *Front Digit Health* (7), p. 1630588.

Zhang, C., Lakens, D., and WA, I. J. 2021. "Theory Integration for Lifestyle Behavior Change in the Digital Age: An Adaptive Decision-Making Framework," *J Med Internet Res* (23:4), p. e17127.

Zhang, J., Oh, Y. J., Lange, P., Yu, Z., and Fukuoka, Y. 2020. "Artificial Intelligence Chatbot Behavior Change Model for Designing Artificial Intelligence Chatbots to Promote Physical Activity and a Healthy Diet: Viewpoint," *Journal of medical Internet research* (22:9), pp. e22845-e22845.

About the authors

Ayesha Thanthrige is a PhD candidate at La Trobe University, specialising in digital health and AI-driven wellness systems. With over a decade of experience in IT as a Business Analyst and Project Manager, she focuses on designing patient-centred, technology-enabled solutions for sustainable healthcare. She holds a BSc (Hons) in Management Information Systems and an MBA, and has contributed to several digital health publications.

Professor **Nilmini Wickramasinghe** is the Optus Chair and Professor of Digital Health at La Trobe University. With over 20 years of international experience, her research focuses on digital health innovation, AI, and value-based care. She is a pioneer in applying digital twins in healthcare and received the Alexander von Humboldt Award for her contributions to the field.

Summary

This study addresses the theory-to-design gap in digital wellness by translating motivational and compassionate care principles into operational system mechanisms. Using Design Science Research Methodology, the MiCARE prototype integrates Self-Determination Theory and the CARE framework into a structured DR–DP–DF model. The system supports autonomy through personalised goal setting, competence through coaching and feedback, and empathy through conversational interaction. A formative usability evaluation demonstrated high usability, strong task completion, and high perceived empathy. The findings show that theory-driven design can be effectively operationalised in

digital wellness systems, contributing a transparent framework and a practical prototype for developing engaging, user-centred healthcare technologies.

The Impact of Generative AI on B2B Marketing Processes: Evidence from Industrial Firms

Mikko Vesterinen, Joel Mero, Mika Skippari, Heikki Karjaluoto
University of Jyväskylä, Jyväskylä, Finland
mivijove@jyu.fi, joel.j.mero@jyu.fi, mika.t.t.skippari@jyu.fi, heikki.karjaluoto@jyu.fi

Generative artificial intelligence (GenAI) is transforming marketing, yet its implications for B2B marketing processes and skills remain insufficiently understood, particularly in industrial contexts where AI adoption has lagged. This study examines how GenAI is used in industrial marketing, its impact on B2B marketing processes, and what skills marketers need. Drawing from 43 semi-structured interviews with AI and marketing leaders from multinational industrial firms and technology providers, we map GenAI applications across marketing information management, planning, and implementation, and categorize them by input type and degree of human augmentation. The findings reveal an uneven impact: GenAI affects implementation most, with increasing automation and customization, whereas its use for information management and planning remains largely unsystematic. As automation increases, marketers' roles shift toward market research, strategy formation, and coordinating activities. The study shows how GenAI reshapes B2B marketing processes and associated skill requirements, offering guidance for systematically embedding GenAI in industrial marketing.

DOI
https://doi.org/10.18690/um.fov.4.2026.32

ISBN
978-961-299-152-4

Keywords:
generative AI, B2B marketing processes, industrial firms, AI applications, marketing skills

University of Maribor Press

1 Introduction

Marketing is among the business functions most affected by GenAI (Grewal et al., 2025), as its applications span core marketing processes (Hermann & Puntoni, 2025). Customers, for example, increasingly use GenAI to search for information and to interact with firms online (Grewal et al., 2025). Additionally, marketing use cases rank among the highest in organizational value creation (Chui et al., 2023). GenAI offers opportunities to reduce marketing resources and expenses, but also challenges marketers to demonstrate their value for the company (Cillo & Rubera, 2025). As a result, marketing processes and required skills are being reshaped, with priorities shifting from tasks where GenAI outperforms humans to those where human judgment and creativity remain superior (Hermann & Puntoni, 2025). However, the impact of GenAI may differ between B2B and B2C contexts due to fundamental differences between them (Mikalef et al., 2023). While B2B marketers have lagged in AI adoption (Moradi & Dass, 2022), their ability to create value increasingly depends on developing skills that meet evolving requirements.

Although research on GenAI, particularly in industrial marketing, remains nascent, prior studies have provided insights into GenAI's applications and broader marketing implications (Chan & Choi, 2025; Grewal et al., 2025; Kshetri et al., 2024). However, none have systematically addressed the use of GenAI in industrial marketing, its impact on B2B marketing processes, and the skills required of marketers. To fill this gap, this study addresses three research questions: (1) In what ways is GenAI used in industrial marketing? (2) How does GenAI influence B2B marketing processes, and how are these expected to develop? (3) What skills must industrial marketers possess to (a) adopt and integrate GenAI, and (b) manage its impact on B2B marketing processes?

This study is based on 43 interviews with three groups: industrial marketing managers, AI managers, and AI consultants. From these interviews, GenAI applications are identified and classified by input type (general vs. custom) and degree of human augmentation. By categorizing applications based on their support for B2B marketing processes, the study enables an assessment of which activities have been augmented or disrupted by GenAI. Findings also clarify the skills industrial marketers need to adopt, integrate, and strategically leverage GenAI

effectively. Managerial implications, in turn, guide AI and marketing leaders driving GenAI transformation in industrial marketing.

2 Literature Review

2.1 B2B Marketing processes

Marketing processes encompass the structured routines through which organizations coordinate and deploy resources to achieve desired performance outcomes (Day, 1994; Herhausen et al., 2020). In B2B marketing, three core processes underpin these activities: marketing information management, marketing planning, and marketing implementation (Mikalef et al., 2023), often complemented by the marketing mix (4Ps) (Moorman & Day, 2016; Morgan et al., 2009; Vorhies & Morgan, 2005). Marketing information management encompasses the acquisition, analysis, and dissemination of market and stakeholder knowledge (Mikalef et al., 2023; Moorman & Day, 2016). Marketing planning involves developing strategies based on market intelligence and organizational resources (Varadarajan, 2010). Marketing implementation captures the execution, coordination, and evaluation of these strategies (Mikalef et al., 2023; Vorhies & Morgan, 2005).

2.2 The impact of GenAI on marketing processes and skills

Before the emergence of GenAI, machine learning had already been applied to marketing research, strategy, and decision-making tasks requiring complex analytical logic (Huang & Rust, 2021). With GenAI's ability to generate novel content through human-like reasoning from vast volumes of unstructured data and its accessible, user-friendly tools, marketing is primed for transformation (Grewal et al., 2025). Literature highlights many GenAI applications across marketing activities. These include generating customer and market insights (marketing information management), drafting and evaluating strategies and positioning statements (marketing planning), and producing personalized content, as well as enhancing lead generation, customer service, and advertising (marketing implementation) (Chan & Choi, 2025; Grewal et al., 2025; Kshetri et al., 2024).

Grewal et al. (2025) categorize GenAI applications along two dimensions: input type (general vs. custom) and the degree of human augmentation. They elaborate that general-input applications rely on broadly accessible data, whereas custom-input applications, such as retrieval-augmented generation (RAG), draw on task-specific data, increasing reliability and task specificity at the cost of greater development complexity. The degree of human augmentation indicates the extent of required human oversight and editing, and thus the level of automation (Grewal et al., 2025). While general-purpose chatbots achieve higher adoption due to their ease of use and flexibility, they lack deep integration into workflows, whereas custom applications enable stronger process integration but exhibit lower adoption rates due to higher complexity (Challapally et al., 2025).

Hermann & Puntoni (2025) argue that human replacement is more likely in marketing information management and implementation tasks, whereas marketing planning, often grounded in tacit knowledge, heuristics, and subjective judgment, remains more amenable to augmentation than substitution. Human augmentation is also often still required to guide AI in customer-facing applications to maintain trust and authenticity (Grigsby et al., 2025). The automation and enhancement potential of GenAI may challenge marketing in cost-focused firms that view it as a cost center, increasing pressure to demonstrate its strategic value (Cillo & Rubera, 2025). At the same time, GenAI reshapes skill requirements (Dwivedi et al., 2023). It frees marketers to focus on higher-value activities while augmenting their reasoning and access to unstructured data (Cillo & Rubera, 2025). Consequently, marketers must develop the skills to use GenAI tools effectively and to critically assess their potential and limitations (Grewal et al., 2025).

In summary, B2B marketing encompasses three core processes: information management, planning, and implementation (Mikalef et al., 2023), which have the potential to be disrupted by GenAI. GenAI applications differ along two principal dimensions: whether they depend on general or customized input, and the extent of human augmentation necessary, which determines their level of task specificity and their potential for automation (Grewal et al., 2025). As GenAI transforms these processes, marketing professionals must cultivate complementary skills to adopt it effectively and manage its impact on their evolving roles (Dwivedi et al., 2023).

3 Methodology

This study employs semi-structured interviews to address the research questions. Given the early stage of GenAI research, an exploratory qualitative approach is appropriate (Magaldi & Berler, 2020). The sample comprised three respondent groups, identified and contacted via LinkedIn and professional networks. The first group included managers and directors of AI, IT, and digital roles at large multinational industrial firms. These informants provided insights into deployed GenAI models and tools, development processes, training initiatives, and adoption strategies. The second group consisted of AI consultants, managers, and executives from technology consulting and cloud platform providers. Owing to their cross-industry client work, they offered a holistic perspective on the use of GenAI in industry. The third group included marketing leaders, managers, and marketing technology specialists from multinational industrial firms, who contributed domain-specific insights into the use of GenAI in marketing. In addition to providing an overview of the current situation regarding GenAI, respondents were asked what types of development they expected in the future.

To expand the participant pool, snowball sampling was used by requesting referrals from interviewees and selecting those meeting the sampling criteria (Noy, 2008). In total, 43 interviews were conducted with representatives from 18 industrial firms and 10 AI and cloud platform vendors. Interviews lasted 32 to 112 minutes (average 61 minutes), and were held remotely via MS Teams, which was also used for transcription. Overall, 1,498 A4 pages of transcripts were analyzed. Table 1 summarizes the respondent groups.

Table 1: Sampling and data collection for the research

	Group 1 AI managers 2/2024 – 9/2025	**Group 2** AI consultants 2/2024 – 2/2025	**Group 3** Marketing managers 1/2025 – 11/2025
Job roles	AI-, IT- and digital manager or director	AI consultant, AI manager, C-level	Head of marketing, marketing manager, MarTech manager
Company industry	Industrial (manufacturing, process, machinery)	Technology consultation, cloud platforms	Industrial (manufacturing, process, machinery)
Sample	23 interviews 14 companies	10 interviews 10 companies	10 interviews 9 companies

ATLAS.ti was used to review, code, and analyze transcripts. The interview data were analyzed using an abductive qualitative approach characterized by systematic combining (Dubois & Gadde, 2002). The analysis proceeded iteratively between interview data and relevant theory. First, open, inductive coding was conducted to identify respondent-centric concepts grounded in participants' descriptions of current GenAI use and future development directions. Through constant comparison across interviews, these codes were refined and grouped into broader themes that describe marketing activities, application types, and the skills required of marketers. In the subsequent stage, abductive reasoning was employed by iteratively comparing the emerging themes with B2B marketing processes to identify connections between the two. This method grounded the findings in empirical insight, while enabling the contribution to ongoing theoretical discussions.

4 Findings

The findings are structured around the B2B marketing processes framework (Mikalef et al., 2023). The interviews identified specific activities in which B2B marketers leverage GenAI, grouped into marketing information management, planning, and implementation. For each activity, the GenAI application type is specified based on the framework of Grewal et al. (2025), and its impact on marketing processes and skills is evaluated. Table 2 summarizes the findings.

Table 2: The impact of GenAI on B2B marketing processes and skills

Marketing process	Impact of GenAI	Requirements for marketer skills
Information management	**Currently:** Sporadic use of GenAI and absent workflows. – Processes largely absent, as information management is seldom viewed as part of industrial marketing. – GenAI use is sporadic and limited to market intelligence with general-input, human-augmented tools. No plans to develop specialized applications. **Future:** Increased focus and established processes. The use of GenAI becomes more systematic. – As GenAI continues to automate and augment implementation activities, marketers' focus shifts to information management.	Skills required as focus shifts to information management: – **Market research:** collecting, processing, analyzing, and disseminating insights from internal and external data. – **Relational skills:** strengthening collaboration and connections with customers and sales.

Marketing process	Impact of GenAI	Requirements for marketer skills
	– Automation and task specificity may increase but require more structured processes enabled by increased marketer focus on this domain.	
Marketing planning	**Currently:** Limited but emerging use, with some firms planning deeper integration. – GenAI use remains largely unsystematic, mainly supporting campaign planning with general-input, human-augmented tools with some exceptions. – Plans exist to introduce custom-input tools and better embed AI in planning processes. **Future:** Increased focus and workflow integration. – As GenAI increasingly automates and augments implementation, focus shifts on planning. – Task specificity may increase in the short term, and partial automation is a longer-term opportunity.	Skills required as focus shifts to marketing planning: – **Strategic thinking:** proactively shaping marketing strategy. – **Data-driven decision-making**: basing decisions on data and insights.
Marketing implementation	**Currently:** Widespread use, with active development toward greater automation and task specificity. – GenAI is widely used in content creation, lead management, and optimization. – All applications require human augmentation. Lead management and optimization employ custom input. **Future:** Reduced need for human augmentation but continued oversight. – Greater automation and task specificity, indicating rapid development in this domain. – Focus shifts to content management, as AI-generated content requires oversight and authentic content becomes more valuable. – Greater coordination and analysis of partners and activities is required as fewer tasks are performed manually in-house.	Skills required as GenAI augments and automates activities: – **Creativity**: maintaining authentic content. – **Orchestration and coordination:** managing marketing assets and partners. – **Marketing analytics:** analyzing, interpreting, and acting on marketing performance data
GenAI-specific skills required from marketers: – **Prompting:** formulating, scoping, and iteratively refining inputs to GenAI. – **AI literacy:** uderstanding what AI can and cannot do and using AI responsibly. – **Experimental mindset:** treating AI adoption as an iterative, learning-oriented process. – **Knowledge of IT processes:** understanding how to progress AI initiatives.		

4.1 Information management

Activities related to marketing information management included market intelligence, comprising competitor, customer, and market analysis. Most used general-input, highly human-augmented AI chatbots (e.g., MS Copilot, ChatGPT) for this purpose, at least sporadically. Their effectiveness often depended on

individual prompting skills, and some respondents expressed dissatisfaction with output quality, attributing this to limited prompting expertise. Also, one respondent reported experimenting with CRM-integrated AI features for customer analysis. A potential GenAI application for market intelligence was discussed in one interview:

For example, artificial intelligence can monitor the web to see which construction projects are starting in different parts of the world and who is involved, which helps us identify companies that could be similar to our current customers. So, that's really interesting, and then we can target them. (Marketing manager)

GenAI applications in information management were limited and unsystematic, and this area presents significant opportunities, especially for automation (Hermann & Puntoni, 2025). Although most companies had a dedicated market intelligence function, this domain was considered important, but the absence of processes and skills limited its development. Marketers were unaccustomed to conducting market research, and, beyond prompting skills, domain expertise gained through closer interactions with customers and frontline personnel would further enhance their information management capacity and use of GenAI. Unified processes were identified as crucial in adopting AI in this domain:

*To do that, you need to get the processes in order, and we don't really have unified processes. We have substantial groundwork to do before we can get to more useful things, such as thinking about automation. For that, you need a unified process, and as long as everyone handles the process however they like, it cannot be automated. (*Head of digital marketing)

Looking ahead, respondents anticipate that as GenAI increasingly automates implementation activities, marketers will reallocate capacity toward more systematic information management. Along this trajectory, the growth of automation and task specificity within this domain is expected to be contingent upon the development of more structured processes and a deeper expertise in the domain.

4.2 Marketing planning

The use of GenAI in marketing planning was limited to marketing campaign planning. Only two respondents reported systematic use, each employing a distinct approach. One described integrating digital marketing data from various channels

into the Azure cloud platform and combining it with internal datasets. A chatbot interface enabled marketers to prompt a custom-input model to analyze marketing performance and generate audience-specific recommendations. The tool provided insights into digital asset performance and suggested improvements to existing materials. This firm also outlined its forthcoming AI marketing planner:

This means that for each division, they specify its primary target audiences and channels. We then connect the AI to our raw data so it can learn from the entire dataset; the goal is to enable it to make strategic recommendations to the target audience, content, budgets, and other levels. (MarTech manager)

In contrast, another approach relied on prompting general-purpose GenAI chatbots for benchmarks and recommendations by querying information on content formats, advertising examples, and campaign structures to support planning, or by providing campaign attributes (such as target audiences, topics, and objectives) and requesting structured campaign suggestions. Marketing planning was generally viewed as less suitable for automation, as strategic decisions require human judgment. Customizing inputs with proprietary data, on the other hand, could enable firm-specific recommendations. GenAI supports the shift from intuition to data-driven decision-making, though AI literacy and systematic processes are essential. Moreover, strategic thinking was highlighted to integrate the insights into marketing strategies. Respondents hoped to offload implementation tasks to GenAI, freeing capacity for planning, but also sought AI support for strategic tasks:

I'd rather see AI used for decision-making, marketing planning, and improving customer understanding — rather than just content production — because understanding the bigger picture is more challenging and creates more value. (...) The value humans add is to strategically focus our efforts (VP, Marketing)

Although current utilization remains predominantly unsystematic, respondents anticipated a gradual transition towards enhanced workflow integration, whereby custom-input applications would facilitate more firm-specific strategic recommendations. Partial automation of planning activities was regarded as a long-term prospect, contingent upon the concurrent development of AI literacy and structured procedures.

4.3 Marketing implementation

Marketing implementation included several GenAI-enhanced activities, such as content creation and curation, lead management, and asset optimization. Most firms encouraged the use of company-specific tools to mitigate confidentiality risks, and marketers mainly relied on general-input, human-augmented tools, while some firms used translation bots trained on brand tone. Plans existed to develop customized bots using proprietary data to generate and repurpose content, potentially linked to performance data for channel-specific optimization. GenAI-generated visuals were restricted in some organizations but used internally in others. One developed the capacity to deploy them in external channels. Respondents emphasized the need for strong human oversight and customized inputs for AI-generated content and sought a more structured way of using GenAI without sacrificing brand integrity:

My hope this year is to find a more structured way to use [GenAI], because when you use the social media channels' own AI tools and other tools, the company's own voice sometimes gets lost. (...) The idea is to emphasize that you should use it to speed up the process, while maintaining authenticity, keeping the main messages in mind, and so on. (Head of marketing)

Lead management was identified as a promising application area. Respondents noted that digital inquiries are diverse and context-specific, often requiring additional clarification before conversion. Website-based service bots and automated follow-up emails were largely in the planning phase, while lead scoring was already in use at two companies. These use cases typically rely on custom input and aim to reduce human interaction. However, their effectiveness remains constrained by marketers' limited trust in AI recommendations:

We still don't trust artificial intelligence-based scoring more than a rule-based version done by a human. (...) People don't trust values that indicate a good lead if they don't understand the scoring criteria. (MarTech manager)

Optimization activities covered copy, visuals, ad formats, and search and generative engine optimization (SEO and GEO). Marketers mainly used AI features embedded in advertising channels or specialized marketing platforms. These tools offered task-specific optimization with minimal human effort but less flexibility than general-purpose models. Marketing analytics was also identified as a promising area,

requiring basic analytical literacy to leverage GenAI effectively. Other skills respondents emphasize include content management and creativity to maintain authenticity; orchestration skills to manage how AI is used to create marketing assets and to manage campaign flows; and an experimental mindset and understanding of IT processes to effectively adopt and develop GenAI tools and applications. Overall, marketing implementation is experiencing the most substantial transformation, though integration into critical workflows remains limited:

However, the process is not yet at its peak or precise enough to significantly increase agility. [GenAI] is used in some places. And then there is that human labor in between, and then at some point, we return to using AI. However, an agile operating model as a whole could be explored further. (Marketing manager)

Respondents anticipated that this trajectory would persist, with the demand for human augmentation decreasing as tools become more advanced and more thoroughly integrated into workflows. Nonetheless, oversight of AI-generated content was projected to remain crucial as manual tasks are increasingly automated, placing greater emphasis on content management, campaign orchestration, and performance analytics.

5 Conclusion

This study advances the literature in four key ways. First, it identifies GenAI applications in industrial marketing, an underexplored area, and categorizes them under B2B marketing processes (Mikalef et al., 2023). Applications in information management and planning were limited to market intelligence and campaign planning, whereas implementation-related uses, content creation, lead management, and optimization, were more prevalent, with content creation most common. Second, it evaluates these applications in terms of required human augmentation and reliance on general versus custom input (Grewal et al., 2025), indicating their level of task specificity and automation. Information management and planning primarily rely on general input and high human augmentation, suggesting unrealized opportunities for automation and customization (Hermann & Puntoni, 2025). In contrast, implementation activities employ more custom-input solutions and show higher automation.

Third, the findings show that GenAI's impact is concentrated in marketing implementation, while its influence on information management and planning remains limited. Development efforts focus primarily on implementation, indicating continued disruption through increased automation and task-specificity. This allows marketers to shift focus toward interpreting large-scale data using tacit and proprietary organizational knowledge inaccessible to GenAI (Hermann & Puntoni, 2025) and to use these insights to proactively shape strategic direction, thereby strengthening marketing's value contribution (Cillo & Rubera, 2025).

Fourth, the study specifies the required skills to leverage GenAI and embrace the new strategic role. As implementation becomes increasingly automated, content management and creativity remain critical to maintaining quality (Grigsby et al., 2025) while campaign orchestration and analytics support workflow improvement. A shift in focus to information management and planning emphasizes the capacity to gather and harness market and proprietary data, which is particularly valuable for industrial firms that rely on detailed customer information (Mikalef et al., 2023), coupled with data-driven decision-making and strategic thinking. Effective GenAI adoption further requires AI literacy, an experimental mindset, technological fluency, prompting skills, and understanding of IT processes.

The findings suggest that AI managers should facilitate the development of high-value GenAI applications in marketing, as transformative use cases remain underdeveloped and marketers often lack the skills to initiate them. Given marketing's role at the sales-customer interface, such applications can directly influence growth and customer orientation. Marketing leaders should systematically integrate GenAI into processes to enhance consistency and to free resources for strategic work. This requires collaboration with IT on targets, training, and governance. Currently, adoption is largely driven by individual initiative, highlighting the need for more proactive leadership.

As an exploratory qualitative study, this research does not quantify impact or comprehensively map GenAI applications but provides a foundation for future research. Given the rapid evolution of GenAI and its limited workflow integration, continued relevance for this research stream requires ongoing dialogue between academia and practice. Finally, the predominantly Finnish-based sample limits this study's generalizability and calls for broader international research.

Acknowledgment

This work was supported by the Foundation for Economic Education.

References

Challapally, A., Pease, C., Raskar, R., & Chari, P. (2025). *The GenAI divide. State of AI in business 2025*. https://www.artificialintelligence-news.com/wp-content/uploads/2025/08/ai_report_2025.pdf

Chan, H.-L., & Choi, T.-M. (2025). Using generative artificial intelligence (GenAI) in marketing: Development and practices. *Journal of Business Research*, *191*, 115276. https://doi.org/10.1016/j.jbusres.2025.115276

Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zemmel, R. (2023, June). *The economic potential of generative AI: The next productivity frontier*. McKinsey. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

Cillo, P., & Rubera, G. (2025). Generative AI in innovation and marketing processes: A roadmap of research opportunities. *Journal of the Academy of Marketing Science*, *53*, 684–701. https://doi-org.ezproxy.jyu.fi/10.1007/s11747-024-01044-7

Day, G. S. (1994). The Capabilities of market-driven organizations. *Journal of Marketing, 58*(4), 37–52. https://doi.org/10.1177/002224299405800404

Dubois, A., & Gadde, L.-E. (2002). Systematic combining: An abductive approach to case research. *Journal of Business Research*, *55*(7), 553–560. https://doi.org/10.1016/S0148-2963(00)00195-8

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Grewal, D., Satornino, C. B., Davenport, T., & Guha, A. (2025). How generative AI is shaping the future of marketing. *Journal of the Academy of Marketing Science*, *53*(3), 702–722. https://doi.org/10.1007/s11747-024-01064-3

Grigsby, J. L., Michelsen, M., & Zamudio, C. (2025). Service ads in the era of generative AI: Disclosures, trust, and intangibility. *Journal of Retailing and Consumer Services*, *84*, 104231. https://doi.org/10.1016/j.jretconser.2025.104231

Herhausen, D., Miočević, D., Morgan, R. E., & Kleijnen, M. H. P. (2020). The digital marketing capabilities gap. *Industrial Marketing Management*, *90*, 276–290. https://doi.org/10.1016/j.indmarman.2020.07.022

Hermann, E., & Puntoni, S. (2025). Generative AI in marketing and principles for ethical design and deployment. *Journal of Public Policy & Marketing*, *44*(3), 332–349. https://doi.org/10.1177/07439156241309874

Huang, M.-H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, *49*(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9

Kshetri, N., Dwivedi, Y. K., Davenport, T. H., & Panteli, N. (2024). Generative artificial intelligence in marketing: Applications, opportunities, challenges, and research agenda. *International Journal of Information Management*, *75*, 102716. https://doi.org/10.1016/j.ijinfomgt.2023.102716

Magaldi, D., & Berler, M. (2020). Semi-structured interviews. In: Zeigler-Hill, V., Shackelford, T.K. (eds) Encyclopedia of personality and individual differences (pp. 4825–4830). Springer, Cham. https://doi.org/10.1007/978-3-319-24612-3_857

Mikalef, P., Islam, N., Parida, V., Singh, H., & Altwaijry, N. (2023). Artificial intelligence (AI) competencies for organizational performance: A B2B marketing capabilities perspective. *Journal of Business Research*, *164*, 113998. https://doi.org/10.1016/j.jbusres.2023.113998

Moorman, C., & Day, G. S. (2016). Organizing for marketing excellence. *Journal of Marketing*, *80*(6), 6–35. https://doi.org/10.1509/jm.15.0423

Moradi, M., & Dass, M. (2022). Applications of artificial intelligence in B2B marketing: Challenges and future directions. *Industrial Marketing Management*, *107*, 300–314. https://doi.org/10.1016/j.indmarman.2022.10.016

Morgan, N. A., Vorhies, D. W., & Mason, C. H. (2009). Market orientation, marketing capabilities, and firm performance. *Strategic Management Journal*, *30*(8), 909–920. https://doi.org/10.1002/smj.764

Noy, C. (2008). Sampling Knowledge: The hermeneutics of snowball sampling in qualitative research. *International Journal of Social Research Methodology*, *11*(4), 327–344. https://doi.org/10.1080/13645570701401305

Varadarajan, R. (2010). Strategic marketing and marketing strategy: Domain, definition, fundamental issues and foundational premises. *Journal of the Academy of Marketing Science*, *38*(2), 119–140. https://doi.org/10.1007/s11747-009-0176-7

Vorhies, D. W., & Morgan, N. A. (2005). Benchmarking marketing capabilities for sustainable competitive advantage. *Journal of Marketing*, *69*(1), 80–94. https://doi.org/10.1509/jmkg.69.1.80.55505

Structure of Data-Driven Business Models for the Circular Automotive Sector: A Taxonomy

Hannah Scheerer,[1] Hendrik van der Valk[1,2]

[1] TU Dortmund University, Chair of Industrial Information Management, Dortmund, Germany
hannah.scheerer@tu-dortmund.de, hendrik.van-der-valk@tu-dortmund.de
[2] Fraunhofer Institute for Software and Systems Engineering, Dortmund, Germany
hendrik.van-der-valk@tu-dortmund.de

Companies in the automotive sector are under increasing pressure from stricter regulations, shifting customer expectations, and ambitious sustainability goals. The circular economy (CE) provides a promising pathway to address these challenges by promoting resource efficiency, lifecycle extension, and closed material loops. Meanwhile, data-driven business models (DDBM) enable enhanced transparency, optimization, and innovative value creation through digital technologies (DT). While the integration of circular strategies and data-driven approaches has significant transformative potential, circular DDBMs remain scarce in the automotive industry and lack systematic conceptualization. To address this gap, this study develops a taxonomy that structures the current landscape of circular DDBMs using a rigorous, iterative methodology. The resulting taxonomy, based on 42 empirical cases, comprises six meta-dimensions, 16 dimensions, and 68 characteristics, offering a comprehensive framework for analyzing, comparing, and designing circular DDBMs in the automotive sector.

DOI
https://doi.org/10.18690/um.fov.4.2026.33

ISBN
978-961-299-152-4

Keywords:
data-driven business models, circular economy, automotive, taxonomy, digital technologies

University of Maribor Press

1 Introduction

The automotive industry is undergoing continuous transformation driven by regulatory pressure, shifting customer demand, and technological innovation, such as electric vehicles, which challenge traditional linear business models (BM) with high environmental impacts (Grimm & Pfass, 2022; Urbinati et al., 2021). The European Green Deal further intensifies this shift by constraining linear value creation, particularly in a sector responsible for substantial life-cycle greenhouse gas emissions (Pichler et al., 2021). The CE concept offers a pathway to reduce these impacts through circular product and process design, with initial applications such as alternative drivetrains, battery systems, and component reuse already emerging (ACEA, 2018; Urbinati et al., 2021). However, high investment costs, organizational complexity, and supply chain adaptations still hinder large-scale adoption (Hoppe et al., 2025). In this context, digitalization and DTs, including data analytics, tracking, and monitoring, are key enablers of CBMs by enhancing transparency, resource efficiency, and data-driven decision-making (Luoma et al., 2021; One Planet Network, 2025), as they enable systematic supply chain data analysis, sustainability indicators, and improved stakeholder transparency (One Planet Network, 2025).

In the automotive sector, companies have recognized that data must be part of their BM, reflecting a shift from manufacturing to software offerings (Ecorys, 2021). In the CE, DDBMs are already established, but their use is still limited (Circular Republic et al., 2025; Danish Business Authority & EMF, 2021). Currently, a small percentage of circular start-ups use data and DT as the basis for their BM. For example, the report Circular Economy Start-up Landscape 2025 identifies that only 109 European circular start-ups out of 2,525 use artificial intelligence (AI) and data analytics for resource efficiency. The report also found that circular start-ups are represented in the battery, energy, and vehicles and mobility sectors (Circular Republic et al., 2025). Therefore, there is a research gap regarding the potential of circular DDBMs in the automotive sector. Against this background, this study aims to understand DDBMs that support CE implementation for the automotive industry, leading to the research question: **What is the current state of DDBMs for enabling the CE for the automotive industry?**

To answer this question, we created a taxonomy that provides a structured overview of the current state of DDBMs supporting CE implementation in the automotive sector. The paper is structured as follows: We first outline the state of the art in DDBMs and CE, followed by the research method in chapter 3. Chapter 4 presents the taxonomy results, and Chapter 5 evaluates them using three real-world objects. The final chapter concludes with key findings, research opportunities, and limitations.

2 State of the Art

2.1 Business Models

The increasing interest in BMs began in the 90s with the onset of digitalization (Zott et al., 2011), and in recent years, researchers have examined various types of BMs (e.g., Möller et al., 2020). In general, BMs describe how a company operates and organizes itself to create value for its customers, as well as how financial changes occur (Osterwalder & Pigneur, 2010). DDBMs are a subterm of digital BMs that specifically focus on data as the key resource (Engelbrecht et al., 2016; Hartmann et al., 2016). Therefore, companies offer raw data or analyzed results for sale or as a service. Digital BMs refer to digital offerings such as digital services or products (Möller et al., 2019).

To optimize the development and description of BMs, tools, and methods support the process by defining the differences and dynamics (Al-Debei & Avison, 2010). Often cited examples include the Business Model Canvas by Osterwalder and Pigneur (2010) and the methodology by Gassmann et al. (2013). However, these are too superficial to develop DDBMS (Engelbrecht et al., 2016). Therefore, further tools and methods were developed, such as the VISOR framework by El Sawy and Pereira (2013).

DDBMs are present in different areas, i.e., in logistics (e.g, Möller et al., 2020), manufacturing (e.g., Stahl et al., 2023), or mobility (e.g., Sterck et al., 2022). Generally, they help companies with decision-making and strategic planning. In the context of sustainability and CE, DBBMs are specialized, particularly in insights into various data to close the circle. For example, DDBMs are used for circular contexts by tracking material flows to identify potential CO_2 hotspots and adjust accordingly

(Luoma et al., 2021). In addition, companies can upgrade their old products with sensors to monitor their condition and implement circular strategies such as »repair« (Ghoreishi, 2023; Ingemarsdotter et al., 2020). Another example of a circular BM based on data or digital components is the dematerialization. Digital components replace physical products to reduce material and to achieve resource efficiency (Geissdoerfer et al., 2018; One Planet Network, 2025).

2.2 Circular Economy Concept

The concept of a CE is widely accepted. In recent years, the awareness, interest, and research on CE have increased because its application is a new phenomenon (Rosa et al., 2019), although the original idea was developed in the 1960s (Boulding, 1966). The main idea of this concept is that global systems become sustainable (ecological, economic, and social) by keeping products and materials in use as long as possible throughout their life cycle (EMF, 2013; Kirchherr et al., 2023). Further advantages are new job creations and growing revenues (OECD, 2025). The CE concept is operationalized through different circular strategies. The 10R framework by Potting et al. (2017) describes various opportunities, such as Refuse, Reduce, Repair, Remanufacturing, or Recycle, to implement circular strategies. Consequently, the EMF also developed a framework, ReSOLVE, to describe the implementation steps for the circular transition (EMF, 2015), although it primarily focuses on BM elements. In the literature, taxonomies and typologies of CBMs exist to describe their patterns and current characteristics(e.g., Lüdeke-Freund et al., 2019; Rosa et al., 2019). However, these studies primarily focus on the technical implementation of CE and largely omit the role of DDBMs as an integral component of CBMs.

DTs and data are key enablers of CBMs, as they support both BM innovation and the implementation of circular strategies (Rosa et al., 2020). However, research remains fragmented, particularly in the automotive sector. Existing studies either focus on specific data types of BMs (e.g., Luoma et al., 2021) or isolated DT use cases of CBMs (e.g., Ingemarsdotter et al., 2020). A holistic perspective on DDBMs is still missing, revealing a clear research gap.

3 Taxonomy Design and Approach

Taxonomies classify objects into dimensions and characteristics, enabling systematic comparison and analysis, and are widely used in BM research to describe and structure BMs, thereby supporting the understanding of complex landscapes (Möller et al., 2022; Szopinski et al., 2019).

This study applies the taxonomy development approach by Kundisch et al. (2022), building on Nickerson et al. (2013), comprising six phases from problem identification to communication.

First, the target users and problem are defined, as they shape the meta-characteristic and subsequent development steps. In our case, we identified two target users of the taxonomy. The taxonomy targets researchers and practitioners interested in a deeper understanding of DDBMs in the circular automotive sector and in innovating new BMs. It supports both the structuring of these emerging BM types and the development of new DDBMs for circularity by serving as an inspiration source, particularly when future work derives patterns from these models (Möller et al., 2022). It thereby responds to the identified gap between the high potential of data for enabling CE and its limited exploitation in the automotive industry (Urbinati et al., 2021). Accordingly, the meta-characteristic is defined as »elements of DDBMs for the circular automotive sector«.

Taxonomy development followed an iterative process that combined conceptual-to-empirical and empirical-to-conceptual approaches. The empirical sample concluded 350 companies identified through *StatrUp Autobahn*[1] and *Catena-X* [2](November 2023 – January 2024), of which 58 underwent detailed analysis using publicly available sources such as websites, articles, videos, and news reports. Eligible companies operated data-driven and CE-oriented BMs in the automotive sector. The sample predominantly consisted of start-ups, which were chosen as the primary database due to their innovation-oriented nature (Möller et al., 2022). Due to insufficient information, we excluded 16 companies, leaving 42 for taxonomy development. Based on the VISOR framework, we conduct four iterations (1. conceptual-to-empirical; 2.-4. empirical-to-conceptual) to refine dimensions and characteristics,

[1] https://startup-autobahn.com/
[2] https://catena-x.net/

introduce a CE-specific meta-dimension, and improve conceptual fit. The process starts inductively and is later complemented by empirical analysis of the selected objects.

After the iterations, all objective and subjective stopping conditions according to Nickerson et al. (2013) are met, resulting in a final taxonomy comprising 16 dimensions and 68 characteristics (cf. Table 1). These conditions ensure completeness (objective) and quality (subjective) of the taxonomy development process (Szopinski et al., 2019). To further validate usefulness, we applied three real-world objects as it is suggested in the study by Szopinski et al. (2019). The remaining phases focus on evaluating completeness, demonstrating applicability, and communicating the taxonomy.

Table 3: Business model taxonomy for circular transformation for the automotive sector

Meta-Dim.	Dimension	Characteristics						
Value Proposition	Service Offering	End-to-End Solutions	Digital Marketplace	Analytics and Montoring	Management Tasks	Process Digitalization	Digital Object / Process Mapping	Database
	Purpose	Transparency	Optimization	Collaboration	Traceability	Compliance	Efficiency Improvement	
	Actors Group	Automobile Manufacturer / OEM	Commercial Vehicle Manufacturer	Battery Manufacturer	Automotive Suppliers	Car Dismantler / Recycler	Industry Neutral	
Interface	Customer Interface	Web-based Platform		Mobile App		Plugin App		
	Data Integration Interface	GUI		API		(IoT) Sensors		
Service Platform	Technology	AI / ML		Cloud Technology		Blockchain		
	Data Ressource	Process Data	Vehicle Data	Battery Data	Sensor/ Machine Data	Supply Chain Data	Material Data	External Data
	Data Source	Internal Source				External Source		
	Data Collection	Manuel		Dynamic		Both		
	Data Activity	Data Generation	Data Acquisition	Data Aggregation	Data Processing	Data Analytics	Data Visualization	Data Distribution
Organizational Model	Partnership Network	Ecosystem	Technology Company	Consultancy Company	Association	University / Research Institution	Other	
	Integration of Other Actors	Supplier	Authority	Customer	Partner	None		
Revenue Model	Offer Options	Subscription		Asset Sale		Freemium		
	Price Structure	Fixed Prices			Variable Prices			
Circular Economy	Design Option	Share		Optimize		Loop	Virtualize	
	Focus Object	(Critical) Raw Material		Vehicle Component		Machinery & Production Plants	Value Chain	

Legend: • TWAICE (Use Case 1) • material.one (Use Case 2) • i.Point (Use Case 3)

Source: Own

4 Results

This chapter presents the characteristics and dimensions resulting from taxonomy development. Some of these dimensions and characteristics are defined for DDBMs, such as the dimensions and characteristics from the »Interface« and »Service Platform«. The data components »Data Resource«, »Data Source, »Data Activity«

and »Data Integration Interface« are BM characteristics based on generic DDBMs. The representative characteristics of the circular DDBMs for the automotive sector become clear in the meta dimension »Value Proposition« and especially in the sixth meta-dimension »Circular Economy«. These elements highlight the distinguishing features of BMs compared to generic DDBMs and among themselves. The variety of characteristics demonstrates the uniqueness of circular DDBMs.

4.1 Meta-Dimension: Value Proposition

The meta-dimension *Value Proposition* comprises the dimensions *Service Offering*, *Purpose*, and *Customer Group* with 19 characteristics in total.

1st Dimension: Service Offering

This dimension describes the functionalities offered by circular DDBMs. Identified characteristics include: *End-to-End Solution* (integration of cross-departmental processes), *Digital Marketplace* (digital interaction platform for suppliers and customers), *Analysis and Monitoring* (e.g., emissions calculation, predictive maintenance), *Management Tasks* (decision and communication support), *Process Digitalization* (replacement of manual documentation by digital workflows), *Object Mapping/Digital Process* (digital representation of objects and processes, e.g., digital product passport), and *Database* (data collection and provision).

2nd Dimension: Purpose

This dimension captures the problems addressed by the offering: *Transparency* (e.g., identifying CO_2 hotspots), *Optimization* (product and process improvement through analytics), *Collaboration* (coordination across actors), *Traceability* (tracking products, materials, suppliers), Compliance (legal and ESG requirements), and *Efficiency Improvement* (operational and cost optimization).

3rd Dimension: Customer Group

Target groups within the automotive ecosystem include *OEMs* (e.g., BMW, or Audi)**,** *Commercial Vehicle Manufacturers*, *Battery Manufacturers*, *Automotive Suppliers* (e.g., n-tier suppliers in value chain), *Car Dismantlers/Recyclers*, and *Industry-Neutral Companies* (e.g., other industries such as pharmaceutical or railway sectors).

4.2 Meta-Dimension: Interface

The Interface meta-dimension deals with the digital and physical interfaces between the provider and the customer. In the context of circular DDBMs, the interface serves as a technological component that facilitates transactions (Möller et al., 2020).

4th Dimension: Delivery Mechanism

Three delivery modes were identified: Web-based *Platform*, *Mobile App*, and *Plugin* (e.g., SAP integration).

5th Dimension: Data Integration Interface

Data can be integrated through *Application Programming Interface* (*API)* or *Graphical User Interface* (GUI), Manual Entry (e.g., questionnaires, Excel uploads), or *Internet of Things (IoT) sensors* for real-time machine and component data acquisition.

4.3 Meta-Dimension: Service Platform

The *Service Platform* focuses on technological and data-related components and includes five dimensions.

6th Dimension: Technology

Relevant enabling technologies are *AI / Machine Learning (ML)* (e.g., predictive analytics, data cleansing), *Cloud Technology* (centralized storage and accessibility), and *Blockchain* (traceability and data reliability).

7th Dimension: Data Resource

Primary resources include *Process Data*, *Vehicle Data*, *Battery Data*, *Supply Chain Data*, and *Material Data*; *External Data* serves as secondary data.

8th Dimension: Data Source

Data originates from *Internal Sources* (e.g., Product Lifecycle Management (PLM), Enterprise Resource Planning (ERP), and Customer Relationship Management) or *External Sources* (e.g., supplier databases, ecoinvent, the International Material Data System (IMDS)). Many of these reference databases serve as a basis for secondary data analysis.

9th Dimension: Data Collection

Companies enable *Manual Upload*, *Automated Collection*, or *Both*.

10th Dimension: Data Activity

Seven activities were identified, based on Hartmann et al. (2016) taxonomy: *Data Generation* (data creation by sensors attached to machines and products), *Data Acquisition* (gather external information, e.g., identify supply chain disruptions), *Data Aggregation* (providing customer with, i.e., life cycle information), *Data Processing* (Translating data into one unified language), *Data Analytics* (data calculations, as well as predictive analytics), *Data Visualization* (dashboard reports, i.e., product carbon footprint), and *Data Distribution* (distribution of information to another stakeholder).

4.4 Meta-Dimension: Organizational Model

This meta-dimension describes actor relationships and collaboration structures within circular data ecosystems.

11th Dimension: Partnership Network

Partnerships support knowledge exchange and ecosystem development. Six partner types were identified: *Technology Companies*, *Consulting Companies*, *Associations*, *Universities/Research Institutions*, *Ecosystems* (e.g., Catena-X), and *Others*.

12th Dimension: Integration of Other Actors

Platforms integrate *Suppliers/Vendors*, *Authorities*, *Customers*, *Partners*, or *No Additional Actor Group*.

4.5 Revenue Model

This dimension is divided into two dimensions, with the characteristics referring to the definitions provided by Hartmann et al. (2016) and Bock and Wiener (2017).

13th Dimension: Revenue Model

Three monetization logics were identified: *Subscription fee* (predetermined schedule), *Asset Sale* (fixed price, i.e., hardware or data generation), and *Freemium* (can be used by customers free of charge or at low cost based on user volume).

14th Dimension: Pricing Structure

Offerings follow either *Fixed Pricing* or *Variable Pricing*, depending on service scope and customization.

4.6 Meta-Dimension: Circular Economy

This meta-dimension was added during the second iteration to emphasize the importance of the CE. It is differentiated into two dimensions: *Design Options* and *Focus Object*.

15th Dimension: Design Options

This dimension refers to possible design strategies supporting CE principles and BMs. Four of six ReSOLVE options according to EMF (2015) are met: *Share* (reuse, maintenance, secondary markets), *Optimize* (waste and CO_2 reduction through data-driven improvements), *Loop* (recycling and material circulation), and *Virtualize* (replacement of physical processes through digital solutions).

16th Dimension: Focus Object

Circular initiatives target *(Critical) Raw Materials* (e.g., raw material for the production of electric battery), *Vehicle Components* (e.g., electric battery or mechanical component), *Machinery and Production Plants*, or the broader *Value Chain* (supply chain or reverse logistics).

5 Evaluation

Following of Kundisch et al. (2022) and Szopinski et al. (2019), we evaluate the taxonomy's utility through three automotive real-world use objects. By decomposing DDBMs into their constituent components, the application provides a modular framework for understanding CE implementation. Additionally, we chose these three objects to show the variety of circular DDBMs for the automotive sector. While the qualitative analysis confirms the taxonomy's descriptive power, a detailed assessment of revenue streams and pricing structures is limited by corporate non-disclosure, despite the observed prevalence of subscription-based models. Overall, the application supports researchers and practitioners in understanding the variety of DDBMs in the circular automotive sector and serves as an inspiration source for innovating new BMs. Additionally, we confirm our evaluation goal of systematically describing the current state of these BMs and identifying their specific, circular, data-driven characteristics.

1. Case Example: TWAICE[3]

TWAICE offers a software-as-a-service (SaaS) platform that optimizes electronic battery health performance and reduces risks by analyzing and monitoring battery status. The software provides continuous, reliable information about the battery's status, eliminating the need for costly, time-consuming physical tests. Through data-driven analysis and simulation, TWAICE supports optimized battery design and improved performance, as well as transparent battery health assessment. Improved efficiency is achieved by reducing development, production, and operating costs. TWAICE's main target customers are battery manufacturers and OEMs, who benefit from faster development cycles, improved cell qualification, and a shared understanding of battery aging. The solution leverages cloud and ML technologies, integrating data from battery management systems and other customer databases via API interfaces for real-time monitoring. Key data activities include collection, processing, analysis, and dashboard visualization. The results of continuous analysis and monitoring are demonstrated in Key Performance Indicators. TWAICE's offerings support CE goals by enabling improved battery design, extended battery lifetime, and battery reuse through battery certification, addressing the Loop and Share options of the ReSOLVE framework.

2. Case Example: material.one[4]

material.one is a subsidiary of the big German IT consulting company Arvato. material.one is specialized for ensuring sustainable compliance of materials and supply chains by simplifying the implementation of quality and environmental requirements of products and supply chains, such as verification of recycled materials or CO_2 emissions. The category *Process digitalization* refers to digitizing and networking business processes related to these requirements.

A digital platform functions as a centralized digital database in which all product and material data are stored and is accessible anytime and anywhere to exchange the data with suppliers, for example. material.one's platform and data exchange, and continuous, data- and information-based exchange between the involved actors, enable transparency and collaboration that are important to achieving sustainability.

[3] https://www.twaice.com/
[4] https://material.one/

material.one's function of digitized and automatically executed business processes improved efficiency by reducing the manual effort of these processes.

The platform fosters collaboration among manufacturers, laboratories, and suppliers by connecting these actors within one system. Typical customers include OEMs such as BMW and Mercedes-Benz, as well as automotive suppliers, while the solution itself is industry-neutral. The SaaS platform uses cloud technology, processes mainly material and process data from ERP, PLM, or Quality Management systems, and standardizes them into comparable proofs. Data analysis is primarily descriptive, with automated comparisons of requirements and test results that are visually prepared and shared. The platform integrates suppliers and partners, cooperates with ecosystems like Catena-X, and relies on Microsoft Azure for secure cloud infrastructure. By digitally managing compliance proofs such as recycled content, CO_2 footprint, and hazardous substances, material.one supports CE implementation, with the value chain as the primary focus object and a strong link to product and material data across the lifecycle.

3. Case Example: IPOINT[5]

The case example IPOINT offers a comprehensive portfolio of software solutions that can be classified as an *End-to-End Solution* for implementing the CE across the entire product life cycle. Its offerings cover digital product passports, management solutions, and analysis and monitoring tools, enabling CO_2 footprint calculation, ESG-compliant reporting, risk management, and strategic consulting. By analyzing material flows and value chains, IPOINT promotes transparency, traceability, and compliance, helping companies identify CO_2 hotspots and supply chain risks. The customer target group is industry-neutral but also includes OEMs such as BMW and Audi, as well as tier-1 suppliers. The solutions rely on cloud technology for data collection, processing, storage, and visualization, supported by AI/ML for automated environmental calculations and blockchain for the digital product passport. Data from internal systems (ERP, PLM) and external databases such as IMDS and ecoinvent are integrated via APIs. IPOINT operates on a subscription-based SaaS model complemented by consulting and training services. Collaboration with suppliers, partners, and ecosystems like Catena-X supports CE implementation.

[5] https://www.ipoint-systems.com/

Overall, IPOINT addresses the EMF design options *Optimize* and *Virtualize* by enabling data-driven eco-design, life cycle assessment, and value chain optimization focused on sustainability and circularity.

Overall, these three real-world objects demonstrate the complexity of circular implementation barriers in the automotive industry. TWAICE's use case is limited to a new type of vehicle: electric vehicles. In this case, a significant degree of competition characterizes the European automotive industry, which fosters the development of vehicles comparable to those produced by competitors (Ecorys, 2021). Electric battery development is still in the early stages, leading to outages and unprofitability (Plotnikov et al., 2023). Therefore, TWAICE developed a BM to overcome barriers in the development process and drive innovation. They enable more insights, reaching safety and a higher lifetime of electric batteries through their use of data and digital technologies. Consequently, material.one indirectly influences the circular economy in this sector. A major challenge to successfully implementing circular strategies is the lack of transparency and information. However, supply chain documentation is increasing because legal requirements are becoming more stringent. Consequently, stakeholders must confirm and share important documentation and guidelines. Compliance with CO_2 emissions and recycling material reporting requirements is especially important. Their platform enables contactless and manual data and information exchange. By connecting digital technologies and reducing manual processes through digitalization, they enable transparent information exchange among supply chain partners. The third real-world object, IPOINT, is different from the others. Originally an IT material compliance company, they have recently expanded to sustainability. They enable material compliance, as does material.one, but they also calculate CO_2 emissions for entire supply chains or specialized products. With its wide range of solutions, IPOINT software is a significant step forward in solving individual circular barriers and challenges for its customers.

6 Conclusion, Limitations, and Outlook

This study presents a taxonomy of 42 empirical objects of DDBMs in the automotive industry, demonstrating the current variety of circular DDBMs in the sector. The resulting taxonomy comprises six meta-dimensions, 16 dimensions, and 68 characteristics, capturing service offerings, value propositions, stakeholder

groups, data resources, and applied technologies. The analysis shows that companies frequently use DTs such as cloud technology and AI to deliver CE–related value propositions, mainly through data analytics, environmental calculations, and data provision.

From a scientific perspective, the findings highlight the positive role of data and DTs in promoting CE and sustainability, while also revealing significant untapped potential due to the still limited variety of DDBMs. For practitioners, the taxonomy provides a structured approach to analyzing, adapting, and innovating BMs to close or extend the life cycles of materials and products in the automotive sector. Limitations include the restricted sample size, the focus on two automotive platforms, and the subjective interpretation of characteristics. Future research should expand the empirical basis, examine barriers to DDBM adoption, and further explore how data and digitalization can accelerate the transition toward a circular automotive industry.

References

ACEA. (2018). *Circular Economy*. https://www.acea.auto/fact/circular-economy/

Al-Debei, M. M., & Avison, D. (2010). Developing a Unified Framework of the Business Model Concept. *European Journal of Information Systems*, *19(3)*, 359–376. doi:10.1057/ejis.2010.21

Bock, M., & Wiener, M. (2017). Towards a Taxonomy of Digital Business Models – Conceptual Dimensions and Empirical Illustrations. In *38th International Conference on Information Systems (ICIS),* Seoul, Südkorea.

Boulding, K. E. (1966). The economics of the coming spaceship earth. *Environmental Quality in a Growing Economy*, 3–14. doi:10.4324/9781315064147-2

Circular Republic, UnternehmerTUM, betterventures, & TUM VentureLabs. (2025). *Circular Economy Start-up Landscape 2025: Development and Trends in the European Market*. https://www.circular-republic.org/startup-landscape-report-2025/

Danish Business Authority, & EMF. (2021). *Looping on data: Best practices and barriers for sharing data in circular business models*. https://content.ellenmacarthurfoundation.org/m/b9dda900dd7c171c/original/Looping-on-data-Best-practices-and-barriers-for-sharing-data-in-circular-business-models.pdf

Ecorys. (2021). *The Future of the EU Automotive Sector: Publication for the Committee on Industry, Research and Energy, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament*. https://www.europarl.europa.eu/RegData/etudes/STUD/2021/695457/IPOL_STU(2021)695457_EN.pdf

El Sawy, O. A., & Pereira, F. (2013). VISOR: A Unified Framework for Business Modeling in the Evolving Digital Space. In O. A. El Sawy & F. Pereira (Eds.), *SpringerBriefs in Digital Spaces. Business Modelling in the Dynamic Digital Space* (pp. 21–35). Springer Berlin Heidelberg. doi:10.1007/978-3-642-31765-1_3

EMF. (2013). *Towards the Circular Economy Vol. 1: An Economic and Business Rationale for an Accelerated Transition*. https://www.ellenmacarthurfoundation.org/towards-the-circular-economy-vol-1-an-economic-and-business-rationale-for-an

EMF. (2015). *Growth within: A Circular Economy Vision for a Competive Europe.* https://www.ellenmacarthurfoundation.org/growth-within-a-circular-economy-vision-for-a-competitive-europe

Engelbrecht, A., Gerlach, J., & Widjaja, T. (2016). Understanding the Anatomy of Data-Driven Business Models - Towards an Empirical Taxonomy. In *Proceedings of the 24th European Conference on Information Systems (ECIS),* Istanbul, Türkei. https://core.ac.uk/download/pdf/301369681.pdf

Gassmann, O., Csik, M., & Frankenberger, K. (2013). *Geschäftsmodelle entwickeln: 55 innovative Konzepte mit dem St. Galler Business Model Navigator*. Hanser.

Geissdoerfer, M., Vladimirova, D., & Evans, S. (2018). Sustainable Business Model Innovation: A Review. *Journal of Cleaner Production*, *198*, 401–416. doi:10.1016/j.jclepro.2018.06.240

Ghoreishi, M. (2023). The Role of Digital Technologies in a Data-driven Circular Business Model. *Journal of Business Models*, *11*(*1*), 78–81. doi:10.54337/jbm.v11i1.7245

Grimm, A., & Pfass, M. (2022). *Transformation der Wertschöpfung in der Automobilbranche: Teilbericht Arbeitspaket 4: Entwicklungen im Regime* (Working Paper Forschungsförderung No. 249). Hans-Böckler-Stiftung.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing Value from Big Data – A Taxonomy of Data-Driven Business Models Used by Start-Up Firms. *International Journal of Operations & Production Management*, *36*(*10)*, 1382–1406. doi:10.1108/IJOPM-02-2014-0098

Hoppe, C., Schoormann, T., Winkelmann, S., & Möller, F. (2025). Tensions in implementing a circular economy – Empirical insights from the automotive industry. *Computers & Industrial Engineering*, *204*, 111090. doi:10.1016/j.cie.2025.111090

Ingemarsdotter, E., Jamsin, E., & Balkenende, R. (2020). Opportunities and Challenges in IoT-enabled Circular Business Model Implementation: A Case Study. *Resources, Conservation and Recycling*, *162*, 105047. doi:10.1016/j.resconrec.2020.105047

Kirchherr, J., Yang, N.-H. N., Schulze-Spüntrup, F., Heerink, M. J., & Hartley, K. (2023). Conceptualizing the Circular Economy (Revisited): An Analysis of 221 Definitions. *Resources, Conservation and Recycling*, *194*, 107001. doi:10.1016/j.resconrec.2023.107001

Kundisch, D., Muntermann, J., Oberländer, A. M., Rau, D., Röglinger, M., Schoormann, T., & Szopinski, D. (2022). An Update for Taxonomy Designers. *Business & Information Systems Engineering*, *64*(*4)*, 421–439. doi:10.1007/s12599-021-00723-x

Lüdeke-Freund, F., Gold, S., & Bocken, N. M. P. (2019). A Review and Typology of Circular Economy Business Model Patterns. *Journal of Industrial Ecology*, *23*(*1)*, 36–61. doi:10.1111/jiec.12763

Luoma, P., Toppinen, A., & Penttinen, E. (2021). Role and Value of Data in Circular Business Models – a Systematic Literature Review. *Journal of Business Models*, *9*(*2)*, 44–71. doi:10.5278/jbm.v9i2.3448

Möller, F., Bauhaus, H., Hoffmann, C., Niess, C., & Otto, B. (2019). Archetypes of Digital Business Models in Logistics Start-Ups. In *Proceedings of the 27th European Conference on Information Systems*.

Möller, F., Stachon, M., Azkan, C., Schoormann, T., & Otto, B. (2022). Designing Business Model Taxonomies – Synthesis and Guidance from Information Systems Research. *Electron Markets*, *32*(*2)*, 701–726. doi:10.1007/s12525-021-00507-x

Möller, F., Stachon, M., Hoffmann, C., Bauhaus, H., & Otto, B. (2020). Data-Driven Business Models in Logistics: A Taxonomy of Optimization and Visibility Services. In *Proceedings of the 53rd Hawaii International Conference on Systems Sciences (HICSS)* (pp. 5379–5388). ScholarSpace.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A Method for Taxonomy Development and its Application in Information Systems. *European Journal of Information Systems*, *22*(*3)*, 336–359. doi:10.1057/ejis.2012.26

OECD. (2025). *The Circular Economy in Cities and Regions of the European Union: OECD Urban Studies*. doi:10.1787/e09c21e2-en

One Planet Network. (2025). *Digitization of Circular Business Models: Digital Solutions Enabling the Transition towards Circular Economy.*

Osterwalder, A., & Pigneur, Y. (2010). *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. Hoboken.

Pichler, M., Krenmayr, N., Schneider, E., & Brand, U. (2021). EU Industrial Policy: Between Modernization and Transformation of the Automotive Industry. *Environmental Innovation and Societal Transitions*, *38*, 140–152. doi:10.1016/j.eist.2020.12.002

Plotnikov, M., Schreynemackers, P., Joachimsthaler, C., Jarmer, J.-P., Kuhlmann, G., Karlstedt, F., Kropkowski, L., Kürpick, C., Müller, S., & Schrade, C. (2023). *Batterie-Logistik. Whitepaper*. Fraunhofer-Gesellschaft. doi:10.24406/publica-1247

Rosa, P., Sassanelli, C., & Terzi, S. (2019). Towards Circular Business Models: A systematic literature review on classification frameworks and archetypes. *Journal of Cleaner Production*, *236*, 1–17. doi:10.1016/j.jclepro.2019.117696

Rosa, P., Sassanelli, C., Urbinati, A., Chiaroni, D., & Terzi, S. (2020). Assessing Relations between Circular Economy and Industry 4.0: A Systematic Literature Review. *International Journal of Production Research*, *58(6)*, 1662–1687. doi:10.1080/00207543.2019.1680896

Stahl, B., Häckel, B., Leuthe, D., & Ritter, C. (2023). Data or Business First?—Manufacturers' Transformation Toward Data-driven Business Models. *Schmalenbach Journal of Business Research*, *75(3)*, 303–343. doi:10.1007/s41471-023-00154-2

Sterck, F., Peukert, C., Hunke, F., & Weinhardt, C. (2022). Understanding Car Data Monetization: A Taxonomy of Data-Driven Business Models in the Connected Car Domain. *Wirtschaftsinformatik 2022 Proceedings*. https://aisel.aisnet.org/wi2022/digital_business_models/digital_business_models/7

Szopinski, D., Schoormann, T., & Kundisch, D. (2019). Because your Taxonomy Is Worth It: Towards a Framework For Taxonomy Evaluation. In *Proceedings of the 27th European Conference on Information Systems*. https://aisel.aisnet.org/ecis2019/

Urbinati, A., Franzò, S., & Chiaroni, D. (2021). Enablers and Barriers for Circular Business Models: an Empirical Analysis in the Italian Automotive Industry. *Sustainable Production and Consumption*, *27*, 551–566. doi:10.1016/j.spc.2021.01.022

Zott, C., Amit, R., & Massa, L. (2011). The Business Model: Recent Developments and Future Research. *Journal of Management*, *37(4)*, 1019–1042. doi:10.1177/0149206311406265

About the authors

Hannah Scheerer is a research assistant at the chair of Industrial Information Management at the TU Dortmund University in Dortmund, Germany. She currently researches circular economy business models and their relation with data as well as the general transformation of incumbents from linear to circular business models. She holds a M. Sc. in Industrial Information Management from the TU Dortmund University.

Hendrik van der Valk is the chief engineer at the chair of Industrial Information Management at the TU Dortmund University in Dortmund, Germany. He also works as a part-time job as scientific coordinator at the Fraunhofer Institute for Software and Systems Engineering, Dortmund. He holds a PhD in Mechanical Engineering from TU Dortmund University

Towards Pedagogically Structured Human–AI Engagement in a Peer Learning Context

Minna Silvennoinen, Sanna Paananen,
Sami Määttä, Satu Aksovaara

Jamk University of Applied Sciences Jyväskylä, School of Professional Teacher Education, Jyväskylä, Finland
minna.silvennoinen@jamk.fi, sanna.paananen@jamk.fi, sami.maatta@gmail.com, satu.aksovaara@jamk.fi

Generative AI is increasingly integrated into higher education, yet prior research has mainly examined AI use at the individual level, with less attention on AI engagement variation across learning tasks, AI roles, and collaborative contexts. This study examined perceived AI usability across a pedagogically structured peer-learning process where generative AI supported different, individual or collaborative learning activities. The context was a compulsory undergraduate course where student teams worked on authentic projects. Responses from 411 students indicate that generative AI was perceived as most usable during collaborative content structuring and least usable during individual reflection. Students' perceived study ability showed small but consistent associations with perceived AI usability. Similar patterns emerged within teams, suggesting that ways of engaging with generative AI may partly develop as shared practices with study-related self-regulation potentially playing a role. Findings highlight the importance of examining AI engagement as a phenomenon shaped by pedagogical task design, collaborative learning, and the roles assigned to generative AI in learning activities.

DOI
https://doi.org/10.18690/um.fov.4.2026.34

ISBN
978-961-299-152-4

Keywords:
generative AI, human–AI collaboration, peer learning, higher education, pedagogical design

University of Maribor Press

1 Introduction

As generative AI (GenAI) becomes an increasingly embedded element of everyday academic practices in higher education (HE) (Carmi, 2025), attention is focusing on how such use becomes pedagogically meaningful. Students differ in how they engage with new technologies (Zou Y. et al., 2025), and GenAI use is known to range from shallow copy-and-paste practices to more reflective interaction (Chen et al., 2026), Yet, little is known about how such use varies across task types, AI roles, and social contexts (Belkina et al., 2025; Qian, 2025). This raise questions about how structured tasks could support intentional human–AI collaboration, highlighting the need for more structured pedagogical approaches that support deeper and more reflective engagement with GenAI in higher-order learning processes (Qian, 2025).

Previous research has largely examined students' use of GenAI at the individual level (Ruiz-Rojas et al., 2024; Salinas-Navarro et al., 2024). However, in collaborative learning contexts, AI engagement may also be shaped by team characteristics and interaction practices. Emerging studies suggest that the roles GenAI takes in collaborative activities can vary depending on team dynamics and users' technological expertise (e.g., Law et al., 2025; Anderson et al., 2025). Nevertheless, relatively little research has examined how pedagogically structured AI roles within learning tasks relate to shared patterns of AI engagement.

Understanding task-specific patterns of human–AI collaboration requires attention to learner-related psychological factors, such as perceived study ability, which may influence how students integrate GenAI into their learning activities (Atchley et al., 2024). Examining these factors may explain why some students engage more actively with GenAI while others make more limited use of it, particularly in tasks requiring analytical framing or reflection.

Universities of applied sciences (UAS) face a growing need to prepare students for AI engagement in professional practice. This human–AI collaboration — the ability to work effectively in teams composed of both humans and AI systems (Atchley et al., 2024) becomes increasingly important for future employability (e.g., Irish et al., 2025; Portocarrero Ramos et al., 2025). Recent research has highlighted the importance of integrating GenAI into collaborative, project-based learning environments that support knowledge co-creation and student agency in

orchestrating interactions between humans and AI (e.g., Law et al. 2025). This raises questions relating to pedagogical structuring of human-AI collaboration within HE (McPhee & Jerowsky, 2025) and further highlights the importance of examining AI engagement as a phenomenon shaped by pedagogical design, social contexts, and learner-related characteristics.

This study investigates how students report GenAI usability across phases of a pedagogically structured peer-learning process in which GenAI takes different supportive or facilitative roles. Research questions are:

RQ1: How do students' perceived AI usability differ across individual and collaborative learning tasks where GenAI serves different roles?

RQ2: How is students' study ability associated with their perceived AI usability across varying GenAI roles?

RQ3: To what extent does students' perceived AI usability with varying GenAI roles show similarities within project teams?

This study contributes to the emerging literature on AI in HE by moving beyond individual use to examine perceived AI usability and patterns of AI use within peer-learning contexts across task types and pedagogically structured uses of GenAI. The findings provide empirical insight into how task design and social context may shape AI use, contributing to understanding how pedagogical structuring can support meaningful human–AI collaboration in HE.

2 Background

Previous research has shown that students commonly employ GenAI for idea generation, text production, revision, and conceptual clarification (Krause et al., 2024; Qian, 2025). Yet, although adoption is widespread, research indicates that engagement often remains uneven and tied to isolated uses of GenAI, as well as being characterised by individual trial and error (Almassaad et al., 2024; Belkina et al., 2025; Krause et al., 2024; Huang et al., 2024). Although recent studies have explored the use of GenAI in tasks requiring higher-order cognitive processes, such as problem-solving, and reflection, findings remain mixed and highly dependent on

pedagogical design and task structure, particularly in terms of guided versus unguided use (Guo et al. 2025; Balart et al. 2025). At the same time, students report uncertainty regarding appropriate and responsible uses of GenAI within learning, pointing to continuing inconsistencies in institutional guidance and a lack of pedagogically structured support (e.g. Abdelaal et al., 2025; Balabdaoui et al., 2024).

In response, pedagogical frameworks for AI integration have been proposed. AI literacy models often conceptualize AI-related competences through hierarchical learning frameworks, such as Bloom's taxonomy (e.g., Zhou & Schofield, 2024) or extend existing digital competence traditions, such as TPACK or DigCompEdu, to the context of GenAI (e.g., Zou D. et al., 2025). However, much of this work focuses on defining competencies or guiding teachers' integration of AI.

Although human–AI collaboration involving GenAI has received increasing attention in recent research (e.g., Jung & Jin, 2025; Jia et al. 2026), less is known about how GenAI use unfolds within pedagogically structured peer-learning contexts. Research conceptualizes peer learning as a socially shared process in which knowledge is co-constructed through interaction, negotiation, and shared regulation among learners (Järvelä et al., 2023). GenAI use in learning may support collaborative processes, while its role and use within these contexts may also be shaped by group-level dynamics and shared practices, since the use is not solely an individual decision (e.g. Sichterman et al., 2025; Pham et al., 2025; Feng, 2025). These team-level dynamics may lead to shared patterns of AI engagement and mirror emerging forms of human–AI collaboration within teamwork. Peer-learning contexts provide a natural setting for introducing new technologies into collaborative work, where students can explore and negotiate GenAI use while working toward shared outcomes.

While pedagogical structuring shapes how GenAI is embedded within learning tasks, students' engagement with these structures may also depend on individual learner-related psychological factors. These factors may play an important role in shaping how students engage with GenAI in academic contexts. Students differ in their psychological readiness to adopt and integrate new technologies, which make perceived study ability particularly relevant for understanding patterns of GenAI use (Acosta-Enriquez et al., 2024; Hadar Shoval, 2025). Perceived study ability can be understood in relation to self-regulated learning, which involves planning,

monitoring, and managing one's learning (Zimmerman, 2002; Schraw & Dennison, 1994). Students with stronger self-regulatory capacities may be better able to integrate GenAI into their learning, whereas lower perceived study ability may be associated with more limited use. In peer learning contexts, these individual characteristics may also interact with team-level dynamics, potentially contributing to shared patterns of GenAI use.

3 Research Design and Methodology

3.1 Context, Participants and Research Design

This study examined students' perceived AI usability in a course integrating GenAI. The study was guided by the **AI Assistant Framework**, in which GenAI use was pedagogically orchestrated through learning task instructions positioning GenAI in roles presented in Table 1. The approach draws on principles of design-based research, where pedagogical interventions are embedded in authentic learning contexts and examined in practice (Wang & Hannafin, 2005).

The study was conducted in a compulsory 10 ECTS undergraduate course organised around peer-learning principles, where students worked in multidisciplinary teams on authentic working-life projects. The course was implemented in two consecutive years (2024–2025) using the same pedagogical design, and the data from both cohorts were combined for analysis.

A total of 544 students (79.6%) consented to participate in the study, of whom 411 answered all perceived AI usability items. Of the participants, 49.8% were male and 46.5% female (3.7% missing). Students represented degree programmes in business, health and wellbeing, and technology. They worked in teams of four to five members (M = 4.64), forming 149 project teams.

GenAI was integrated into five phases of the learning process (P1-P5), with task instructions guiding how students engaged with GenAI in each project phase. The tasks differed in both their social configuration (individual vs. team) and their functional role in the learning process, see Table 1.

Table 1: Pedagogically structured AI-supported tasks across project phases

Learning phase	Learning task	GenAI role	Mode
P1 Ideation	Client meeting	Ideation support	Individual
P2 Analysis	Competence identification	Skill analysis support	Team
P3 Problem solving	Teamwork challenges and process reflection	Teamwork coaching	Team
P4 Production	Project presentation	Content structuring	Team
P5 Reflection	Reflection	Reflection support	Individual

The design enabled comparison of perceived AI usability across phases, analysis of its association with study ability and team-level similarity, and comparisons within students across tasks.

3.2 Measures

Perceived AI usability was measured across the five phases of the AI Assistant Framework (P1–P5) using two self-report items per phase on a five-point Likert scale. The items assessed perceived AI competence ("I was able to use AI in finding a solution/ completing the task") and perceived ease of use ("Using AI in the task was easy"). For each phase, a composite perceived AI usability score was calculated as the mean of the two items. Internal consistency for the two-item scales was acceptable across phases (Table 2).

Study ability was measured using a scale derived from the HowULearn instrument (Parpala & Lindblom-Ylänne, 2012), based on the Metacognitive Awareness Inventory (Schraw & Dennison, 1994). The scale includes two dimensions—knowledge about cognition and regulation of cognition—measured on a five-point Likert scale. Cronbach's α for the scale was .88.

3.3 Analytical Strategy

Patterns of **perceived AI usability** across the five phases (P1–P5) were examined using descriptive statistics, and phase differences were tested with a repeated-measures ANOVA. To examine associations between students' **study ability** and perceived AI usability, Pearson correlations (two-tailed) were calculated between the study ability scale and each of the five perceived AI usability scores P1–P5. Pairwise deletion was used to maximize the number of available cases.

To assess **team-level similarity** in perceived AI usability (RQ3), **intraclass correlation coefficients** (ICC) were calculated for each phase using a one-way random-effects model with team as the grouping factor. Variance components were obtained using UNIANOVA in SPSS. The actual average team size (k = 4.64) was applied in the ICC formulas. ICC(1) represents the proportion of variance attributable to team membership, whereas ICC(2) indicates the reliability of aggregated team-level means. All analyses were conducted using IBM SPSS Statistics 31, with statistical significance evaluated at $p < .05$. ICC values between .05 and .20 were interpreted as indicating small to moderate clustering.

4 Results

RQ1. Differences in perceived AI usability across learning tasks

Perceived AI usability across the P1–P5 was examined. Descriptive statistics (*M* and *SD*), reliability estimates, and intercorrelations among the task-specific perceived AI usability scores are presented in Table 2.

Table 2: Descriptive statistics and intercorrelations for perceived AI usability

Phase	α	*M*	*SD*	P1	P2	P3	P4	P5
P1 Ideation (Individual) **Ideation support**	0.87	2.68	1.14	—				
P2 Analysis (Team) **Facilitating skill analysis**	0.89	2.86	1.16	0.65	—			
P3 Problem solving (Team) **Teamwork coaching**	0.88	2.77	1.16	0.64	0.80	—		
P4 Production (Team) **Content structuring**	0.88	3.04	1.18	0.49	0.51	0.55	—	
P5 Reflection (Individual) **Reflection support**	0.86	2.49	1.13	0.72	0.67	0.72	0.60	—

Note. M = mean; SD = standard deviation. Cronbach's α is reported for each two-item scale. All correlations are Pearson's r (two-tailed), p < .001.

Internal consistency for the two-item scales was acceptable across tasks (Cronbach's α = .86–.89). Intercorrelations among the perceived AI usability scores were positive and statistically significant (all $p < .001$), ranging from $r = .49$ between P1 and P4 to $r = .80$ between P2 and P3. Results of the repeated-measures ANOVA and Bonferroni-adjusted pairwise comparisons are presented in Table 3.

Table 3: Repeated-measures ANOVA and pairwise comparisons for perceived AI usability

A. Omnibus ANOVA (Greenhouse–Geisser corrected)

Source	df1	df2	F	p	Partial η^2
Learning task	3.387	1388.621	35.717	< .001	.080

Note. Greenhouse–Geisser correction was applied due to violation of sphericity (Mauchly's W = .684, p < .001).

B. Pairwise comparisons (Bonferroni-adjusted)

Comparison	Mean difference	SE	p	95% CI Lower	95% CI Upper
P1 vs P2	-0.180	0.047	.002	-.313	-.047
P1 vs P3	-0.094	0.048	.537	-.230	.043
P1 vs P4	-0.359	0.058	< .001	-.521	-.196
P1 vs P5	0.193	0.042	< .001	.076	.311
P2 vs P3	0.086	0.036	.180	-.016	.189
P2 vs P4	-0.179	0.057	.019	-.340	-.018
P2 vs P5	0.373	0.046	< .001	.244	.503
P3 vs P4	-0.265	0.055	< .001	-.420	-.111
P3 vs P5	0.287	0.042	< .001	.168	.406
P4 vs P5	0.552	0.051	< .001	.409	.696

Note. SE = standard error; P1 = Ideation; P2 = Analysis; P3 = Problem solving; P4 = Production; P5 = Reflection. All p-values are Bonferroni-adjusted.

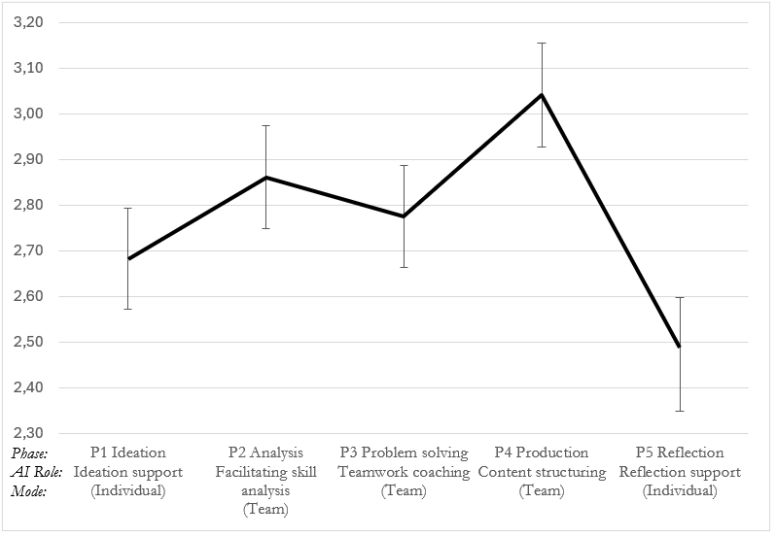

Figure 1: Mean perceived AI usability across learning tasks (P1–P5)

The repeated-measures ANOVA indicated significant differences in perceived AI usability across the P1–P5, $F(3.39, 1388.62) = 35.72$, $p < .001$, partial $\eta^2 = .08$ (see Table 3). Bonferroni-adjusted pairwise comparisons further indicated that perceived AI usability was highest in P4 Production and lowest in P5 Reflection, while P1–P3 formed a middle cluster with only one significant difference (between P1 and P2). Descriptively, perceived AI usability appeared somewhat lower in the individual tasks P1 and P5 than in the team-based tasks P2–P4. Figure 1 illustrates the mean perceived AI usability across the five tasks ($N = 411$), with error bars representing 95% confidence intervals.

RQ2. Study ability and perceived AI usability

To address RQ2, correlations between study ability and perceived AI usability across the five learning tasks were examined. The results are presented in Table 4.

Table 4: Correlations between study ability and perceived AI usability

Phase	*r*	*p*
P1 Ideation	.10	.032
P2 Analysis	.13	< .001
P3 Problem solving	.09	.029
P4 Production	.055	.236
P5 Reflection	.11	.020

Note. r = Pearson correlation coefficient.

Study ability showed generally weak associations with perceived AI usability across learning tasks (see Table 4). The strongest positive correlation was observed for P2 ($r = .13$, $p < .001$). Smaller positive correlations were found for P1 ($r = .10$, $p = .032$), P3 ($r = .09$, $p = .029$), and P5 ($r = .11$, $p = .020$). No statistically significant association was found for P4 ($r = .055$, $p = .236$).

RQ3. Team-level similarity in perceived AI usability

To address RQ3, intraclass correlation coefficients (ICC) were calculated to examine team-level similarity in perceived AI usability across the five learning tasks. The results are presented in Table 5.

Table 5. Intraclass correlation coefficients for perceived AI usability tasks

Phase	ICC(1)	ICC(2)
P1 Ideation	.10	.34
P2 Analysis	.15	.45
P3 Problem solving	.09	.32
P4 Production	.12	.38
P5 Reflection	.09	.33

Note. ICC(1) indicates the proportion of variance attributable to team membership; ICC(2) indicates the reliability of team-level means.

ICC(1) values ranged from .09 to .15, indicating small but noticeable similarity in perceived AI usability within teams across learning tasks. The highest ICC(1) value was observed for P2 (.15), followed by P4 (.12). Lower values were observed for P1 (.10), P3 (.09), and P5 (.09). ICC(2) values ranged from .32 to .45, remaining below the commonly suggested threshold for reliable aggregation of team-level means.

5 Discussion

This study provides new insight into how perceived AI usability varies across phases of the learning process, within student teams, and in relation to study ability, and illustrates how structured GenAI roles in peer-learning contexts may shape students' engagement with AI. Firstly, perceived AI usability varied across the phases of the learning process (RQ1), with the highest levels observed during collaborative production (P4) and the lowest during individual reflection (P5). This may suggest that AI is perceived as more useful in phases involving content development and the preparation of shared outputs. Similar observations have been reported in earlier studies showing that students often apply GenAI for idea generation, drafting, and content development (e.g., Krause et al., 2024; Chen et al., 2026).

Secondly the results highlight the role of **social learning context**, while perceived AI usability was higher when addressing teamwork challenges P3 than as individual reflection support P5, suggesting that collaborative contexts may provide additional opportunities for students to experiment with GenAI when engaging in tasks involving reflection. Positioning of GenAI within the social structure of the learning process may thus influence how useful students perceive AI support to be. ICC analyses indicate also small but consistent within-team similarity in perceived AI

usability, strongest while GenAI was supporting skill analysis (P2) and moderately visible in content structuring (P4), both requiring shared interpretation or coordinated team outputs. In contrast, phases involving more reflection (P3 & P5), showed lower similarity within teams. The findings can be interpreted through a peer-learning perspective, suggesting that AI use is not only an individual behaviour but is shaped through shared practices and collaborative processes within student teams (see also e.g. Sichterman et al. 2025). This aligns with research on peer learning and shared regulation, where learning emerges through interaction and co-construction rather than individual action alone.

Thirdly, reflective learning places high demands on **metacognitive regulation**, while effective use of GenAI similarly requires learners to formulate goals and critically evaluate AI outputs (e.g. Wang, 2023). As Levin et al. (2025) note, these metacognitive capacities are not automatically activated, which may explain the low perceived usability of AI in P5. In line with this, study ability showed weak but consistent associations with perceived AI usability and students with higher study ability perceived AI as slightly more usable.

Taken together, these findings suggest that perceived AI usability reflects both individual metacognitive experiences and shared practices within teams, as indicated by variation across learning phases and modest within-team similarity. When learning activities require collective sensemaking or coordinated outputs, results are consistent with CSCL research on shared learning practices (Fischer et al., 2013). However, ICC varied across learning phases and team similarity appeared linked to the pedagogical role assigned to GenAI in each task, suggesting that shared human–AI collaboration practices may develop when AI engagement is embedded in structured collaborative learning activities.

This study relied on self-reported measures, which may not fully reflect actual practices, limiting the construct validity of the measures. The findings are also based on a single course context, within one institution and team interaction processes were not directly measured. This limits conclusions about the underlying mechanisms through which shared AI practices emerge.

At the same time, the study was situated in a task-based peer-learning environment that reflects common features of project-based higher education. The integration of GenAI into real-world tasks provides a context that closely resembles how AI is used in both educational and professional settings. While the findings cannot be directly generalised, they may offer indicative insights into how pedagogically structured AI use functions in similar collaborative learning environments

The findings contribute to discussions on pedagogical frameworks for AI integration by suggesting that GenAI can function as a collaborative partner whose role varies across learning phases. This aligns with broader frameworks emphasising AI-supported problem solving, collaboration, and learner agency (e.g., UNESCO, 2024), while extending them by showing how these principles become visible in practice through pedagogically structured AI roles. In contrast to frameworks that primarily emphasise AI competencies or instructional use (e.g., Zhou & Schofield; Zou D. et al., 2025), this perspective positions GenAI roles within the learning process itself. The study thus shifts attention from general AI competencies toward the design of learning processes in which meaningful and context-sensitive AI use can emerge.

Conclusions

This study advances understanding of how pedagogical design can support meaningful human–AI collaboration in HE. It shows how structured tasks and peer-learning contexts can make human–AI collaboration more visible, purposeful, and accessible to students with different levels of metacognition. The pedagogical focus should therefore be on instructional guidance that supports meaningful use of GenAI, particularly in reflective and metacognitively demanding tasks, rather than assuming such engagement emerges spontaneously.

Future research should move beyond viewing human–AI collaboration as an individual skill and examine how pedagogical design, learners' self-regulatory capacities, and collaborative contexts shape human–AI practices, framing AI use as a socio-technical learning process shaped by tasks and team dynamics.

End notes

Use of Generative AI in the Writing Process. Generative AI tools were used to support language editing and to assist in identifying potentially relevant literature during the literature search. All sources were independently verified by the authors, who take full responsibility for the manuscript.

References

Abdelaal, M. S., AbdElaal, E.-S., Janssen, A., Smith, E., & Diaby, A. L. (2025). Integrating GenAI in higher education: The critical voices shaping pedagogy and integrity. In S. Barker, S. Kelly, R. McInnes, & S. Dinmore (Eds.), *Futurefocused: Educating in an era of continuous change. Proceedings ASCILITE 2025* (pp. 188–198). https://doi.org/10.65106/apubs.2025.2639

Acosta-Enriquez, B. G., Arbulú Ballesteros, M. A., Arbulu Perez Vargas, C. G., Huamaní Jordan, O., & López Roca, C. (2024). Analysis of college students' attitudes toward the use of ChatGPT in their academic activities: Effect of intent to use, verification of information and responsible use. *BMC Psychology*, 12(1), 255. https://doi.org/10.1016/j.heliyon.2024.e38315

Almassaad, A., Alajlan, H., & Alebaikan, R. (2024). Student Perceptions of Generative Artificial Intelligence: Investigating Utilization, Benefits, and Challenges in Higher Education. *Systems*, *12*(10), 385. https://doi.org/10.3390/systems12100385

Anderson, E., Lin, G. C., Farid, A., Fenech, M., Hanks, B., Klopfer, E., ... & Nguyen, A. (2025). Exploring Genai Technologies within collaborative learning. In Proceedings of the 18th International Conference on Computer-Supported Collaborative Learning-CSCL 2025, pp. 499-506. International Society of the Learning Sciences.

Atchley, P., Pannek, H., Wofford, K., Hopkins, M., & Atchley, R. A. (2024). Human and AI collaboration in the higher education environment: Opportunities and concerns. *Cognitive Research: Principles and Implications, 9*(20). https://doi.org/10.1186/s41235-024-00547-9

A Systematic Literature Review on the Pedagogical Implications and Impact of GenAI on Students' Critical Thinking https://doi.org/10.3390/a19030179

Balabdaoui, F., Dittmann-Domenichini, N., Grosse, H., et al. (2024). A survey on students' use of AI at a technical university. *Discover Education*, 3, 51. https://doi.org/10.1007/s44217-024-00136-4

Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Grundy, S., & Hassan, G. M. (2025). *Implementing generative AI (GenAI) in higher education: A systematic review of case studies*. Computers and Education: Artificial Intelligence, 8, 100407. https://doi.org/10.1016/j.caeai.2025.100407

Carmi, G. (2025). Learning with Generative AI: An Empirical Study of Students in Higher Education. Education Sciences, 15(12), 1696. https://doi.org/10.3390/educsci15121696

Chen, Y., Shi, X., Liu, X., Wang, S., Liu, T. X., & Gašević, D. (2026). Not all students engage alike: Multi-institution patterns in GenAI tutor use. *arXiv*. https://arxiv.org/abs/2602.00447

Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. In M. S. Khine & I. M. Saleh (Eds.), New science of learning: Cognition, computers and collaboration in education (pp. 525–552). *Springer*. https://doi.org/10.1007/978-1-4419-5716-0_26

Feng, S. (2025). Group interaction patterns in generative AI-supported collaborative problem solving: Network analysis of the interactions among students and a GAI chatbot. *British Journal of Educational Technology*, 56, 2125–2145. https://doi.org/10.1111/bjet.13611

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational psychologist*, *48*(1), 56-66. https://doi.org/10.1080/00461520.2012.748005

Guo, Y., Huang, L., Zhang, C., Li, Q., & Chen, M. (2025). ChatGPT in Education: A Systematic Review of Its Impact on Critical Thinking Skills and Dispositions. *Thinking Skills and Creativity*, 102106. https://doi.org/10.1016/j.tsc.2025.102106

Shoval, D. H. (2025). Artificial intelligence in higher education: Bridging or widening the gap for diverse student populations? *Education Sciences*, 15(5), 637.
https://doi.org/10.3390/educsci15050637

Huang, D., Huang, Y., & Cummings, J. J. (2024). Exploring the integration and utilisation of generative AI in formative e-assessments: A case study in higher education. *Australasian Journal of Educational Technology*, 40(4), 1–19. https://doi.org/10.14742/ajet.9467

Irish, A. L., Gazica, M. W., & Becerra, V. (2025). A qualitative descriptive analysis on generative artificial intelligence: Bridging the gap in pedagogy to prepare students for the workplace. *Discover Education*, 4, 48. https://doi.org/10.1007/s44217-025-00435-4

Jia, Y., Li, C., Chan, A., & Li, Q. (2026). Human-AI Collaboration in the STEM Classroom: A Systematic Literature Review of GenAI as a Complement in Higher Education. In International Conference on Technology in Education (pp. 389-401). Springer, Singapore.

Järvelä, S., Nguyen, A., & Hadwin, A. F. (2023). Human and artificial intelligence collaboration for socially shared regulation in learning. British Journal of Educational Technology, 54(5), 1057–1076. https://doi.org/10.1111/bjet.13325

Krause, S., Panchal, B. H., & Ubhe, N. (2024*).* The evolution of learning: Assessing the transformative impact of generative AI on higher education. *arXiv*. https://doi.org/10.48550/arXiv.2404.10551

Law, N., Wang, N., Ma, M., Liu, Z., Lei, L., Feng, S., ... & Tsao, J. (2025). The role of generative AI in collaborative problem-solving of authentic challenges. British Journal of Educational Technology.

Levin, I., Marom, M., & Kojukhov, A. (2025). Rethinking AI in education: Highlighting the metacognitive challenge. *BRAIN: Broad Research in Artificial Intelligence and Neuroscience, 16(Special Issue 1)*, 250–263. https://doi.org/10.70594/brain/16.S1/21

Martínez-Peláez, R., Mena, L. J., Toral-Cruz, H., Ochoa-Brust, A., Potes, A. G., Flores, V., Ostos, R., Pacheco, J. C. R., Félix, R. A., & Félix, V. G. (2025). Can large language models foster critical thinking, teamwork, and problem-solving skills in higher education? A literature review. *Systems*, 13, 1013. https://doi.org/10.3390/systems13111013

Qian, Y. Pedagogical Applications of Generative AI in Higher Education: A Systematic Review of the Field. *TechTrends* **69**, 1105–1120 (2025). https://doi.org/10.1007/s11528-025-01100-1

Parpala, A., & Lindblom-Ylänne, S. (2012). Using a research instrument for developing quality at the university. *Quality in Higher Education*, 18(3), 313–328. https://doi.org/10.1080/13538322.2012.733493

Pham, T. D., Karunaratne, N., Exintaris, B., et al. (2025). The impact of generative AI on health professional education: A systematic review in the context of student learning. *Medical Education*, 59(12), 1280–1289. https://doi.org/10.1111/medu.157

Portocarrero Ramos, H. C., Cruz Caro, O., Sánchez Bardales, E., Quiñones Huatangari, L., Campos Trigoso, J. A., Maicelo Guevara, J. L., & Chávez Santos, R. (2025). Artificial intelligence skills and their impact on the employability of university graduates. *Frontiers in Artificial Intelligence*, 8, Article 1629320. https://doi.org/10.3389/frai.2025.1629320

Ramos, B., & Condotta, R. (2024). Enhancing learning and collaboration in a unit operations course: Using AI as a catalyst to create engaging problem-based learning scenarios. *Journal of Chemical Education, 101*(8), 3246–3254. https://doi.org/10.1021/acs.jchemed.4c00244

Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative Working and Critical Thinking: Adoption of Generative Artificial Intelligence Tools in Higher Education. *Sustainability*, 16(13), 5367. https://doi.org/10.3390/su16135367

Salinas-Navarro, D. E., Vilalta-Perdomo, E., Michel-Villarreal, R., & Montesinos, L. (2024). Using Generative Artificial Intelligence Tools to Explain and Enhance Experiential Learning for Authentic Assessment. *Education Sciences*, 14(1), 83. https://doi.org/10.3390/educsci14010083

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Sichterman, B., Noroozi, O., Boetje, J., van Ginkel, S., Khosravi, H., & Versendaal, J. (2025). Supporting peer learning with artificial intelligence: A systematic literature review. Innovations in Education and Teaching International, 62(5), 1648–1664. https://doi.org/10.1080/14703297.2025.2530118

Wang, L. (2023). Metacognitive strategies used in ChatGPT among college students. *Lecture Notes in Education Psychology and Public Media*, *21*, 283-288. https://doi.org/10.54254/2753-7048/21/20230170

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), 5–23. https://doi.org/10.1007/BF02504682

Zhao, Z., An, Q., & Liu, J. (2025). Exploring AI tool adoption in higher education: Evidence from a PLS-SEM model integrating multimodal literacy, self-efficacy, and university support. *Frontiers in Psychology*, 16. https://doi.org/10.3389/fpsyg.2025.1619391

Zhou, X., & Schofield, L. (2024). Developing a conceptual framework for Artificial Intelligence (AI) literacy in higher education. *Journal of Learning Development in Higher Education*, (31). https://doi.org/10.47408/jldhe.vi31.1354

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zou, D., Xie, H., & Kohnke, L. (2025). Navigating the future: Establishing a framework for educators' pedagogic artificial intelligence competence. *European Journal of Education*, 60, e70117. https://doi.org/10.1111/ejed.70117

Zou, Y., Kuek, F., Feng, W., & Cheng, X. (2025). Digital learning in the 21st century: Trends, challenges, and innovations in technology integration. *Frontiers in Education*, *10*, 1562391. https://doi.org/10.3389/feduc.2025.1562391

Automatic Transformation of Textual Content Using LLM-Technology: a Feasibility Study

Koen Smit, Sam Leewis,
Annemae van de Hoef, Duain Castro
HU University of Applied Science Utrecht, Utrecht, the Netherlands
koen.smit@hu.nl, sam.leewis@hu.nl, annemae.vandehoef@hu.nl. duain.castro@hu.nl

This study investigates the feasibility of using Large Language Model (LLM)-technology to classify and transform textual (educational) content according to CEFR proficiency levels in support of digital accessibility in higher education. We developed a browser-based Proof-of-Concept that performs on-the-fly CEFR classification and transform text to target levels (A1–C2) in Dutch and English. Using a corpus of 120 texts evenly distributed across CEFR levels, we evaluated classification accuracy and transformation effectiveness under multiple temperature settings. Results show modest zero-shot classification performance and systematic mid-level bias, particularly for Dutch. Transformation outcomes were stronger in English, especially for B1–B2 targets, but weak for extreme levels and prone to instability. The findings suggest that while LLMs show promise for automated readability adaptation, reliable deployment requires task-specific tuning, multilingual robustness testing, and human-centered evaluation to ensure meaningful accessibility gains.

DOI
https://doi.org/10.18690/um.fov.4.2026.35

ISBN
978-961-299-152-4

Keywords:
large language models, CEFR, digital accessiblity, text transformation, research paper

University of Maribor Press

1 Introduction

Providing universally accessible educational materials is a critical goal in higher education. Digital accessibility guidelines such as Web Content Accessibility Guidelines (WCAG) emphasize that content must be clear and understandable for diverse learners (World Wide Web Consortium (W3C), 2026), and initiatives like *Easy-to-Read* and *Plain Language* exist specifically to simplify complex texts (Martínez et al., 2024). In practice, however, achieving clarity remains challenging. WCAG's "Understandable" principle advises using straightforward language (short sentences, common words, simple tense) because users with cognitive or language impairments may struggle with complexity (World Wide Web Consortium (W3C), 2024). Similarly, for example, research shows that a large fraction of the Dutch-speaking population requires content at around Common European Framework of Reference for Languages (CEFR) B1 level to ensure comprehension (OECD, 2013; Raad voor Cultuur, 2019). If institutional websites or course materials are written at higher levels (CEFR B2–C2), significant portions of readers may be excluded. Yet many educational texts remain too complex, and manual simplification is labor-intensive and inconsistent (Martínez et al., 2024). Ensuring accessible language is therefore not only a technical or regulatory requirement but also reflects broader human values such as inclusion, equity, and equal participation in education.

Across higher education, digital accessibility has increasingly become embedded in binding regulatory frameworks. In the European context, the WCAG are formally codified in ISO/IEC 40500:2025 and operationalized through EN 301 549 and the European Accessibility Act, thereby extending accessibility requirements to websites, digital documents, and software systems used by public institutions (ETSI, 2025; European Commission, 2025; ISO, 2025; Ministerie van Algemene Zaken, 2017; Rijkswaterstaat, 2025). Although WCAG does not prescribe a fixed reading level, literacy research provides an empirical benchmark for what constitutes broadly accessible communication. Data from the Organisation for Economic Co-operation and Development (OECD)'s Programme for the International Assessment of Adult Competencies (PIAAC) indicate that a considerable proportion of adults perform at literacy levels comparable to CEFR B1 or below (OECD, 2025). In this regulatory and demographic context, the challenge is not merely drafting accessible text but systematically assessing and operationalizing language-level alignment within digital environments.

Recent advances in Large Language Models (LLMs) offer new solutions. These AI models can analyze and transform text, potentially automating readability adaptation. Several studies report dramatic improvements in accessibility when using LLMs to simplify content. For instance, one study found that LLM-generated summaries of medical articles lowered the average reading grade level from about 9.5 to roughly 5–7 while preserving factual accuracy (Romoff et al., 2025). Similarly, LLMs (ChatGPT, Gemini, Claude) reduced health education pamphlets from grade ~10 down to the recommended 5–6 level with no loss of understandability (Will et al., 2025). In educational settings, AI-tools have allowed students to select textbook passages and receive AI-simplified versions, significantly reducing complexity without altering meaning (Johnson et al., 2025). These findings suggest that LLMs can effectively generate "easy-to-read" versions of complex texts while maintaining meaning.

Building on this potential, our research answers the following question: To what extent can an LLM Proof-of-Concept (PoC) reliably classify the accessibility level of Dutch and English texts and transform them into more accessible versions while preserving meaning? We explore this by developing a browser plugin that (1) automatically classifies any text's reading level (in CEFR terms) and (2) transforms the text to different target levels. This design-science PoC allows content creators and users to personalize language complexity on the fly. By focusing on both classification and transformation in two languages, our work goes beyond prior studies to examine feasibility for inclusive, personalized content delivery in higher education.

2 Background and Related Work

Web accessibility standards underscore the importance of plain language. WCAG 2.0 (Guideline 3.1) and related materials advise using clear, simple wording and short sentences to make information comprehensible (World Wide Web Consortium (W3C), 2026). This aligns with broader "plain language" and Universal Design principles that call for well-organized, readable course materials (Martínez et al., 2024). In Europe, text complexity is often gauged by the CEFR levels (A1–C2), which describe learners' language proficiency (Council of Europe Publishing, 2020). For example, studies equate WCAG AAA readability targets to about CEFR B1 (roughly elementary schooling level), see e.g., (Velleman & Van Der Geest, 2014; World Wide Web Consortium (W3C), 2025). In practice, most native adults achieve

around B1–B2; writing above that level risks excluding those with limited literacy. Although CEFR was developed for language learners, it provides a useful scale for "accessibility leveling" of texts. For example, Dutch experts have also proposed native-language frameworks, such as the Referentiekader Nederlandse Taal, to more precisely define readability for native speakers.

Traditional readability metrics (e.g. Flesch-Kincaid (DuBay, 2004)) use surface features like sentence length and word syllables, but these often miss deeper aspects of difficulty. Machine learning approaches have been developed to predict CEFR levels more accurately. For instance, neural models (BERT, RoBERTa) trained on CEFR-annotated corpora can classify Dutch and English texts by reading level. These classifiers capture linguistic patterns beyond formulaic scores, and outperform simple heuristics in predicting text complexity. Large annotated datasets and corpora are now available: for example, the CEFR-based Sentence Profile corpus and multilingual resources like UniversalCEFR provide training material across languages (Zhang et al., 2026). Nonetheless, high-quality CEFR datasets are scarcer for Dutch than for English. Some authors have constructed Dutch readability corpora (e.g. via expert assessment tools, see e.g. a Dutch dataset by Breuker (2023) or an assessment tool by Velleman & Van Der Geest, (2014), but the volume remains limited. Hence, while automated scoring tools exist (and some web tools can estimate CEFR levels), their coverage and reliability vary.

Text simplification has been studied extensively in Natural Language Processing (NLP), which can be used to extract semantics from text (Nadkarni et al., 2011). Early methods used handcrafted rules or statistical models to split sentences and substitute complex words (Siddharthan, 2006). Neural approaches later emerged (encoder-decoder and transformer models) for end-to-end simplification. Recent research focuses on controlled simplification, where a complex sentence is rewritten to match a target proficiency level. Zhang et al. (2026) note that LLMs struggle with large leaps in difficulty but can improve when guided step-by-step. In practice, LLMs (especially GPT-style models) have demonstrated strong single-shot simplification capabilities: for example, one study showed GPT-3.5 Turbo greatly outperformed a traditional translation-based pipeline in simplifying Dutch municipal texts (Vlantis et al., 2024). LLMs have also been used to generate "easy-to-read" Spanish texts by fine-tuning on accessible corpora (Martínez et al., 2024). Furthermore, Leewis et al. (2025) evaluated nine LLMs for CEFR classification of English and Dutch

educational texts, finding that larger models perform best but results vary across models and temperature settings. In this context, temperature is a configurable parameter of an LLM (typically ranging from 0 to 1) that controls the variability of the model's output. Higher values lead to more diverse and less predictable responses, while lower values produce more consistent and deterministic outputs (Peeperkorn et al., 2024).

Empirical work consistently finds that LLM simplifications reduce grade levels and improve readability while largely preserving meaning. In healthcare, ChatGPT and similar models significantly lowered the reading difficulty of patient education materials (to around 6th-grade level) without introducing factual errors (Will et al., 2025). In orthopaedics, GPT-based tools reduced the average reading grade of spine health articles from ~9.5 to ~5–6, with reviewers noting "acceptable accuracy" (Romoff et al., 2025). In educational publishing, an AI-powered "simplifier" in a textbook e-reader reduced language complexity (as measured by syntax and lexicon) while maintaining semantic fidelity (Johnson et al., 2025). These results suggest LLMs can dynamically adapt content to different proficiency targets. Crucially, however, few studies have tested these tools in higher education contexts or in languages other than English.

Our work builds on the notion that both designers as well as learners should control content complexity. Related educational tools (e.g. adaptive learning platforms) tailor material to student proficiency, but typically focus on question difficulty or topic sequencing and less on personalization of the content itself for inclusiveness. Browser extensions like Brisk (Brisk Teaching, 2026) have begun allowing readers to request simplified versions of web text, but these are stand-alone and often not scientifically evaluated. Accessibility checkers (e.g. WAVE, axe) and readability scoring tools exist, yet they usually only flag issues rather than transform text. Our plugin is novel in combining on-the-fly language-level classification with LLM-driven rewriting. By embedding this into a web browser, we enable content creators and consumers to label and transform any page's text in real time (Johnson et al., 2025; Romoff et al., 2025; Will et al., 2025) , and that classifiers can assign CEFR levels based on learned language features. However, to the knowledge of the authors, no published studies have evaluated the reliability of this technology for higher-education materials or directly compared its impact on Dutch and English content.

Our study addresses this gap by prototyping an accessibility tool and assessing its effectiveness for multilingual, inclusive content.

3 Research Method

This study follows a design-science orientation: we develop an LLM-based working artifact (a Chrome browser plugin) to demonstrate feasibility, and we evaluate its technical performance and practical utility through controlled experiments and complementary stakeholder conversations (Hevner et al., 2004). The artifact implements two core functions: (1) automatic CEFR classification of on-page texts and (2) automatic transformation (rewriting) of a given text to any CEFR target (A1–C2) while aiming to preserve meaning.

Experimental design and variables

Each experiment used the same local LLM runtime, i.e., a locally deployed Copilot variant. Due to policy restrictions of the case organization, Copilot is selected as the only fitting LLM. For classification, each input text was submitted with three decoding/temperature settings (0.2, 0.6, and 1.0) to probe stability. The LLM output was then saved for further evaluation. For transformation, for each source text the model was instructed to transform it to each CEFR level (A1–C2), producing one rewritten variant per target level. All prompts used a role-assignment pattern (e.g., "You are an expert English teacher…") and were delivered in the language of the input (English prompts for English texts, Dutch prompts for Dutch texts) to avoid language-mixing artefacts documented in preliminary testing. Non-conforming or empty model outputs were logged as "No Result (NR)". The design decisions (local Copilot LLM, three temperature settings) were chosen to balance reproducibility, privacy, and an exploration of model determinism vs. diversity (Holtzman et al., 2020).

We evaluated the PoC using a corpus of 120 short texts: 60 Dutch and 60 English, evenly distributed across CEFR levels (A1–C2; 10 texts per level). The texts were representative of higher-education web contexts (e.g., course descriptions, policy pages, informational texts). Gold-standard CEFR labels were available in the English dataset used in this study, which consisted of an excerpt from an existing dataset (Montgomerie, 2021). For Dutch, a separate dataset was created by classifying short

texts using an existing text classification tool (Velleman & Van Der Geest, 2014). Regarding the validation of rewriting results, a model trained for English CEFR classification was used to evaluate the rewritten English texts (Lee, 2025). Alternative approaches were not feasible, as they either introduced additional CEFR levels or were not freely accessible. For Dutch, the same classification tool was used to evaluate the rewritten texts.

For the classification experiment, each text was submitted to the locally deployed Copilot LLM. The model was instructed to classify the manually inputted text into one of the CEFR levels (A1–C2). Each text was processed three times using the predefined temperature settings described in the experimental design (0.2, 0.6, and 1.0). Outputs were constrained to a predefined JSON structure containing the assigned CEFR level. Empty or malformed outputs were logged as "No Result" (NR). Classification performance was assessed by comparing the model's assigned level with the known CEFR label of the input text. Results were summarized using overall accuracy and confusion matrices to identify patterns of over- or under-classification across levels. Differences between temperature settings were examined descriptively.

For the transformation experiment, each original text was rewritten to each CEFR target level (A1–C2) using structured prompts. The model output consisted of a rewritten version of the text in JSON format. The rewritten texts were subsequently re-classified to determine whether the intended target level had been achieved. The primary measure for transformation was therefore the proportion of rewritten texts that were classified at the requested CEFR level. In addition to quantitative comparison, outputs were qualitatively inspected to identify recurring patterns, such as systematic over- or under-simplification, instability across temperature settings, or deviations from the intended target level. This descriptive analysis was used to interpret performance trends and inform refinement of prompt design. All prompts and output logs were archived to ensure transparency and reproducibility of the experiments. The data and analysis is available upon request.

4 Results

An initial PoC was developed on the basis of several requirements gathered from stakeholders during exploratory discussions. These conversations were intentionally unstructured and open in nature, organised around thematic expectations concerning potential users of the plugin among the staff of the university of applied sciences.

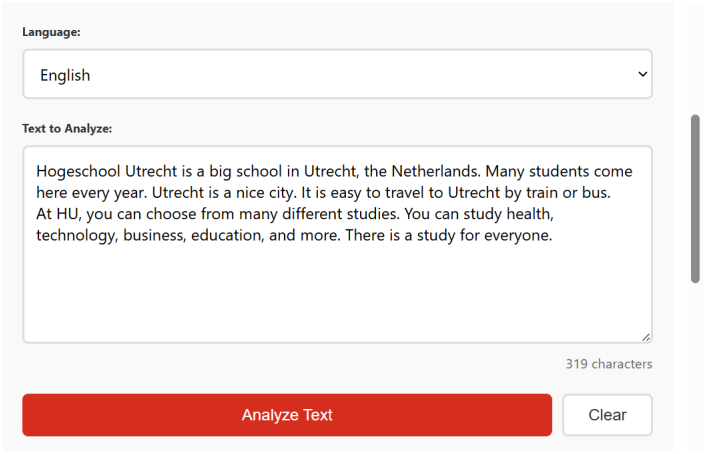

Figure 1: PoC screenshot of the textual input
Source: Own

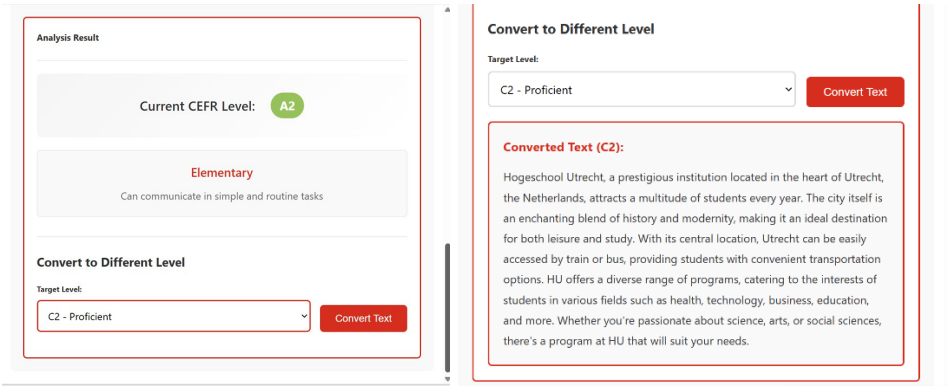

Figure 2: CEFR-based language classification (left) and transformation (right) in the PoC
Source: Own

Given that the PoC represents an early-stage implementation, the primary focus was testing the feasability of utilizing a LLM on classifying and transforming Dutch and English texts. In this early-stage the focus lies on establishing core functionality and testing the performance of the underlying LLM in relation to the front-end interface. Particular attention was paid to ensuring security and privacy, by running the LLM locally. Screenshots of the PoC are presented in Figure 1 (textual input) and Figure 2 (left: classification result, right: transformation functionality).

Classification

Classification performance was modest for both languages. Using the Copilot LLM via the browser plugin, each text was classified at three temperature settings (0.2, 0.6, 1.0). For English, overall accuracy was 33.33% at temperatures 0.2 and 0.6, and 23.33% at 1.0 (Table 2). For Dutch, the best accuracy (35.0%) occurred at temperature 0.2, decreasing to 26.67% at 0.6 and 23.33% at 1.0 (see Table 1). Based on these results, a lower temperature setting (0.2), which produces more deterministic outputs, appears to yield slightly better performance for Dutch. For English, temperatures of 0.2 and 0.6 show comparable results, suggesting that performance is relatively stable across lower to mid-range temperature settings. However, overall differences between settings are modest, indicating that temperature may have only a limited impact on classification performance in this setup.

Table 1: Accuracy metrics for CEFR text classification

Temperature	Total	Correct NL	Accuracy NL	Correct EN	Accuracy EN
0.2	60	21	35%	20	33,33%
0.6	60	16	26,67%	20	33,33%
1.0	60	14	23,33%	14	23,33%

English (Temp 0.2, Figure 3). Classifications are relatively dispersed across levels, but systematic errors are evident. B2 texts are commonly recognised, whereas A2 examples are often mislabelled as A1 or B1. High-level texts (C1, C2) are frequently predicted as B2, indicating a mid-level convergence bias (tendency to cluster around B2). *Dutch (Temp 0.2, Figure 3).* The model shows a strong bimodal tendency towards A1 and B2. All A1 texts were correctly classified, but A2 examples were entirely predicted as A1; most higher-level texts (B1–C2) were classified as B2. This indicates difficulty distinguishing adjacent levels and particular confusion around the mid-

range levels (B1/B2). *English (Temp 0.6, Figure 4)*. The classifications closely mirrored those at temperature 0.2: A1 and B2 texts were consistently identified, while higher-level texts were often predicted as B2. *Dutch (Temp 0.6, Figure 4)*. Similarly, the results resembled those at temperature 0.2, again revealing a strong bimodal tendency toward A1 and B2 classifications. *English (Temp 1.0, Figure 5)*. At this temperature, the LLM demonstrated the greatest difficulty in classification, producing highly sporadic results. Notably, B1–C1 texts were not correctly classified once. *Dutch (Temp 1.0, Figure 5)*. Similarly, this temperature yielded more sporadic results. This is unsurprising, as classification is generally considered a deterministic task. A temperature of 1.0 is therefore too high for this purpose, as it reduces the model's determinism and leads to less consistent predictions.

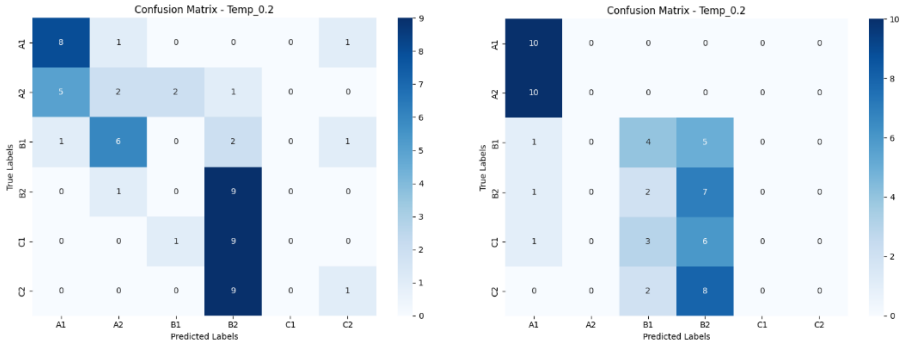

Figure 3: Confusion patterns of English (left) and Dutch (right) CEFR text classification (T = 0.2)

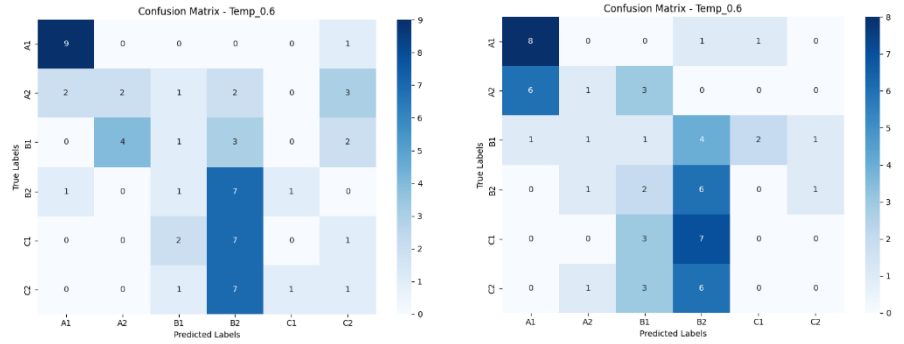

Figure 4: Confusion patterns of English (left) and Dutch (right) CEFR text classification (T = 0.6)

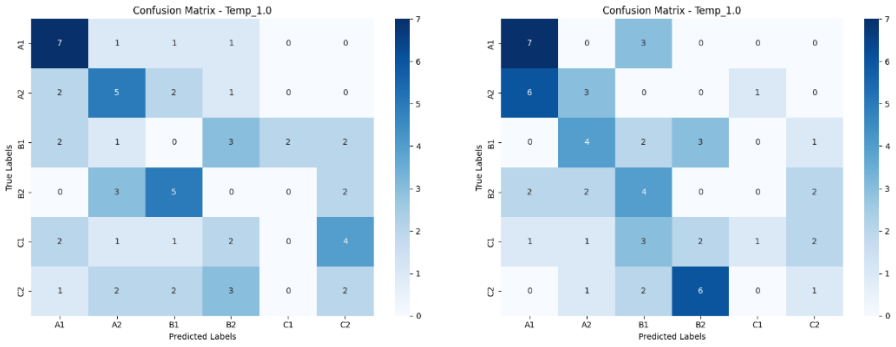

Figure 5: Confusion patterns of English (left) and Dutch (right) CEFR text classification (T = 1.0).

Overall accuracy is low in absolute terms and the confusion matrices show systematic biases rather than random noise. For English the model tends to collapse higher levels into B2; for Dutch it concentrates predictions at A1 or B2, with A2 collapsing to A1 and higher levels mapped to B2. The results suggest that, in the present zero-shot configuration, the Copilot LLM struggles to finely discriminate adjacent CEFR categories and that lower-temperature (more deterministic) decoding generally produced better classification accuracy.

Transformation

Language	Temperature	A1	A2	B1	B2	C1	C2
Dutch	0.2	8/56 (14.29)	13/56 (23.21)	18/54 (33.33)	14/54 (25.93)	12/48 (25.00)	10/55 (18.18)
Dutch	0.6	4/57 (7.02)	10/55 (18.18)	18/53 (33.96)	15/55 (27.27)	11/54 (20.37)	11/53 (20.75)
Dutch	1.0	3/55 (5.45)	12/53 (22.64)	21/54 (38.89)	16/52 (30.77)	11/46 (23.91)	9/50 (18.00)
English	0.2	21/60 (35.00)	24/60 (40.00)	17/60 (28.33)	36/60 (60.00)	11/60 (18.33)	3/60 (5.00)
English	0.6	23/60 (38.33)	29/60 (48.33)	22/60 (36.67)	33/60 (55.00)	13/60 (21.67)	6/60 (10.00)
English	1.0	23/60 (38.33)	26/60 (43.33)	24/60 (40.00)	37/60 (61.67)	23/60 (38.33)	5/60 (8.33)

Table 2: Transformation results of both Dutch and English languages. English totals per cell were 60 transformations in the reported experiment. Dutch totals vary because 77 transformations were produced in a different language and 43 transformations returned no output or an invalid text, reducing the evaluable Dutch sample

The transformation experiments measured whether the LLM could transform each source text to specified CEFR target levels (A1–C2). For each language we report the percentage of transformations that were subsequently classified at the requested target level (i.e., target-match accuracy). The English results are substantially stronger than the Dutch ones; additionally, some Dutch transformations could not be evaluated because they were produced in the wrong language or the output was invalid (see note below).

- English. The LLM transforms most reliably to B2, achieving its highest target-match rates (≈55–62%). Mid-range targets (A1–B1) show moderate performance (≈28–48%), while the highest target (C2) has very low match rates (≈5–10%). The effect of temperature is not consistent across CEFR levels. In some cases (e.g., C1 at Temp 1.0), higher temperature leads to improved target-match accuracy, while in other cases no clear or systematic trend is observed.
- Dutch. Across all targets the accuracies are markedly lower. The best Dutch result is for B1 at Temp 1.0 (38.89%). Several target levels (notably A1) perform poorly and often decline as temperature increases. While lower temperature settings (0.2) sometimes yield slightly more stable results for certain levels such as A1, A2, and C1, higher temperatures occasionally produce better outcomes for intermediate levels such as B1 and B2. Overall, similar to English, no clear monotonic effect of temperature on performance can be observed. The Dutch evaluation is further complicated by missing / invalid outputs (see Table 2 note).

The transformation results indicate that the Copilot LLM, in this zero-shot plugin configuration, is better at producing transformations in English and is most effective for mid-level targets (especially B2). The model performs poorly for very advanced targets (C2) in both languages. For Dutch, the frequency of non-usable outputs and language-switching suggests limited robustness, which might be related to the model's training data distribution.

5 Discussion

The work of Martínez et al. (2024), Romoff M et al. (2025), Will et al. (2025), and Johnson et al. (2025) consistently report that LLMs can lower reading difficulty while largely preserving meaning, especially in health and textbook domains. These studies

emphasize human validation, report strong grade-level reductions, and validate AI simplification in single-language (mostly English) settings. Our work complements and extends this evidence by operationalizing CEFR levels (A1–C2) rather than grade metrics, combining both automatic CEFR classification and controlled transformation, and directly comparing Dutch and English performance. Unlike prior studies that focus on simplification outcomes, we reveal systematic zero-shot biases (mid-level convergence for English; bimodal clustering for Dutch) and practical failure modes.

This study has several limitations that we need to address. First, the PoC relied on a locally deployed Copilot LLM without task-specific fine-tuning for CEFR classification or controlled rewriting. The modest classification accuracy and uneven transformation results may therefore reflect limitations of zero-shot prompting rather than intrinsic limits of LLM technology. Also, although three temperatures were tested, temperature strategy exploration remained limited. Other sampling or prompting strategies were not systematically explored. Lastly, recent work shows that LLM-based approaches can achieve strong performance in text classification tasks, particularly when combined with fine-tuning or few-shot strategies, which often outperform zero-shot prompting approaches (Alizadeh et al., 2024; Bucher & Martini, 2024). This suggests that the relatively modest performance observed in this study may be partly attributable to the use of a locally deployed model without task-specific optimization. Another complexity is that the experimental corpus consisted of 120 short texts, which, while stratified by CEFR level, may not capture the full linguistic variability of higher-education materials. Furthermore, the study equates textual accessibility with CEFR level alignment, whereas accessibility as defined by WCAG encompasses additional structural (e.g., semantic markup, heading structure, navigability) and cognitive (e.g., predictability, clarity of instructions, cognitive load) dimensions not addressed in this evaluation. Another limitation is that we evaluated the classification of Dutch texts and the transformation of both languages' effectiveness through automated re-classification, which may inherit limitations of the underlying classification approach and introduce cyclical evaluation. Unfortunately, alternative approaches were not explored due to the limited availability of suitable open-source classification tools within the project constraints. In the end, we argue that this study shows that the LLM-technology, with proper finetuning and evolution (e.g., using guardrails), might be an adequate tool to support practitioners to increase the accessibility of their (digital) textual content.

6 Future Work

Future research should examine how task-specific model adaptation affects CEFR classification and controlled text transformation. This study relied on a general-purpose LLM in a zero-shot configuration, which limited stability when distinguishing adjacent CEFR levels. Experiments with fine-tuning or instruction tuning tailored to CEFR tasks may improve level discrimination. Comparative testing of prompt designs, few-shot versus zero-shot approaches, and alternative temperature settings could further clarify which configurations produce the most reliable outputs. An additional recommendation would be to conduct a benchmark study comparing different LLMs in order to determine which models perform best for this task. The empirical basis should also be expanded by including longer and more complex texts as well as discipline-specific materials such as STEM, legal, or social science content to assess robustness across domains. In addition, more rigorous human-centered evaluation is needed. Automated CEFR reclassification offers a scalable metric but does not fully capture meaning preservation or accessibility improvements. Future work should therefore include expert CEFR raters, stakeholder comprehension testing, and user-based evaluations. Field deployment studies could further explore how stakeholders use such tools and how automated adaptations are perceived in practice. Finally, the performance differences observed between English and Dutch highlight the need for systematic multilingual research, including analysis of cross-language robustness, language-switching outputs, and the effects of language-specific tuning.

7 Conclusion

This study contributes both theoretically and practically to the development on automated readability adaptation in higher education. Theoretically, it shows how CEFR can be operationalized as a scale for digital accessibility research by combining CEFR-based classification with controlled transformation in a design-science artifact. The results also nuance expectations about LLMs: in a zero-shot setting they exhibit systematic biases (mid-level convergence in English and bimodal clustering in Dutch) and limited ability to distinguish adjacent CEFR levels, highlighting limits to out-of-the-box robustness. Practically, the study presents a browser-based plugin for on-the-fly classification and rewriting and identifies key implementation lessons. While LLM-driven rewriting can produce useful adaptations, particularly for mid-

range English levels (B1–B2), issues such as language mixing, invalid outputs, and low reliability for extreme levels (A1, C2) still limit deployment readiness. The local deployment approach further illustrates a privacy-conscious integration strategy. These findings point to a pragmatic agenda for future work: targeted model tuning, richer multilingual corpora, structured human evaluation, and socio-technical assessment when moving from PoC to prototypes and real-world implementation.

Acknowledgments

We would like to thank the team of the Digital Learning Environment of HU University of Applied Sciences Utrecht in the Netherlands for their support in this project.

References

Alizadeh, M., Kubli, M., Samei, Z., Dehghani, S., Zahedivafa, M., Bermeo, J. D., Korobeynikova, M., & Gilardi, F. (2024). *Open-Source LLMs for Text Annotation: A Practical Guide for Model Setting and Fine-Tuning* (arXiv:2307.02179). arXiv. https://doi.org/10.48550/arXiv.2307.02179

Breuker, M. (2023). CEFR Labelling and Assessment Services. In G. Rehm (Ed.), *European Language Grid* (pp. 277–282). Springer International Publishing. https://doi.org/10.1007/978-3-031-17258-8_16

Brisk Teaching. (2026). *Personalize instruction with AI, right in your workflow*. Brisk Teaching. https://www.briskteaching.com/

Bucher, M. J. J., & Martini, M. (2024). *Fine-Tuned 'Small' LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification* (arXiv:2406.08660). arXiv. https://doi.org/10.48550/arXiv.2406.08660

Council of Europe Publishing. (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment - Companion volume*. Council of Europe. https://www.coe.int/en/web/common-european-framework-reference-languages

DuBay, W. H. (2004). *The Principles of Readability*.

ETSI. (2025). *Draft EN 301 549* (Version 4.1.0). ETSI. https://www.etsi.org/deliver/etsi_EN/301500_301599/301549/04.01.00_20/en_301549v040100ev.pdf

European Commission. (2025). *The EAA comes into effect in June 2025. Are you ready?* European Commission. https://accessible-eu-centre.ec.europa.eu/content-corner/news/eaa-comes-effect-june-2025-are-you-ready-2025-01-31_en

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–106. https://doi.org/10.2307/25148625

Holtzman, A., Buys, J., Du, L., Forbes, M., & Choi, Y. (2020). *The Curious Case of Neural Text Degeneration* (arXiv:1904.09751). arXiv. https://doi.org/10.48550/arXiv.1904.09751

ISO. (2025). *ISO/IEC 40500:2025* (2nd edn). https://www.iso.org/standard/91029.html

Johnson, B. G., Jerome, B., Dittel, J. S., & Campenhout, R. V. (2025). Improving Textbook Accessibility through AI Simplification: Readability Improvements and Meaning Preservation. *Proceedings of the Sixth International Workshop on Intelligent Textbooks (iTextbooks 2025)*.

Lee, W. (2025). *CEFR Classifier* [Computer software]. https://huggingface.co/dksysd/cefr-classifier

Martínez, P., Ramos, A., & Moreno, L. (2024). Exploring Large Language Models to generate Easy to Read content. *Frontiers in Computer Science*, *6*, 1394705. https://doi.org/10.3389/fcomp.2024.1394705

Ministerie van Algemene Zaken. (2017, October 31). *Taalniveau B1* [Webpagina]. Dienst Publiek en Communicatie. https://www.communicatierijk.nl/vakkennis/rijkswebsites/aanbevolen-richtlijnen/taalniveau-b1

Montgomerie, A. (2021). *CEFR Levelled English Texts* [Data set]. https://www.kaggle.com/datasets/amontgomerie/cefr-levelled-english-texts

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544–551. https://doi.org/10.1136/amiajnl-2011-000464

OECD. (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. OECD. https://doi.org/10.1787/9789264204256-en

OECD. (2025). *Education at a Glance 2025: OECD Indicators*. OECD Publishing. https://doi.org/10.1787/1c0d9c79-en

Peeperkorn, M., Kouwenhoven, T., Brown, D., & Jordanous, A. (2024). *Is Temperature the Creativity Parameter of Large Language Models?* (arXiv:2405.00492). arXiv. https://doi.org/10.48550/arXiv.2405.00492

Raad voor Cultuur. (2019). *Lees! Een oproep tot een leesoffensief*. Den Haag: Raad Voor Cultuur. https://www.raadvoorcultuur.nl/documenten/2019/06/24/lees

Rijkswaterstaat. (2025). *Schrijftips voor taalniveau B1 in toepasbare regels*. Informatiepunt Leefomgeving. https://iplo.nl/digitaal-stelsel/toepasbare-regels/maken-testen/schrijfwijzer/schrijftips-taalniveau-b1/

Romoff, M., Brunette, M., Peterson, M. K., Hashmi, S. Z., & Kim, M. S. (2025). The role of large language models in improving the readability of orthopaedic spine patient educational material. *Journal of Orthopaedic Surgery and Research*, *20*(1), 531. https://doi.org/10.1186/s13018-025-05955-1

Siddharthan, A. (2006). *Syntactic simplification and text cohesion* (p. 195 pages) [PDF]. Computer Laboratory, University of Cambridge. https://doi.org/10.48456/TR-597

Velleman, E., & Van Der Geest, T. (2014). Online Test Tool to Determine the CEFR Reading Comprehension Level of Text. *Procedia Computer Science*, *27*, 350–358. https://doi.org/10.1016/j.procs.2014.02.039

Vlantis, D., Gornishka, I., & Wang, S. (2024). Benchmarking the Simplification of Dutch Municipal Text. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2217–2226.

Will, J., Gupta, M., Zaretsky, J., Dowlath, A., Testa, P., & Feldman, J. (2025). Enhancing the Readability of Online Patient Education Materials Using Large Language Models: Cross-Sectional Study. *Journal of Medical Internet Research*, *27*, e69955. https://doi.org/10.2196/69955

World Wide Web Consortium (W3C). (2024). *Web Content Accessibility Guidelines 2.2*. https://www.w3.org/TR/WCAG22/

World Wide Web Consortium (W3C). (2025). *Understanding Success Criterion 3.1.5: Reading Level*. World Wide Web Consortium (W3C). https://www.w3.org/WAI/WCAG21/Understanding/reading-level?utm_source=chatgpt.com

World Wide Web Consortium (W3C). (2026, March 12). *Use Clear and Understandable Content*. World Wide Web Consortium (W3C). https://www.w3.org/WAI/WCAG2/supplemental/objectives/o3-clear-content/

Zhang, J., Qiu, X. Y., Lu, L., Huang, Z., Hu, Y., Wu, Y., & Lu, J. (2026). *Let's Simplify Step by Step: Guiding LLM Towards Multilingual Unsupervised Proficiency-Controlled Sentence Simplification* (arXiv:2602.07499). arXiv. https://doi.org/10.48550/arXiv.2602.07499

(Open) Data Stewardship in the Public Sector: Evidence from a Survey of Slovenian Public Administration

Tina Jukić, Sanja Vrbek, Damijana Keržič
University of Ljubljana, Faculty of Public Administration, Ljubljana, Slovenia
tina.jukic@fu.uni-lj.si, sanja.vrbek@fu.uni-lj.si, damijana.kerzic@fu.uni-lj.si

In the context of digital transformation and open government data initiatives, data have become a key strategic resource of contemporary public administration. Based on an online survey, this paper examines how (open) data stewardship is organised in Slovenian public administration. The findings show that data stewardship is present but weakly institutionalised. It is most often performed as an additional task rather than as a formal organisational role. Respondents primarily carry out operational tasks related to data collection, verification and user support, while more strategic responsibilities remain less developed. The results also reveal considerable training needs, particularly in areas such as legal compliance, data quality standards and mechanisms for opening data. This study provides empirical insights into organisational practices and capacity gaps relating to data stewardship in public administration.

DOI
https://doi.org/10.18690/um.fov.4.2026.36

ISBN
978-961-299-152-4

Keywords:
open government data, data stewardship, data governance, data roles and responsibilities, data competencies

University of Maribor Press

1 Introduction

In the context of digital transformation, data have become a key strategic resource of contemporary societies and public administration. Open government data (OGD), in particular, have attracted increasing attention due to their potential to support innovation, economic development and improved public service delivery (Ruijer et al., 2020; Wirtz et al., 2022). At the same time, OGD are recognised for their contribution to transparency, accountability and citizen participation, thereby strengthening democratic governance (Ruijer et al., 2020).

The growing importance of data has been reflected in international and European policy frameworks. The OECD emphasises that public sector data should be treated as a public good and made proactively available for reuse (OECD, 2020). Similarly, the European Union has placed data governance at the centre of its digital transformation agenda through initiatives such as the European Data Strategy and the Digital Decade (European Commission, 2020; 2021). These developments are complemented by a regulatory framework governing access to and reuse of public sector data. In Slovenia, these principles have been incorporated into national legislation and strategic documents, which further emphasise the role of data in public sector modernisation.

Despite these developments, research shows that the value of open data does not arise automatically from data publication alone. Simply releasing datasets rarely leads to societal or economic benefits without appropriate organisational capacities and active data use (Charalabidis et al., 2018; Hitz-Gamper et al., 2019). The creation of public value from OGD therefore depends not only on legal and technical frameworks, but also on organisational practices that ensure data quality, accessibility and reuse.

In this context, data stewardship plays an increasingly important role as a set of organisational practices and responsibilities supporting the management, quality control and (re)use of data. However, empirical knowledge about how data stewardship is organised in practice, particularly within public administration, remains limited. Existing literature also suggests that stewardship roles often remain heterogeneous and only partially formalised across organisational contexts,

reflecting the evolving nature of public sector data governance (Tammaro et al., 2019; OECD, 2023).

To address this gap, this study examines how (open) data stewardship is organised in Slovenian public administration. Based on a survey conducted among public sector organisations, the study analyses the profile of employees involved in data stewardship, the organisational arrangements through which these activities are carried out, and the associated responsibilities and competencies.

The study addresses the following research question: How is (open) data stewardship organised in Slovenian public administration, particularly with regard to organisational arrangements, employee profiles, responsibilities, competencies and perceived strengths and improvement needs?

The remainder of the paper is structured as follows. Section 2 presents the conceptual framework of data stewardship. Section 3 describes the methodology and empirical results, while Section 4 concludes with key findings and implications.

2 Data Stewardship: Roles, Responsibilities and Competencies

The growing importance of data-driven governance and open data policies has increased the relevance of data stewardship in research and organisational contexts. Data stewardship is generally understood as a set of practices ensuring that data are properly managed, documented, preserved and made accessible throughout their lifecycle (Tammaro et al., 2019). Despite its increasing importance, the role remains conceptually heterogeneous and not fully standardised, with different terms such as data steward, data curator or data custodian used across disciplinary and organisational contexts (Tammaro et al., 2019; Tammaro & Casarosa, 2014).

Nevertheless, there is broad agreement that effective data stewardship is essential for enabling data sharing, reuse and long-term value creation (Manu & Gala, 2021). Data stewards are typically responsible for managing data throughout the lifecycle while ensuring accessibility, reliability and reusability, and are expected to combine domain knowledge with an understanding of organisational processes (Tammaro et al., 2019; Jablonski, 2022).

Given the limited literature on open data stewardship in public administration, this study draws on broader research on data stewardship and data curation. The literature review focused on scientific and professional publications from 2014 onwards. These were identified through searches in Google Scholar using combinations of the terms "data steward" or "data curator" with keywords such as "roles", "responsibilities", "profiles", "skills" and "competencies". Since the public administration-specific literature was relatively scarce, the analytical framework was built through a synthesis of more general stewardship literature and practice-oriented sources.

2.1 Responsibilities of Data Stewards

Data stewardship is a multidimensional practice encompassing technical, organisational and governance-related activities. Core responsibilities include ensuring data access, preservation, quality assurance and compliance with relevant standards, policies and legal frameworks (Tammaro et al., 2019; The Turing Way Community, 2022).

Across the data lifecycle, data stewards are involved in activities such as data collection, storage, maintenance, validation and metadata management, which support data discoverability and reuse (Manu & Gala, 2021). Metadata management is particularly important for enabling accessibility and reuse, while compliance with regulatory frameworks, including data protection requirements, represents a key responsibility.

In addition, data stewards contribute to organisational data governance by supporting the development of standards, monitoring data management practices and facilitating data publication and reuse (The Turing Way Community, 2022). They also play a role in promoting good data practices through training, awareness-raising and support for colleagues.

An important aspect of the role is coordination across organisational actors. Data stewards often act as intermediaries between data producers, users and other stakeholders, facilitating communication and fostering a data-driven organisational culture (Manu & Gala, 2021). Operational responsibilities may also include defining procedures for data collection, managing access rights and security levels,

monitoring data usage and supporting the standardisation of data practices (DataGalaxy, 2023).

2.2 Skills and Competencies of Data Stewards

The literature highlights a diverse set of competencies required for effective data stewardship, combining technical and soft skills. Technical competencies include data management, data analysis, database systems and metadata standards, as well as the ability to ensure data quality and maintain data infrastructures (Demchenko, 2016).

At the same time, data stewardship requires strong organisational and interpersonal competencies. These include communication, collaboration, problem-solving and project management, which are essential for operating across technical and organisational domains (Feng & Richards, 2018; Murillo & Yoon, 2024). Soft skills are particularly important for explaining complex data-related issues to non-technical stakeholders and supporting coordination among actors.

Legal and policy-related knowledge is also critical, especially in public sector contexts. Data stewards must understand regulatory frameworks, including data protection requirements and rules governing access to public sector information (Fitsilis et al., 2024). In addition, understanding organisational processes and data governance arrangements is emphasised as a key competence in the public sector (OECD, 2023).

Overall, the role of the data steward is inherently multidimensional, combining technical expertise, organisational coordination and legal awareness. Despite its growing importance, empirical evidence on how these responsibilities and competencies are implemented in public administration remains limited. This gap motivates the present study, which examines (open) data stewardship in Slovenian public administration.

3 Empirical Assessment of (Open) Data Stewardship in Slovenian Public Administration

3.1 Methodological framework

Based on the reviewed literature, a conceptual framework of data stewardship responsibilities and competencies was developed and used to design the survey instrument.

The survey aimed to analyse the profile of (open) data stewards, existing organisational arrangements for data stewardship, and the associated responsibilities and competencies in Slovenian public administration.

The questionnaire was developed based on a synthesis of scientific and professional literature on data stewardship and data curation (e.g., Tammaro et al., 2019; Hudson-Vitale et al., 2017; The Turing Way Community, 2022). Responsibilities were grouped into thematic categories reflecting activities across the data lifecycle, while the competency framework included technical, organisational, legal and soft skills. The questionnaire consisted of 46 questions (117 variables) organised into four sections: (1) respondent profile and organisational context, (2) responsibilities of data stewards, including self-assessed competence and training needs, (3) understanding of the data ecosystem, and (4) soft skills and personal attributes.

Data were collected in January 2025 through an online survey distributed to municipalities (212), ministries (19), bodies within ministries (41), government offices (14), and public agencies (18). The survey link was sent to general organisational e-mail addresses with a request for internal distribution to employees involved in (open) data stewardship. Participation was anonymous, and multiple respondents from the same organisation were possible.

A total of 96 responses were included in the analysis (44 partially and 52 fully completed questionnaires). Since the survey was distributed indirectly via organisations, the total number of potential respondents could not be determined and the response rate could not be calculated. The questionnaire included multiple thematic sections and item batteries, and not all respondents answered every question. Consequently, the number of valid responses varies across individual items

and tables. Partial responses were retained in the analysis in order to preserve available information, particularly because some respondents may not have been familiar with specific stewardship tasks or organisational practices. Therefore, the reported mean values are based on valid responses for each individual item.

Among respondents, 48% were employed in ministries or bodies within ministries, 38% in municipalities, 8% in government offices and 6% in public agencies. Among municipal respondents, the largest shares came from municipalities with up to 5,000 inhabitants (13%) and those with 5,001–15,000 inhabitants (17%).

3.2 Survey Results

3.2.1 Organisational Knowledge and Time Allocation for Data Stewardship

The results indicate a high level of familiarity with organisational context, with average scores between 5.0 and 5.3 on a six-point scale. Respondents reported the highest familiarity with organisational purpose and mission (5.3) and sector (5.2), followed by organisational processes and roles in policy and service delivery (5.0). This suggests that most respondents have a solid understanding of their organisational environment.

At the same time, data stewardship is rarely a full-time role. Most respondents (66%) reported devoting up to 10% of their working time to these tasks, with only a small minority exceeding 30%. This indicates that data stewardship is typically performed as an additional responsibility, pointing to a fragmented and weakly institutionalised approach.

3.2.2 Employees Involved in (Open) Data Stewardship

The results show that data stewardship is typically performed by a small number of employees. Among organisations with at least one responsible employee, most reported between one and four staff involved, indicating limited capacity and low formalisation.

Stewardship activities are most frequently located in IT or digitalisation departments (37%), although respondents also come from a wide range of administrative and sector-specific units. This confirms the cross-cutting nature of data stewardship, while also indicating the absence of a clearly defined organisational "home".

In terms of job positions, data stewardship tasks are most often carried out by mid-level and senior staff, such as advisers, (under)secretaries and heads of units. The diversity of roles further suggests that stewardship is not yet established as a formal function but is performed on an ad hoc basis depending on organisational context.

3.2.3 Assessment of Organisational Stewardship Systems

Respondents identified several key advantages of existing data stewardship arrangements, including improved data accessibility and transparency, support for decision-making, and benefits related to digitalisation and automation. Some also highlighted clearer organisational structures, improved data quality and increased standardisation.

At the same time, important improvement needs were identified. These include technological upgrades and automation, clearer allocation of responsibilities, additional staffing and training, and stronger coordination and exchange of good practices. Respondents also emphasised the need to improve data quality, traceability and security, as well as to increase awareness and engagement among staff.

3.2.4 Responsibilities of (Open) Data Stewards

Respondents assessed the extent to which 21 different tasks represented their actual responsibilities on a six-point Likert scale. None of the tasks reached the maximum possible score, and the averages ranged from 1.9 to 4.0 on a six-point scale (Table 1).

The most strongly represented responsibilities were checking data reliability (4.0), cooperating with and assisting users (3.9), collecting data (3.6), detecting and correcting inconsistencies (3.6), data storage (3.5), and ensuring legal compliance (3.5). These findings show that respondents are primarily engaged in operational and

support-oriented tasks related to data handling, quality control and assistance to users.

Moderately frequent responsibilities included developing internal data management plans and strategies (3.4), documentation management (3.3), anonymisation (3.3), participation in drafting regulations (3.2), promoting open data (3.2), and monitoring developments in the field (3.1). These tasks have a more strategic character, but they still do not appear to be dominant components of the role.

The least frequently reported tasks were more specialised responsibilities, such as metadata management (2.9), indexing (2.5), implementing quality standards (2.4), establishing permanent repositories (2.1), and managing usage rights (1.9). These tasks are highly specialised, while respondents represent very heterogeneous professional backgrounds, which may partly explain this result.

Table 1: Responsibilities of Data Stewards in the Slovenian Public Administration Context (n = 51–55*)

Responsibilities / Tasks	Mean
Verifying data reliability	4.0
Cooperation with and support for users	3.9
Data acquisition / collection	3.6
Identifying and resolving inconsistencies	3.6
Data storage	3.5
Ensuring compliance with applicable legislation	3.5
Developing plans and strategies for data management within the organisation	3.4
Documentation management	3.3
Data anonymisation	3.3
Participation in the development of regulations	3.2
Promoting the use of open data	3.2
Monitoring developments in the field of (open) data management	3.1
Maintaining data integrity and security	3.1
Implementing mechanisms for opening data	3.0
Monitoring access and conditions of use	2.9
Metadata management	2.9
Ensuring and implementing data removal / deletion protocols	2.8
Indexing	2.5
Implementing quality standards	2.4
Establishing persistent repositories	2.1
Managing usage rights	1.9

* Note: n varies across items due to item non-response.

3.2.5 Competence Gaps, Training Needs, and Skills

Table 2: Comparison of Reported Responsibilities, Self-Assessed Capabilities and Training Needs Across Data Stewardship Tasks (n = 51–55*)

	Responsibility	Capability	Need for training	Difference Responsibility vs. Capability	Difference Capability vs. Need for training
Verifying data reliability	4.0	4.3	4.4	-0.2	-0.1
Identifying and resolving inconsistencies	3.6	3.8	4.3	-0.2	-0.5
Performing data anonymisation	3.3	3.7	4.2	-0.4	-0.5
Data storage	3.5	3.6	4.1	0.0	-0.5
Metadata management	2.9	3.1	4.1	-0.2	-1.0
Implementing quality standards	2.4	2.4	4.3	0.0	-2.0
Data acquisition / collection	3.6	3.8	4.2	-0.2	-0.3
Maintaining data integrity and security	3.1	3.0	4.2	0.0	-1.2
Implementing mechanisms for opening data	3.0	2.9	4.3	0.1	-1.4
Ensuring and implementing data removal / deletion protocols	2.8	3.1	4.1	-0.3	-1.0
Documentation management	3.3	3.5	4.1	-0.3	-0.5
Indexing	2.5	2.6	4.2	-0.1	-1.6
Establishing persistent repositories	2.1	2.2	4.0	-0.1	-1.8
Managing usage rights	1.9	2.0	3.6	-0.1	-1.6
Monitoring access and conditions of use	2.9	3.1	4.0	-0.1	-0.9
Cooperation with and support for users	3.9	4.0	4.2	-0.2	-0.2
Developing plans and strategies for data management within the organisation	3.4	3.3	4.3	0.0	-0.9
Ensuring compliance with applicable legislation	3.5	3.4	4.4	0.2	-1.0
Participation in the development of regulations	3.2	3.2	4.2	0.0	-1.0
Monitoring developments in the field of (open) data management	3.1	2.9	4.3	0.2	-1.3

* Note: n varies across items due to item non-response.

Respondents' perceived abilities were generally closely aligned with their responsibilities, although some discrepancies were observed (Table 2). In several areas, such as metadata management, anonymisation and data collection, respondents reported slightly higher capability than responsibility, suggesting some underutilised capacity. Conversely, slightly lower self-assessed capabilities were observed in more strategic and risk-sensitive areas, including data integrity and

security, legal compliance, opening data and the development of internal data management strategies.

Training needs were consistently high across tasks, particularly in legal compliance, data quality standards, opening data and strategic data management. The largest gaps between current capability and training needs were observed in more specialised areas, especially quality standards, persistent repositories, indexing and usage rights.

Overall, the results indicate a systematic need for competence development, especially in strategic, legal and technically specialised aspects of data stewardship.

Additional question blocks showed that respondents generally reported a high understanding of the data ecosystem in which they operate, with means ranging from 4.9 to 5.2. They also rated their soft skills positively, especially problem-solving (5.0), self-initiative attitude (5.0), analytical thinking (4.9), teamwork (4.8), and adaptability (4.8). By contrast, they rated themselves lower in explaining complex IT content (3.7), inspiring others (4.1), and leadership (4.2).

Finally, although respondents reported relatively strong understanding of new technologies such as artificial intelligence and the Internet of Things (4.8) and a solid awareness of information and cybersecurity concepts (4.5), they rated themselves lower in terms of familiarity with digital and data projects outside their own organisation (4.1) and outside their own sector (4.2). This points to limited cross-organisational and cross-sectoral awareness, suggesting that broader networking and exchange of practice could strengthen stewardship capacity further.

Overall, the results portray open data stewardship in Slovenian public administration as a function that is present but weakly institutionalised, largely operational, unevenly distributed across organisations, and marked by substantial unmet training needs.

4 Conclusion

This study addressed the research question of how (open) data stewardship is organised in Slovenian public administration, particularly in terms of organisational arrangements, employee profiles, responsibilities, competencies and perceived improvement needs.

The findings show that (open) data stewardship is present across a wide range of public administration organisations but remains weakly institutionalised. In most cases, stewardship activities are performed as an additional responsibility rather than as a distinct organisational function. The role is most often located within IT or digitalisation departments, although it also appears across various administrative and sector-specific units, reflecting the cross-cutting nature of data stewardship.

The results further show that respondents are primarily engaged in operational tasks related to data collection, verification, storage and user support, while more specialised and strategic responsibilities appear less developed. Respondents identified strengths related to improved data accessibility, transparency and support for decision-making, but also highlighted several improvement needs, including clearer allocation of responsibilities, additional staffing and training, stronger coordination, and improvements in data quality management. Considerable training needs were identified particularly in areas such as legal compliance, data quality standards, mechanisms for opening data and the development of data management strategies.

Overall, the findings point to the need for a more structured and strategic approach to (open) data stewardship in Slovenian public administration, including clearer role definition, stronger institutional support and systematic competence development.

The study contributes to the emerging literature on public sector data governance by providing empirical evidence on data stewardship practices in a national public administration context. However, the findings should be interpreted with caution due to the use of self-reported data and the possibility of multiple respondents from the same organisation. Future research could further explore data stewardship arrangements through comparative and longitudinal approaches.

Acknowledgment

The research was funded by the Ministry of Digital Transformation of the Republic of Slovenia and the Slovenian Research and Innovation Agency (V5-2389 and P5-0093).

References

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., & Ferro, E. (2018). The multiple life cycles of open data creation and use. In Y. Charalabidis, A. Zuiderwijk, & M. Janssen (Eds.), The world of open data (Public Administration and Information Technology, Vol. 28). Springer. https://doi.org/10.1007/978-3-319-90850-2_2

DataGalaxy. (2023). 8 Pillars Of A Successful Data Governance Framework. https://www.datagalaxy.com/en/blog/8-pillars-of-data-governance-frameworks/

Demchenko, Y. (2016). EDISON data science competence framework (CF-DS) and data science body of knowledge (DS-BoK). Research Data Alliance. https://www.rd-alliance.org/wp-content/uploads/2024/05/rda7-ig-etrd03-edison-data-competences-v02.pdf

European Commission. (2020). A European strategy for data. European Commission. https://digitalstrategy.ec.europa.eu/en/policies/strategydata

European Commission. (2021). 2030 Digital compass: The European way for the digital decade. European Commission.

Feng, Y., & Richards, L. (2018). A review of digital curation professional competencies: Theory and current practices. Records Management Journal, 28(1), 62–78. https://doi.org/10.1108/RMJ-09-2016-0022

Fitsilis, P., Damasiotis, V., Dervenis, C., Kyriatzis, V., & Tsoutsa, P. (2024). Effective data stewardship in higher education: Skills, competences, and the emerging role of open data stewards. arXiv. https://doi.org/10.48550/arXiv.2410.20361

Hitz-Gamper, B. S., Neumann, O., & Stürmer, M. (2019). Balancing control, usability and visibility of linked open government data to create public value. International Journal of Public Sector Management, 32(5), 451–466. https://doi.org/10.1108/IJPSM-02-2018-0062

Hudson-Vitale, C., Imker, H., Johnston, L. R., Carlson, J., Kozlowski, W., Olendorf, R., & Stewart, C. (2017). SPEC kit 354: Data curation. Association of Research Libraries.

Jablonski, J. (2022). Identifying and engaging data stewards. Pythian. https://www.pythian.com/blog/business-insights/identifying-and-engaging-data-stewards

Manu, T. R., & Gala, B. (2021). Data curation activities in research data repositories: Best practices. In Proceedings of the International Conference on Science and Technology Research Data Analysis (ICSTRDA 2021) (pp. 43–52).

Murillo, A. P., & Yoon, A. (2024). Trends in data curation: Competencies and tools. Proceedings of the Association for Information Science and Technology, 61(1), 1047–1050. https://doi.org/10.1002/pra2.1181

OECD. (2020). Open, useful and re-usable data (OURdata) index: 2019. OECD Publishing.

OECD. (2023). Data stewardship, access, sharing and control: A Going Digital III module synthesis report. OECD.

Ruijer, E., Grimmelikhuijsen, S., van den Berg, J., & Meijer, A. (2020). Open data work: Understanding open data usage from a practice lens. International Review of Administrative Sciences, 86(1), 3–19.

Tammaro, A. M., & Casarosa, V. (2014). Research data management in the curriculum: An interdisciplinary approach. Procedia Computer Science, 38, 138–142. https://doi.org/10.1016/j.procs.2014.10.023

Tammaro, A. M., Matusiak, K. K., Sposito, F. A., & Casarosa, V. (2019). Data curator's roles and responsibilities: An international perspective. Libri, 69(2), 89–104. https://doi.org/10.1515/libri-2018-0090

The Turing Way Community. (2022). The Turing Way: A handbook for reproducible, ethical and collaborative research (Version 1.0.2). Zenodo. https://doi.org/10.5281/zenodo.3233853

Wirtz, B. W., Weyerer, J. C., Becker, M., & Müller, W. M. (2022). Open government data: A systematic literature review of empirical research. Electronic Markets, 32, 2381–2404. https://doi.org/10.1007/s12525-022-00582-8

About the authors

Tina Jukić is an Associate Professor in the field of Informatics in Public Administration at the Faculty of Public Administration, University of Ljubljana. Her research focuses on various aspects of the digital transformation of the public sector, with particular emphasis in recent years on technologies supporting co-creation with public service users and on open public sector data. She has led and participated in numerous national and international research projects and regularly presents and publishes the results of her research at reputable national and international conferences and in scientific journals.

Sanja Vrbek is an Assistant Professor in the field of Public Administration at the Faculty of Public Administration, University of Ljubljana. Her research focuses on co-creation and innovative practices in the public sector, with particular emphasis on digital solutions that support collaboration between public organisations and their stakeholders. She actively participates in several national and international research projects, including projects addressing the impact of digitalisation and organisational change in public administration and the development of innovative tools for co-creation. She regularly presents the results of her research at scientific and professional conferences and publishes in national and international academic publications.

Damijana Keržič is an Assistant Professor of Informatics in Public Administration at the Faculty of Public Administration, University of Ljubljana. She received her PhD in Social Informatics from the Faculty of Social Sciences, University of Ljubljana, in 2022. In her doctoral dissertation, she focused on learning analytics in blended learning in higher education. Her research interests include e-learning, educational data mining, and educational technologies in higher education. She regularly participates in various research projects related to the digitalisation of the learning process. Her work has been published in professional and scientific articles and presented at numerous academic and professional conferences.

Integrating AI into Small Firm Knowledge Work: A Qualitative Case Study

Tanja Korhonen, Heli Kemppainen
Kajaani University of Applied Sciences, Master School, Kajaani, Finland
tanja.korhonen@kamk.fi, heliekemppainen@gmail.com

The study investigates how artificial intelligence can be integrated into the knowledge work of a micro-enterprise to improve productivity, efficiency and work well-being. A longitudinal single-case design, supported by a master's thesis and follow-up interview, is used to examine how lightweight readiness practices, human-AI task division and iterative workflow adjustments shape adoption under resource constraints. The experimental part was conducted qualitatively within a single firm. The findings indicate that AI can reduce the time spent on preparatory tasks and, in doing so, ease deadline pressure. At the same time, expert oversight is maintained through explicit checkpoints, which appear to help mitigate risks related to automation bias. In this case, the analysis sheds light on AI readiness at a micro level, particularly in how work moves along the assistive-to-advisory spectrum in human-AI collaboration and toward more AI-supported workflows. It also offers practical guidance for small knowledge work firms that are looking for realistic and responsible ways to begin using AI.

DOI
https://doi.org/10.18690/um.fov.4.2026.37

ISBN
978-961-299-152-4

Keywords:
AI-integration, micro-enterprise, knowledge work, organisational readiness, human-AI collaboration

University of Maribor Press

1 Introduction

Artificial intelligence (AI) is spreading quickly in knowledge-intensive work. It offers clear potential to improve productivity and efficiency, but it also raises questions about its effects on employees and whether organisations are ready to adopt it. For micro and small companies, the potential is significant, but so are the challenges. These firms often work with limited resources, fragmented data and a limited capacity for change.

Prior literature (Iansiti & Lakhani, 2020; Colson, 2019) describes AI as a general-purpose technology that can reshape how organisations operate, learn and make decisions. At the same time, it stresses that benefits do not come from technology alone. They depend on having the right capabilities and on integrating AI carefully into everyday work. In this article, AI mainly refers to generative AI.

We examine the research question (RQ): *How can Artificial Intelligence be effectively integrated into the knowledge work of a micro-enterprise to improve productivity and work well-being?* Knowledge work is understood here as the acquisition, creation, combination or application of knowledge (Davenport et al., 1996). The question is timely and responds to a clear gap in the literature. Most existing research focuses on large organisations (Enholm et al., 2022; Jöhnk et al., 2021), while small and medium-sized enterprises (SMEs, see European Commission, 2003) face different challenges. These include lower levels of AI readiness, which can put them at a disadvantage (Yang et al., 2024). At the same time, SMEs form a large part of many economies and account for a significant share of knowledge-intensive work, which makes their AI adoption both economically and socially important.

Generative AI has recently made advanced AI tools easier to access through simple interfaces and cloud services. This makes the RQ especially relevant for small firms, as these changes can lower the barriers to adoption (Brynjolfsson et al., 2023; Longueville et al., 2025). However, there is still limited empirical research on how small firms can use these tools effectively. Most studies focus on large organisations or specific technical solutions rather than broader ways of integrating AI into everyday work (Russo, 2024; Weinberg, 2025).

Including work well-being alongside productivity and efficiency responds to a growing concern in the literature. While AI can improve performance, it may also increase stress, mental workload or job insecurity if it is not introduced carefully (Neumann et al., 2022; Radhakrishnan et al., 2022).

AI often improves productivity by reducing routine tasks and making work more efficient, although most evidence comes from large organisations (Soulami et al., 2024). In these settings, structured implementation and greater resources tend to strengthen the gains (Brynjolfsson et al., 2023; OECD, 2023). In SMEs, AI can also improve performance and efficiency, but the results depend more on factors such as skills, training and organisational support. At the same time, limited resources and other adoption barriers still play a significant role (Sánchez et al., 2025; Wahab & Radmehr, 2024; Fan et al., 2025; Soomro et al., 2024).

Research on micro-enterprises is still scarce. Existing studies focus mainly on conditions for adoption rather than outcomes, suggesting that AI use is shaped by entrepreneurial orientation and technological readiness. Evidence on productivity and well-being effects remains limited and mostly indirect (Nevi et al., 2025), which reflects a broader gap noted in earlier reviews (Soulami et al., 2024; OECD, 2023).

This RQ bridges theory and practice by examining not just whether AI creates value, but *how* it can be integrated in resource-constrained contexts. Existing frameworks for AI readiness and business value (Enholm et al., 2022; Jöhnk et al., 2021) have been developed primarily from large-firm perspectives and require adaptation to small-firm realities. Through this longitudinal case study, we provide empirical insights into how small firms can structure practices and workflows to gain AI benefits while preserving human-centred work.

The study is based on two empirical sources from the same micro-enterprise in professional services. First, we draw on a master's thesis from 2025 that examined AI adoption in a micro-sized firm (Kemppainen, 2025). This was followed by a semi-structured interview with the same firm in 2026, which focused on outcomes, challenges and effects on well-being one year after the initial recommendations. These materials are considered alongside existing research on the business value of AI, organisational and AI readiness, human-AI collaboration in knowledge work, and the workforce effects of machine learning (ML).

This article first reviews prior research on AI value creation, readiness, human-AI task allocation, decision processes, and small-firm AI adoption challenges. It then outlines the research design, presents results integrating empirical insights with theory into a practical model for resource-constrained knowledge work, and concludes with implications, limitations, and future research directions.

2 Previous research

Research on AI in business has shifted from focusing mainly on technology to looking at how organisational capabilities and practices shape outcomes. Studies have examined what supports or hinders AI adoption, as well as its direct effects, such as better decisions and more efficient processes, and broader impacts on performance and sustainability. They also show that many organisations struggle to create value because of skill gaps, resistance to change and unclear goals (Enholm et al., 2022). Research also distinguishes between automation, where tasks are replaced, and augmentation, where AI supports human work. It shows that value depends not only on the technology but also on factors such as management support, strategy and organisational culture (Radhakrishnan et al., 2022).

In SMEs, studies often combine the Technology-Organisation-Environment (TOE) framework, with Diffusion of Innovations (DOI) theory, which explains how new ideas and technologies spread over time. Using these perspectives, research identifies common barriers such as limited data access, lack of skills, cultural resistance and weak governance, and suggests step-by-step ways to support adoption (Sánchez et al., 2025).

Employees may worry about deskilling or job loss when new AI tools are introduced. This highlights the need for clear change management and for presenting AI as a support rather than a threat (Mohd Rasdi & Umar Baki, 2025).

Readiness is often used to explain how organisations adopt digital technologies and AI. Earlier research describes it as a set of capabilities and shows that missing these capabilities is a common reason why initiatives fail. Because of this, readiness should be assessed before moving to wider use (Lokuge et al., 2019).

More recent AI-focused studies identify practical factors that matter, such as data quality, model transparency, skills and governance. These extend the traditional TOE view to better fit AI contexts (Jöhnk et al., 2021; Pumplun et al., 2019). A common finding is that readiness and adoption develop together over time and strengthen each other (Jöhnk et al., 2021).

For SMEs, readiness challenges are often more pronounced. Research shows that smaller firms tend to have lower levels of AI readiness, which can make it harder for them to deal with adoption challenges than for larger organisations (Yang et al., 2024). At the same time, recent studies suggest that agile firms can build practical readiness through lighter approaches, such as cloud-based SaaS tools, flexible structures and small-scale pilots that are developed over time (Oluwamola et al., 2025).

Evidence from a boutique branding consultancy also shows that generative AI can support everyday work by speeding up content creation, reducing managerial workload and gradually strengthening the organisation's knowledge base (Yurong, 2024). These findings offer a useful example for other SMEs. Overall, this supports the idea that creating value from AI often requires changes to the business model as well (Åström et al., 2022).

Human-AI symbiosis research shows that AI supports human cognition in complex, data-rich tasks, while humans provide contextual, intuitive, and ethical judgment under uncertainty. The dominant view is one of intelligence augmentation rather than substitution (Jarrahi, 2018). Decision-making studies similarly advocate a shift from "data-driven" workflows, where humans perform most processing, to "AI-driven" workflows in which machines handle data-intensive analysis and generate options, and humans focus on deliberation guided by values, strategy, and context (Colson, 2019; Iansiti & Lakhani, 2020).

In terms of the workforce, machine learning is often seen as a general technology that changes work rather than simply replacing it. Jobs consist of many different tasks, and not all of them are suitable for automation. As a result, work is changing and expanding rather than disappearing, although the effects are uneven (Brynjolfsson & Mitchell, 2017). From a management perspective, AI is often built

on earlier data and analytics capabilities. Organisations can move forward more easily if they build on these existing strengths, if skills, governance and data practices develop at the same time (Davenport, 2018).

Small companies can unlock value from AI if they adopt a deliberate approach to its adoption, ensuring that the foundations for its use are in place. They should also design human-AI task allocations that protect well-being and embed AI into decision workflows with clear guidelines for accuracy, transparency, and responsibility (Horvat et al., 2023; Uren et al., 2023; Cheong, 2024).

3 Methodology

This study brings together two empirical sources from the same micro-sized professional services company and combines them with a focused review of the literature. One source is a master's thesis that examines AI use in a knowledge work firm with fewer than ten employees using an action research approach (Kemppainen, 2025). The article itself is based on a single case study, which is suitable for exploring how and why things happen in real organisational settings where events cannot be controlled (Yin, 2019). The case also makes it possible to follow AI use over time in a micro-enterprise and to capture changes and emerging practices.

The semi-structured interview was conducted in February 2026 with the firm's owner-manager to understand how the earlier recommendations had been implemented. The discussion covered areas such as task types, tools in use, changes in workflows, perceived productivity, error patterns, governance and effects on well-being. The interview lasted 53 minutes and was transcribed immediately afterwards.

The analysis followed an abductive approach, moving between the data and existing theory. In practice, this meant combining inductive observations with theoretical interpretation and revisiting both as the analysis progressed (Thompson, 2022; van Hulst et al., 2025). This approach made it possible to use established concepts while still paying attention to new and unexpected insights from the case.

Two established frameworks were used to make sense of the empirical material. First, the AI readiness framework by Jöhnk et al. (2021) helped us look at the firm's preparedness across five areas: strategy, culture, organisation, technology and data.

It allowed us to see which areas were already in place, which still needed work, and how a small firm's more lightweight approach compared to the framework, which was originally developed with larger organisations in mind. Through this lens, elements such as the firm's simple policy ("AI-draft, human-decide"), its prompt library and its choice of tools could be seen as practical, scaled-down versions of strategic, cognitive and IT readiness.

Second, the framework by Enholm et al. (2022) was used to analyse outcomes and how value was created. It distinguishes between factors that support or hinder adoption, short-term effects such as better decisions and more efficient processes, and broader impacts on performance and sustainability. It also stresses that value comes not only from technology but from how it is combined with organisational capabilities and change management. Using this framework, we were able to link the firm's reported productivity gains, changes in error patterns and effects on well-being to the different dimensions of the model.

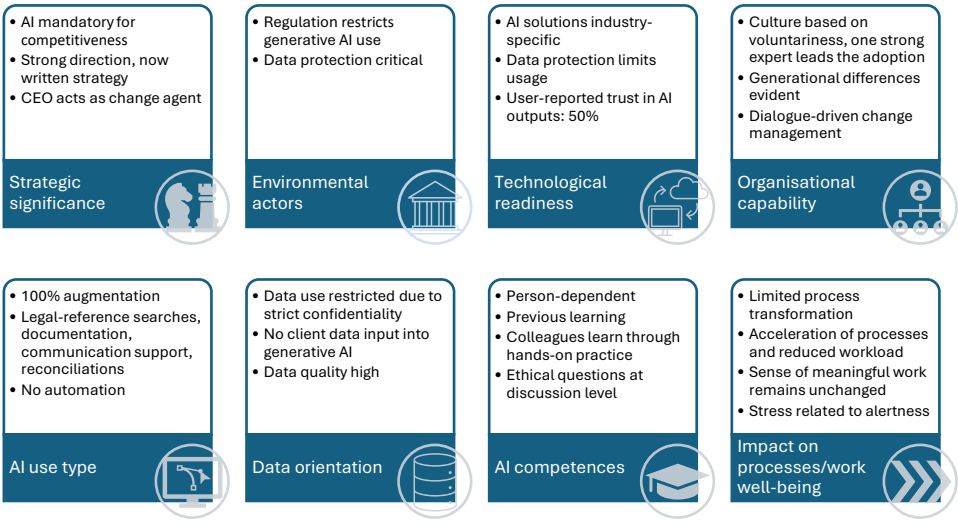

Figure 1: Theme categories in integrating AI into micro firm knowledge work
Source: Own

The analysis moved back and forth between the data and the theoretical frameworks. The interview transcript was first coded inductively, focusing on themes such as tasks, tools, workflows, outcomes and challenges. These themes were then compared with the concepts in the frameworks by Jöhnk et al. (2021) and Enholm

et al. (2022), which helped to identify both matches and differences. From this process, eight theme categories were formed, as shown in Fig. 1: Strategic significance, Environmental actors, Technological Readiness, Organisational capability, AI use type, Data orientation, AI competences and Impact on processes/work well-being.

This study has some limitations. It focuses on a single case, relies on self-reported well-being, and productivity. The findings relate to one professional services micro-enterprise in a specific national context, so they should be applied to other settings with caution. However, the year-long follow-up and the use of established theoretical frameworks provide a solid basis for broader interpretation and further theory development (Yin, 2019).

4 Results

The results show that AI use in the micro-enterprise developed step-by-step over twelve months. Readiness was supported by simple governance practices and gradual upskilling, along with clearer data and processes that made tasks easier to handle with AI. Work was structured so that drafting and validation were kept separate, with humans responsible for checking outputs. Over time, AI outputs were gradually included in decision-making, with clear human checkpoints in place. The findings are presented in terms of productivity and efficiency, work well-being, and integration practices.

The firm has primarily implemented generative AI in the following areas: document drafting, templated email responses, client memos, and first-pass reconciliations of routine bookkeeping anomalies. Based on interview data rather than direct time measurement, the greatest gains were reported in preparatory work, with staff estimating notable time savings in drafting standard communications and some improvements in preparing checklists and workpapers. In contrast, core accounting judgments and final filings were human led, with AI providing alternative explanations for consideration. This distribution of labour was consistent with the thesis roadmap, which emphasised the removal of administrative friction from expert time while preserving human authority over compliance risk and client-specific nuances (Kemppainen, 2025).

These patterns reflect a shift in how tasks and decisions are organised. AI takes on most of the computational work, while humans focus on interpretation, context and ethical judgement. How well this works depends not only on the technology, but also on how workflows are redesigned and how clearly roles and responsibilities are defined (Jarrahi, 2018; Colson, 2019).

The firm's experience is similar to findings from other knowledge-intensive settings. In these contexts, generative AI has helped to speed up information creation and reduce managerial workload, but it still requires clear human oversight (Yurong, 2024).

It was observed that error patterns underwent a shift rather than a complete disappearance. The company reported a decrease in typographical and formatting errors, but occasional "plausible-but-wrong" statements were noted when prompts were ambiguous or when AI extrapolated outdated regulation. To reduce the risk of automation bias, the firm introduced an explicit "two-layer check" for outputs to external parties. This process was as follows: first, the drafter verified content against source documents; second, a risk-weighted review was applied for tax and compliance claims.

This response reflects the concept of readiness work, which places significant emphasis on governance, data quality, and transparent indicators for responsible adoption. It is also consistent with the literature on workforce management, which asserts the importance of maintaining human involvement in situations where the suitability of tasks for ML is only partial (Jöhnk et al., 2021; Brynjolfsson & Mitchell, 2017).

Based on interview data rather than measurable indicators, the effects on work-related well-being were mostly positive. Employees reported reduced deadline pressure and greater control over tasks, although prompt design and monitoring of AI errors remained initial stressors that eased over time. The mixed effects line up with earlier findings that AI can either reduce or increase strain depending on process design and maturity (Enholm et al., 2022).

4.1 Theoretical implications

The case shows that AI readiness in micro enterprises can develop in a simple, practice-based way rather than through a formal program, while still covering different aspects and evolving over time (Lokuge et al., 2019; Jöhnk et al., 2021; Pumplun et al., 2019). The findings also suggest that productivity and well-being improve when AI supports work step-by-step and humans remain in control through clear checkpoints (Jarrahi, 2018).

The case further shows a gradual move toward AI-supported workflows, where AI generates options and humans make the final decisions. This is in line with earlier research on AI-driven operating models and the role of governance in creating value (Colson, 2019; Iansiti & Lakhani, 2020; Enholm et al., 2022; Davenport, 2018). Together, these observations help describe how readiness and human-AI collaboration take shape in a small firm with limited resources and highlight the importance of clear checkpoints and governance in making AI use sustainable.

4.2 Practical implications

For managers in micro and small knowledge-work firms, the following steps offer a structured approach to AI adoption (see Fig.2).

Plan and select. Assess the current state and define strategic objectives (Sánchez et al., 2025). Use a concise readiness canvas to define use cases, expected outputs, verification rules, and data sources. This supports iterative and pragmatic AI adoption in SMEs (Horvat et al., 2023; Kemppainen, 2025; Weinberg, 2025). Managers can start with standardised, high-volume tasks and extend AI use gradually as skills and governance practices develop. Prioritise high-impact, low-complexity projects and choose the right tools and partners. (Agbaakin, 2025; Kemppainen, 2025; Longueville et al., 2025; Sánchez et al., 2025)

Pilot and evaluate. Start with small-scale pilots to validate AI before full implementation (Sánchez et al., 2025). Reduce automation bias by requiring source cues for AI-generated factual statements. This strengthens verification practices and supports responsible AI use. (Brynjolfsson et al., 2023; Enholm et al., 2022; Kemppainen, 2025) Define success metrics (KPIs) (Agbaakin, 2025).

Upskill and encourage. Embed learning into daily work using the "see one, do one, teach one" micro-training approach, and encourage a culture of experimentation and positive engagement with AI (Jöhnk et al., 2021; Kemppainen, 2025; Lokuge et al., 2019). Apply change-management practices that reduce resistance, clarify responsibilities and reinforce that AI is intended to augment rather than replace human capabilities. (Sánchez et al., 2025)

Measure and scale. Use KPIs to evaluate AI impact and scale accordingly (Sánchez et al., 2025).

These steps operationalise readiness and human-AI symbiosis principles associated with sustainable value creation (Davenport, 2018; Enholm et al., 2022; Jöhnk et al., 2021).

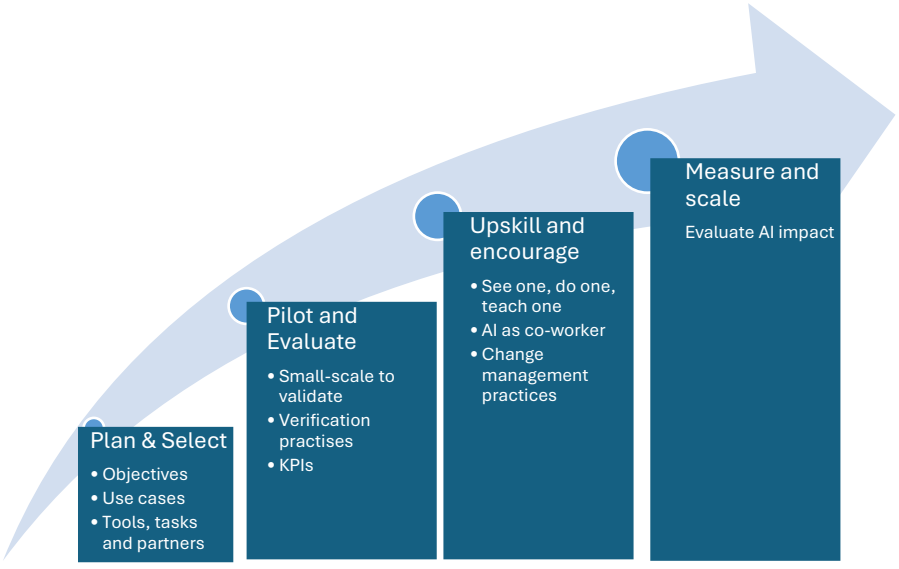

Figure 2: Approach to AI adoption in micro or small firms
Source: Own

5 Discussion and conclusions

The findings show that integrating AI into small-firm knowledge work is challenging yet achievable. The case shows that a small firm can improve efficiency by adopting AI gradually and with a human focus. This interpretation echoes workforce research that recognises major transformation without assuming simple substitution (Brynjolfsson & Mitchell, 2017) and with managerial guidance promoting a gradual, capability-based shift from analytics to AI (Davenport, 2018).

In our case, the micro-enterprise shows readiness in a practical way rather than through a formal program, while still achieving clear benefits. This supports the scalability of readiness (Lokuge et al., 2019) and relates with AI-specific readiness indicators (Jöhnk et al., 2021; Pumplun et al., 2019). It further suggests that lightweight mechanisms such as simple policies, prompt libraries, and customized tools can constitute functional readiness in resource-constrained contexts.

The case adds to human-AI collaboration research by showing how support roles can develop into more advisory roles, and how clear safeguards against automation bias help maintain trust and well-being (Jarrahi, 2018). It supports the proposition that AI-driven operating architectures can be realized incrementally in small firms when decision processing shifts are matched by clear human checkpoints (Colson, 2019; Iansiti & Lakhani, 2020). In this micro-enterprise context, data governance emerged as a central constraint. Data-related limitations shape feasible AI use cases more than technical capabilities do. A related observation is that data-driven innovation across European firms is strongly conditioned by concerns over fair data use, ecosystem governance and regulatory boundaries (Heikkilä et al., 2023).

Our case firm's move to AI-generated previews suggests that the evolutionary path observed in large enterprises (Davenport, 2018) is also achievable in micro-enterprises when governance, skills, and workflows develop in parallel. This contributes to emerging work on generative AI in professional services by illustrating how large-firm frameworks can be adapted to small-firm contexts (Russo, 2024; Sánchez et al., 2025; Weinberg, 2025).

Small knowledge-work firms can begin with high-volume, low-risk drafting tasks. Formalise "AI-draft, human-decide" as policy, standardise prompts, and measure cycle time. Gradually extend into advisory uses where AI proposes alternatives and explanations that humans evaluate. They should invest in data hygiene and documentation, require traceability for factual statements, and analyse AI mistakes to update prompts and checklists. These steps help connect high-level AI goals with everyday work.

The study has limitations. It is a single-case design in one national and professional context with self-reported productivity estimates and qualitatively assessed well-being effects, without direct quantitative measurement (e.g., time-tracking or validated well-being indicators). While single-case studies enable analytical generalization to theoretical propositions (Yin, 2019), the findings' transferability to other sectors, firm sizes, or cultural contexts requires empirical validation. Future research should therefore validate these findings using quantitative approaches, such as time-and-motion studies, digital-trace data, or standardized well-being measures, in addition to examining how the assistive-to-advisory pattern of human-AI collaboration differs across sectors and how specific readiness subdimensions influence outcomes in micro-enterprises. It should also explore how the minimal readiness bundle scales as firms grow or adapt to changing regulations such as the EU AI Act (European Union, 2024) and clarify construct boundaries by linking digital-innovation readiness with AI-readiness frameworks (Lokuge et al., 2019; Jöhnk et al., 2021; Pumplun et al., 2019).

Effective AI integration in small-firm knowledge work is mainly an organisational design task. Building a minimal readiness bundle, dividing work so that humans and AI each use their strengths, and adding safeguards against automation bias allow firms to improve productivity and efficiency without harming well-being. This case shows that such progress is achievable even with limited resources when adoption is deliberate, iterative and human-centred.

Endnote

The authors acknowledge the use of artificial intelligence tools in supporting the literature search, improving the English language and formulation of the text, and assisting in the proofreading of the manuscript.

Acknowledgment

This publication is part of the project AIRA – AI Solutions for Sustainable and Productive Knowledge Work, co-funded by the European Union. The authors would like to especially thank Antti Rimpiläinen (Kajaani University of Applied Sciences) and Mira Kekkonen (Oulu University of Applied Sciences) for their contributions to identifying relevant literature.

Co-funded by the European Union

References

Agbaakin, O. (2025). Leveraging Artificial Intelligence as a Strategic Growth Catalyst for Small and Medium-sized Enterprises. Computers and Society. Cornell University. https://doi.org/10.48550/arXiv.2509.14532

Åström, J., Reim, W., & Parida, V. (2022). Value creation and value capture for AI business model innovation: A three-phase process framework. Review of Managerial Science, 16(7), 2111–2133. https://doi.org/10.1007/s11846-022-00521-z

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work (NBER Working Paper No. 31161). National Bureau of Economic Research. https://doi.org/10.3386/w31161

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. Science, 358(6370), 1530–1534. https://doi.org/10.1126/science.aap8062

Cheong, B.C. (2024). Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. Front. Hum. Dyn. 6:1421273. doi: 10.3389/fhumd.2024.1421273

Colson, E. (2019). What AI-driven decision making looks like. Harvard Business Review. https://hbr.org/2019/07/what-ai-driven-decision-making-looks-like

Davenport, T. H., Jarvenpaa, S., & Beers, M. (1996). Improving knowledge work processes. MIT Sloan Management Review, Summer 1996.

Davenport, T. H. (2018). From analytics to artificial intelligence. Journal of Business Analytics, 1(2), 73–80. https://doi.org/10.1080/2573234X.2018.1543535

De Longueville, B., Sanchez, I., Kazakova, S., Luoni, S., Zaro, F., Daskalaki, K., & Inchingolo, M. (2025). The Proof Is in the Eating: Lessons Learnt from One Year of Generative AI Adoption in a Science-for-Policy Organisation. AI, 6(6), 128. https://doi.org/10.3390/ai6060128

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial intelligence and business value: A literature review. Information Systems Frontiers, 24(5), 1709–1734. https://doi.org/10.1007/s10796-021-10186-w

European Commission. (2003). SME definition. Internal Market, Industry, Entrepreneurship and SMEs. https://single-market-economy.ec.europa.eu/smes/sme-fundamentals/sme-definition_en

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689, 1–258. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Fan, X., Soomro, M. A., Soomro, S. A., Sohu, J. M., & Shah, N. (2025). Artificial intelligence adoption and technostress in manufacturing SMEs. Journal of Manufacturing Technology Management. https://doi.org/10.1108/jmtm-02-2025-0120

Heikkilä, M., Ahmad, F., & Heikkilä, J. (2023). Unlocking the potential of data-driven business models: An empirical investigation into the role of ecosystems and fair data use. Bled eConference Proceedings. https://doi.org/10.18690/um.fov.6.2023.6

Horvat, D., Kroll, H., & Jäger, A. (2023). AI readiness: An integrated socio-technical framework. In HCI International 2023 Posters (pp. 539–546). Springer. https://doi.org/10.1007/978-3-031-36121-0_69

Iansiti, M., & Lakhani, K. R. (2020). Competing in the age of AI: Strategy and leadership when algorithms and networks run the world. Harvard Business Review Press.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. Business Horizons, 61(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Jöhnk, J., Weißert, M., & Wyrtki, K. (2021). Ready or not, AI comes—An interview study of organizational AI readiness factors. Business & Information Systems Engineering, 63(1), 5–20. https://doi.org/10.1007/s12599-020-00676-7

Kemppainen, H. (2025). Utilizing Artificial Intelligence in the Expert Work of a Small Accounting Firm. Master's thesis. Kajaani University of Applied Sciences.

Lokuge, S., Sedera, D., Grover, V., & Xu, D. (2019). Organizational readiness for digital innovation: Development and empirical calibration of a construct. Information & Management, 56(3), 445–461. https://doi.org/10.1016/j.im.2018.09.001

Mohd Rasdi, R., Umar Baki, N. (2025). Navigating the AI landscape in SMEs: Overcoming internal challenges and external obstacles for effective integration. PLOS ONE 20(5): e0323249. https://doi.org/10.1371/journal.pone.0323249

Neumann, O., Guirguis, K., & Steiner, R. (2022). Exploring artificial intelligence adoption in public organizations: A comparative case study. Public Management Review, 26(1), 114–141. https://doi.org/10.1080/14719037.2022.2048685

Nevi, G., Pizzichini, L., Bastone, A., & Dezi, L. (2025). Adoption of artificial intelligence by micro and small health enterprises: effects of entrepreneurial orientation on the TOE model. European Journal of Innovation Management, Emerald Group Publishing Limited, vol. 28(11), pp. 209-241.

OECD. (2023). The impact of AI on the workplace: Evidence from OECD case studies of AI implementation. OECD Publishing. https://doi.org/10.1787/2247ce58-en

Oluwamola, O., Adekunle, A., & Ogunleye, B. (2025). Business transformation through AI adoption in agile enterprises: Design methodologies, architectures, deployment scenarios, governance, and future directions. International Journal of Scientific Research and Management, 13(3), 1–18. https://doi.org/10.18535/ijsrm/v13i03.ec07

Pumplun, L., Tauchert, C., & Heidt, M. (2019). A new organizational chassis for artificial intelligence: Exploring organizational readiness factors. In Proceedings of the 27th European Conference on Information Systems (ECIS).

Radhakrishnan, R., Duan, Y., & Soylu, A. (2022). Understanding organizations' artificial intelligence journey: A qualitative approach. Pacific Asia Journal of the Association for Information Systems, 14(6), 1–34. https://doi.org/10.17705/1pais.14602

Russo, D. (2024). Navigating the complexity of generative AI adoption in software engineering. ACM Transactions on Software Engineering and Methodology, 33(6), Article 154. https://doi.org/10.1145/3652154

Sánchez, E., Calderón, R., & Herrera, F. (2025). Artificial intelligence adoption in SMEs: Survey based on TOE–DOI framework, primary methodology and challenges. *Applied Sciences, 15*(12), 6465. https://doi.org/10.3390/app1512646

Soomro, B. A., Fan, X., Sohu, J. M., Soomro, S. A., & Shaikh, F. A. (2024). Organizational support and AI-related anxiety in SMEs. *Kybernetes*. https://doi.org/10.1108/k-07-2024-1889

Soulami, B., Benchekroun, H., & Galiulina, E. (2024). Human-centered AI adoption and worker productivity: A systematic review. *Frontiers in Artificial Intelligence*. https://doi.org/10.3389/frai.2024.1473872

Thompson, J. (2022). A Guide to Abductive Thematic Analysis, *The Qualitative Report,* vol. 27, no. 5, pp. 1410-1421.

Uren, V., Edwards, J. S., & Jiménez, A. (2023). Technology readiness and the organizational journey towards AI adoption: An empirical study. International Journal of Information Management, 68, Article 102588. https://doi.org/10.1016/j.ijinfomgt.2022.102588

van Hulst, Merlijn, & E. L. Visser. (2025). Abductive Analysis in Qualitative Research. Public Administration Review 85(2), pp. 567–580.

Wahab, M. D. A., & Radmehr, M. (2024). The impact of AI assimilation on firm performance in small and medium-sized enterprises. *Heliyon, 10*(8), e29580. https://doi.org/10.1016/j.heliyon.2024.e29580

Weinberg, B. D. (2025). A framework for the adoption and integration of generative AI in midsize organizations and enterprises (FAIGMOE). Journal of Business Research, 170, 114–128.

Yang, L., Holtz, D., Jaffe, S., Suri, S., Teevan, J., Iqbal, S., & Horvitz, E. (2024). Artificial intelligence adoption in a professional service industry: A multiple case study. Technological Forecasting and Social Change, 199, Article 123251. https://doi.org/10.1016/j.techfore.2024.123251

Yin, R. K. (2019). Case study research and applications: design and methods (Sixth edition). SAGE.

Yurong, C. (2024). Unlocking SME potential. In AI and Business Innovation (pp. 45–62). CRC Press. https://doi.org/10.1201/9781003469551-2

Location-Aware Machine Learning to Forecast Urban Foot Traffic

Katayoon Pourmahdi, Bahareh Naseri,
Ilia Gugenishvili, Jozsef Mezei,
Anna-Greta Nyström

Åbo Akademi University/Faculty of Social Sciences, Business and Economics, Turku, Finland
katayoon.pourmahdi@abo.fi, bahareh.naseri@abo.fi, ilia.gugenishvili@abo.fi, jozsef.mezei@abo.fi, anna-greta.nystrom@abo.fi

DOI
https://doi.org/10.18690/um.fov.4.2026.38

ISBN
978-961-299-152-4

Smart cities increasingly rely on intelligent tools to manage urban mobility, reduce congestion, and support evidence-based public governance. This study examines how machine learning (ML) models can support urban mobility management by predicting footfall and multimodal traffic flows. Five ML models were trained and evaluated on a dataset of 261,539 hourly observations collected from a 16-sensor monitoring network across four urban locations in Turku, Finland. The modelling framework integrates temporal, meteorological, socio-economic, and spatial covariates to capture location-specific mobility flows. To prevent temporal leakage, model evaluation used a strict chronological train–test split, supported by five-fold time-series cross-validation; all metrics are reported with empirical standard deviations (±SD). XGBoost recorded the strongest predictive performance (R^2 = 0.720 ±0.071 in the City Centre, 0.839 ±0.099 in the University corridor, and 0.570 ±0.105 in the cultural district). Location 4, a low-volume remote sensor, yielded R^2 = 0.179 ±0.231, reflecting the inherent predictability limits of a near-zero-volume sensor (avg. 2.4 pedestrians/h). The findings guide sensor deployment planning, real-time mobility management, and data infrastructure design in smart city contexts.

Keywords:
smart city,
footfall flows,
mobility management,
machine learning,
neural networks

University of Maribor Press

1 Introduction

Urban governance and management practices evolve rapidly under the influence of various factors, such as climate change, migration, and digitalization (Peráček & Kaššaj, 2026). In the digital age, urban management involves working with large volumes of urban data and requires coordination among authorities, stakeholders, and citizens to co-design and co-produce solutions. The situation places the emphasis on the ability to translate data into actionable insights for effective decision-making (Gil et al., 2026). Technological tools and practices, such as sensors, real-time monitoring, and spatial data analytics, are important in that process (Jamal et al., 2026; Lingani et al., 2019).

Policymakers seeking to improve urban governance and decision-making have adopted machine learning (ML) methods (Jamal et al., 2026) to forecast traffic flow and pedestrian movement patterns (Haghshenas et al., 2026; Koech & Sobhani, 2026). Despite the recent growth in the use of urban data infrastructure, the application of ML to support urban governance remains under-researched across different contexts (Lingani et al., 2019; Wolniak et al., 2024). This gap is particularly evident in the area of urban mobility, which encompasses both pedestrian movements and multimodal traffic. The topic is one of the most complex and resource-intensive facing city administrations. Much of the existing research evaluates the technical performance of ML methods for traffic prediction and overall predictive accuracy (Kim & Chan, 2025; Wu et al., 2024). It typically reports global performance metrics aggregated across cities or sensor networks, with limited attention to model variability across different spatial contexts. As a result, we know relatively little about how and why ML model performance changes across heterogeneous urban contexts and its implications for the design of context-aware mobility prediction systems. An additional research gap concerns the limited translation of predictive outputs, such as footfall forecasts, congestion risks, or short-term demand surges, into actionable insights for urban planners, traffic managers, and public administrators. Although ML methods can generate accurate predictions, few studies connect these outputs to concrete governance applications or real-time operational decision-making.

Concurrently, the consequences of insufficient mobility management are wide-ranging: traffic congestion driven by inefficient traffic control, excess demand, population growth, or weather disruptions continues to degrade air quality, extend travel times, and diminish the quality of life for city residents (Lingani et al., 2019). In contrast, accurate and context-sensitive predictions of footfall and multimodal traffic flows offer opportunities to identify emerging congestion hotspots, optimize traffic control, and support responsive, sustainability-oriented urban mobility strategies.

This study aims to predict footfall and multimodal traffic flows to support effective urban management and decision-making, thereby contributing to safer streets, reduced pollution, and a more efficient use of urban infrastructure. To do so, we apply five ML models to predict day-to-day traffic trends using mobility data alongside meteorological and socio-economic datasets. We report location-specific model performance and identify the key factors driving predictions, with the goal of producing findings that are directly useful to urban planners and policymakers. The following research questions guide this study:

RQ1: In which urban contexts, and with what performance variability, can ML models reliably predict pedestrian footfall and traffic flows?
RQ2: Which contextual factors are the most influential predictors of urban pedestrian and traffic flows, and what do interpretable ML techniques reveal about their relative contributions?

To address these questions, we draw on a multi-sensor dataset of 261,539 hourly observations collected across 16 monitoring devices in Turku, Finland. Predictor variables include short-term temporal lag features, cyclical time-of-day encodings, meteorological conditions (temperature and precipitation), and socio-economic and spatial covariates (household income and distance from the City Centre).

2 Literature Review

ML has broadened the decision-support options available to mobility managers in smart cities. They can now access large volumes of real-time traffic data, road network information, and traveller behaviour patterns (Kim & Chan, 2025). Consequently, city managers can dynamically reallocate traffic flows, reduce

congestion, and enhance overall mobility efficiency (Alhasnawi et al., 2025). A growing body of research recognizes that potential and explores predictive models underpinning smart mobility applications. A smaller (but growing) research strand expressly addresses the science of location-specific or context-stratified mobility prediction, indicating that model performance varies substantially across urban settings (Schläpfer et al., 2021). Comparative studies assessing multiple classification and regression techniques, including deep neural networks (DNNs), support vector machines (SVMs), decision trees, random forests (ensembles of decision trees aggregated by majority vote or averaging), and discrete choice models (Ali, 2025; Fernández-Delgado et al., 2014), establish that SVMs and random forests are consistently among the top-performing classifiers. For example, Fernández-Delgado et al. (2014) demonstrate that tree ensembles and kernel methods rival or outperform deep networks across a broad range of classification tasks. That learning guided the choice of XGBoost (a regularized gradient-boosted tree ensemble) over random forest: XGBoost incorporates L1/L2 regularization and is computationally more efficient on tabular data (Chen & Guestrin, 2016), making it better suited to the sensor-level modelling context. These findings underscore the advantages of non-linear models for capturing the complexity of urban mobility behaviour. Traditional statistical methods, such as logistic and linear regression, often fail to reproduce these dynamics with comparable accuracy (Dabiri & Heaslip, 2018; Fernández-Delgado et al., 2014; Pulugurta et al., 2013).

Recent studies signal the relevance of integrating spatial and temporal relationships in predictive models for footfall and traffic flows, and deep learning for footfall and traffic flow prediction (cf. Asher et al., 2023; Ng et al., 2017). For instance, Koech and Sobhani (2026) and Murcio and Wang (2025) demonstrate variations in pedestrian activity across several locations in Dublin and the UK, emphasizing the need for context-aware modelling. Similarly, Wang et al. (2025) showed that including weather variables alongside temporal features improves short-term traffic flow predictions, with their hybrid CNN–BiLSTM model outperforming five baseline models.

The above studies establish that ML model performance is sensitive to algorithmic design and also to the spatial and socio-economic characteristics of the data environment. Such sensitivity has direct implications for how predictive outputs, particularly footfall estimates, should be interpreted and operationalized in practice

in urban locations. That also underscores the need for models that account for contextual variability to support reliable, actionable decision-making in smart mobility contexts and systems.

3 Modelling

3.1 Data Preprocessing

The data used in this study were gathered in Turku, Finland, from a network of 16 pedestrian and traffic monitoring devices (referred to as device01–device16) deployed across four spatially distinct urban locations. The raw dataset comprises 261,539 hourly observations recorded continuously between March 31 and November 1, 2025. Each data point corresponds to a specific device, measurement scenario (defined here as a unique directional or aggregated counting mode of a sensor, e.g., inbound, outbound, or total flow, as recorded by the device firmware), date, and hour, and includes counts of six traffic categories: pedestrians (Human), cyclists (Bike), cars (Car), buses (Bus), trucks (Truck), and other vehicles (Other Vehicle). Across the full dataset, pedestrian counts exhibit the widest variation (mean = 172.1, SD = 287.2, range: 0–11,466), reflecting strong spatial and temporal heterogeneity in foot-traffic intensity. Car counts (mean = 23.4, SD = 124.2) and bicycle counts (mean = 8.9, SD = 24.6) show comparatively lower but still substantial variance, while counts for buses, trucks, and other vehicles are sparse.

In addition to traffic counts, each record includes contextual covariates: hourly air temperature (range: −6.1 to 29.9 °C) and precipitation (range: 0–10.6 mm) sourced from Finnish Meteorological Institute (FMI) data, the average household income of the surrounding postcode (range: €29,878–€42,046)[1], and the geodesic distance of each device from the City Centre (range: 0.05–1,529 m). The 16 monitoring devices generated sixty-four directional and aggregated scenarios, enabling both direction-specific and total-flow analyses. For modelling purposes, the devices were grouped into four locations based on geographic proximity and urban function: Location 1 (City Centre / Tori, devices01 and 03–12, n = 128,194 records after preprocessing), Location 2 (Hämeenkatu / University corridor, device02, n = 10,259), Location 3

[1] Average rather than median income is reported here because Statistics Finland publishes household income at the postal-code level as an arithmetic mean. We acknowledge that right-skewed income distributions may cause this proxy to overweight high earners.

(Logomo / Ratapihankatu cultural district, devices13–15, n = 25,337), and Location 4 (remote building entrance, device16, n = 2,263).

Before model training, temporal lag features (lags 1, 2, and 3 of pedestrian, car, bus, truck, and bicycle counts) and cyclical hour encodings (sine and cosine transformations of the hour-of-day variable) were appended to each record. Missing values were present in two forms: (i) structurally missing lag values, arising inevitably from lag construction at the beginning of each device's time series (the first three records per device lack valid lag-1, lag-2, and lag-3 observations); and (ii) sporadic missing values in the temperature variable, attributable to sensor recording gaps, which were imputed using the column median before lag construction. All remaining records with any missing value were removed via listwise deletion. This preprocessing yielded a final analytical dataset of 166,053 complete observations (a retention rate of 63.5% of the raw 261,539 records), which was used for all model training and evaluation.

3.2 Input Features

A total of twenty-eight input features were constructed and used for model training. They were organized into five conceptually distinct groups: temporal, lag-based, meteorological, socio-economic, and spatial. Table 1 provides a full list of feature names, group membership, and a brief description of each feature.

3.3 Model Development

We addressed the two research questions by selecting five ML models to span a range of algorithmic complexity and interpretability. The selection includes a parametric linear baseline (linear regression), a shallow non-linear tree-based model (Decision Tree), a powerful gradient-boosting ensemble (XGBoost), a fully connected artificial neural network (ANN), and a recurrent sequential architecture (LSTM). The following subsections describe the architecture and training configuration of each model in detail.

Linear Regression: parametric baseline assuming a linear relationship between features and traffic counts; used to quantify the non-linearity gain of more complex models.

Table 1: Complete list of input features used for model training (n = 28)

Feature Groups	Feature Names	n	Description
Temporal (cyclical)	Hour_sin, Hour_cos	2	Sine and cosine transformations of the hour-of-day variable (0–23) to encode circular time continuity.
Temporal (categorical)	Weekdays	7	Binary dummy variables for each day of the week (one-hot encoded, 7 columns).
Pedestrian	Human_lag1, Human_lag2, Human_lag3	3	Pedestrian counts at t−1, t−2, and t−3 hours capture short-term autocorrelation in foot traffic.
Car	Car_lag1, Car_lag2, Car_lag3	3	Car counts at t−1, t−2, and t−3; top-ranked predictors in feature importance analysis.
Bus	Bus_lag1, Bus_lag2, Bus_lag3	3	Bus counts at t−1, t−2, and t−3 reflect the co-occurrence of public transport with pedestrian flows.
Truck	Truck_lag1, Truck_lag2, Truck_lag3	3	Truck counts at t−1, t−2, and t−3; second-highest importance group after car lags.
Bicycle	Bike_lag1, Bike_lag2, Bike_lag3	3	Bicycle counts at t−1, t−2, and t−3 capture active mobility co-movement with pedestrians.
Meteorological	Average temperature [°C], Precipitation [mm]	2	Hourly air temperature (−6.1 to 29.9°C) and precipitation (0–10.6 mm) from the Finnish Meteorological Institute.
Socio-economic	Average income of households	1	Average annual household income of the device's postal-code area (€29,878–€42,046).
Spatial	Distance from the City Centre	1	Geodesic distance of each monitoring device from the Turku City Centre (0.05–1,529 m).

Decision Tree: non-parametric, rule-based model (maximum depth 6) providing interpretable predictions; serves as a transparent non-linear baseline complementing the parametric linear regression baseline.

XGBoost: gradient-boosted decision tree ensemble (200 estimators, depth 4, learning rate 0.05) with L1/L2 regularization; one of the best-performing models for structured tabular data. XGBoost was selected over random forest, which also ranks among the top-performing models (Fernández-Delgado et al., 2014), because it offers equivalent or superior predictive performance on tabular sensor data while providing built-in L1/L2 regularization, native handling of missing values, and gain-based SHapley Additive exPlanations (SHAP) feature importance (Doshi-Velez & Kim, 2017).

ANN: fully connected feedforward network with three hidden layers (256–128–64 neurons), batch normalization, ReLU activation, and dropout; trained over 150 epochs using the Adam optimiser.

LSTM: A two-layer recurrent network (128 hidden units, sequence length of 6 hours, dropout rate of 0.2) designed to capture temporal dependencies in sequential traffic data; trained on a 30,000-sequence subsample over 100 epochs.

All models were evaluated on a chronological 80/20 train-test split, with performance reported as mean ±SD across 5-fold time-series cross-validation. A LAG_ONLY baseline (XGBoost on t−1 lags only) confirmed that full models learn beyond autocorrelation. Input features were standardized using a StandardScaler fitted on the training set alone.

4 Findings

Under the revised temporal split (chronological 80/20 train–test), XGBoost delivered the strongest location-stratified performance across all four monitoring zones (best: Location 2, $R^2 = 0.839$ ±0.099; Location 1, $R^2 = 0.720$ ±0.071). The ANN performed comparably (Location 2 $R^2 = 0.832$ ±0.103; Location 1 $R^2 = 0.717$ ±0.061). Parametric baselines were weaker (LR: $R^2 = 0.511$–0.514; DT: $R^2 = 0.537$–0.645 at Locations 1–2). The LAG_ONLY autocorrelation baseline (XGBoost using t−1 features only) yielded $R^2 = 0.141$–0.425, confirming substantial contextual gains (Location 1: $\Delta R^2 = +0.295$; Location 2: $\Delta R^2 = +0.698$; Location 3: $\Delta R^2 = +0.389$; Location 4: $\Delta R^2 = +0.290$) beyond simple autocorrelation.

When models were trained separately for each of the four monitoring locations, prediction accuracy varied, revealing strong spatial heterogeneity in predictability. Location 2 (Hämeenkatu / University corridor) yielded the highest accuracy across all models, with XGBoost achieving $R^2 = 0.839$ (±0.099) and RMSE = 118.9. This location exhibits a highly regular noon-hour pedestrian peak (≈ 600 persons/h at 12:00), consistent with lunchtime activity at the University corridor, which ML algorithms can capture effectively. In contrast, Location 3 (Logomo / Ratapihankatu cultural venue) produced the weakest results (XGBoost $R^2 = 0.570$ ±0.105), reflecting the irregular, event-driven footfall patterns typical of arts and entertainment districts. Location 4 (Remote / Outlier, device 16, avg. 2.4

pedestrians/h) yielded the lowest predictive accuracy across all models (XGBoost $R^2 = 0.179 \pm 0.231$, RMSE = 1.46). The near-zero pedestrian counts at this remote building entrance produce a degenerate series with minimal signal for ML models to exploit. Notably, the LAG_ONLY baseline also failed at this location ($R^2 = -0.111$), confirming that Location 4's low predictability is intrinsic to the measurement context rather than a limitation of the modelling approach.

All locations exhibit strong diurnal rhythms, with pedestrian activity concentrated during the daytime and approaching zero in the early hours of the morning. Location 1 (City Centre / Tori) shows the broadest activity window, reflecting the intensity of the commercial core throughout the day. Location 2 shows the sharpest single peak (12:00, ≈ 600 persons/h), consistent with midday university activity. Location 3 displays a broader, flatter profile with a secondary afternoon peak around 14:00 (≈ 208 persons/h). These contrasting temporal signatures confirm that time-of-day encoding, and lag features are essential inputs for modelling.

Table 2 summarises the full performance results. XGBoost feature importance (gain-based) and SHAP beeswarm analysis (Location 1) are provided as supplementary figures. Across Locations 1–3, the dominant predictor was Hour_cos (mean gain = 0.277), followed by Human_lag1 (0.117) and average household income (0.108). Location-specific rankings reveal an important divergence: at Locations 1 and 2, temporal features dominate (Hour_cos gains of 0.288 and 0.295, respectively), whereas at Location 3 (Logomo), household income ranked first (gain = 0.325), substantially above Hour_cos (0.248), reflecting the socio-economic stratification of the arts and entertainment district catchment. Distance from the City Centre was a consistent spatial predictor across all main locations (mean gain = 0.031).

5 Discussion

With regard to RQ1, XGBoost and ANN emerged as the most accurate models for hourly foot-traffic and consistently outperformed parametric baselines. In the location-stratified evaluation using a strict temporal split, XGBoost achieved the strongest predictive performance ($R^2 = 0.720 \pm 0.071$ in the City Centre, 0.839 ± 0.099 in the University corridor, and 0.570 ± 0.105 in the cultural district), consistent with broader evidence that ensemble methods outperform parametric baselines for urban traffic forecasting (Wang et al., 2025). In contrast, the LSTM

model underperformed relative to its theoretical potential, primarily due to training on a limited subsample and early overfitting; a limitation also noted by Dabiri and Heaslip (2018), who highlight that deep sequential models require substantial training data to realize their full capacity. The slight advantage of ANN at the high-volume City Centre location suggests that neural networks scale more effectively with larger and more heterogeneous training sets (Ali, 2025), whereas XGBoost is preferable when data are limited, or interpretability is required (Doshi-Velez & Kim, 2017).

With regard to RQ2, XGBoost feature importance and SHAP analysis reveal a consistent predictor hierarchy. The cosine encoding of hour-of-day (Hour_cos) was the dominant predictor across all main locations (gain = 0.288–0.295), capturing the strong diurnal rhythm of pedestrian flows. Human_lag1 ranked second at Locations 1 and 2 (gain = 0.122–0.201), reflecting short-term autocorrelation. A noteworthy finding is that at Location 3 (Logomo), average household income ranked first (gain = 0.325), above Hour_cos (0.248), reflecting the socio-economic stratification of the arts district catchment. Distance from the City Centre was a consistent spatial predictor (mean gain = 0.031), while meteorological variables exerted secondary effects. The SHAP beeswarm plot (Location 1) confirms that high Hour_cos values drive strongly positive contributions (SHAP $\geq +100$), and that distance from the City Centre shows negative contributions at large distances, consistent with declining pedestrian density away from the urban core. This finding is consistent with those of Koech and Sobhani (2026) and Murcio and Wang (2025) showing that temporal and spatial variables are significant in urban movement prediction.

Prediction performance is driven more by urban land-use type than by data volume. The University corridor achieved the highest accuracy despite having the fewest training records (n = 10,259), because its pedestrian pattern was highly regular and institutionally structured. In contrast, the Logomo cultural venue produced the weakest results ($R^2 = 0.570 \pm 0.105$) due to irregular, event-driven footfall; incorporating event calendar data at such locations is a clear direction for future improvement. Location 4 (remote sensor, avg. 2.4 pedestrians/h) yielded $R^2 = 0.179 \pm 0.231$; the LAG_ONLY baseline also failed here ($R^2 = -0.111$), confirming that low predictability at this site is intrinsic to the measurement context rather than a modelling limitation. This establishes a case for a tiered strategy: location-specific models for atypical sensors, and shared models for high-volume locations.

Table 2: Prediction performance of ML models by location

Locations	Model	R^2 (±SD)	RMSE (±SD)	MAE (±SD)	MSE	MAPE (%)
Location 1: City Centre / Tori: devices 01, 03–12; market square area; avg. distance from City Centre = 69 m; n = 128,193 records.	ANN	0.717 ±0.061	109.34 ±22.03	44.38 ±6.45	11,955	142.9%
	XGB	**0.720 ±0.071**	**108.73 ±23.10**	**43.11 ±3.38**	**11,822**	**149.6%**
	LR	0.514 ±0.017	143.28 ±16.57	71.02 ±3.92	20,529	363.8%
	DT	0.645 ±0.077	122.53 ±23.80	52.91 ±4.78	15,015	155.9%
	LAG_ONLY	0.425 ±0.043	155.82 ±14.54	85.25 ±5.76	24,280	694.6%
Location 2: Hämeenkatu / University corridor: device 02; main commercial street and University corridor; avg. distance = 1,088 m; n = 10,260 records.	ANN	0.832 ±0.103	121.44 ±42.47	71.53 ±22.27	14,748	44.8%
	XGB	**0.839 ±0.099**	**118.94 ±50.44**	**75.06 ±27.48**	**14,147**	**69.9%**
	LR	0.511 ±0.078	207.42 ±51.32	131.54 ±26.99	43,024	143.6%
	DT	0.537 ±0.179	201.86 ±42.35	119.05 ±24.18	40,748	66.9%
	LAG_ONLY	0.141 ±0.215	274.90 ±62.66	175.63 ±29.96	75,570	200.5%
Location 3: Logomo / Ratapihankatu: devices 13–15; cultural venue and railway district; avg. distance = 625 m; n = 25,337 records.	ANN	0.497 ±0.117	82.65 ±12.04	46.08 ±4.06	6,830	129.5%
	XGB	**0.570 ±0.105**	**76.37 ±12.29**	**42.58 ±4.00**	**5,833**	**122.0%**
	LR	0.432 ±0.041	87.83 ±13.84	56.60 ±3.76	7,714	343.2%
	DT	0.515 ±0.083	81.10 ±13.37	44.08 ±3.94	6,578	111.0%
	LAG_ONLY	0.181 ±0.057	105.46 ±13.23	72.63 ±5.88	11,122	720.4%
Location 4: Remote / Outlier: device 16; remote building entrance/exit; avg. distance = 1,529 m; n = 2,263 records.	ANN	−0.196 ±0.417	1.77 ±0.74	1.34 ±0.41	3.1	88.5%
	XGB	**0.179 ±0.231**	**1.46 ±0.67**	**1.10 ±0.30**	**2.1**	**74.5%**
	LR	−0.163 ±0.196	1.74 ±0.66	1.37 ±0.24	3.0	93.6%
	DT	−1.191 ±1.159	2.39 ±1.02	1.30 ±0.38	5.7	86.5%
	LAG_ONLY	−0.111 ±0.232	1.70 ±0.63	1.33 ±0.21	2.9	90.2%

These findings offer location-specific planning insights. High temporal predictability at the City Centre supports operational decisions such as maintenance scheduling and crowd management. The University corridor's structural patterns make it

suitable for events and timetable-driven transport planning. At Logomo, the dominance of socio-economic variables points toward long-term decisions such as zoning and cycling infrastructure investment.

6 Concluding Remarks

The current research demonstrates that ML, in particular XGBoost and ANN, can accurately predict foot-traffic flows when models are tailored to the spatial and functional context of each location, as confirmed under conservative temporal evaluation. Under chronological 80/20 train–test splits with 5-fold time-series cross-validation, XGBoost achieved $R^2 = 0.720 \pm 0.071$ (City Centre) and 0.839 ± 0.099 (University corridor), confirming genuine predictive capability beyond simple autocorrelation (LAG_ONLY baseline: $R^2 = 0.425$ and 0.141, respectively; $\Delta R^2 = +0.295$ to $+0.698$). Performance at the irregular cultural venue (Location 3: $R^2 = 0.570 \pm 0.105$) and the low-volume outlier sensor (Location 4: $R^2 = 0.179 \pm 0.231$) highlights the boundary conditions of the approach. The location-stratified modelling approach, cross-modal feature relationships, and land-use sensitivity presented here provide practical guidance, scoped to the Turku context, for smart city traffic management and urban mobility planning.

Limitations include: seasonal coverage (March–November only; winter unvalidated); LSTM subsampling (full-dataset evaluation needed); single-city scope (validation in other urban contexts required); weather as observed covariate (operational deployment would require forecast values); meteorological variables (sunshine duration and cloud cover are unavailable at hourly resolution from the FMI for the study period and are not included); and evaluation uncertainty ($\pm$SD reflects fold variance, not hyperparameter uncertainty). Future work should extend seasonal coverage, incorporate event calendars and richer spatial features, and pilot the approach in a live operational setting.

Acknowledgement

The authors acknowledge the help of the City of Turku in establishing the monitoring network and thank the relevant municipal departments for providing access to the mobility dataset used in this study. The study is part of the Turku Urban Research Programme and was funded through the MUVIT project (2025-2026).

Endnotes

Data Availability: The dataset was collected in Turku, Finland, through a municipal pedestrian and traffic monitoring network. The data are available upon reasonable request from the corresponding author.

AI Disclosure: The authors used AI-assisted tools for language editing and text drafting in the preparation of this manuscript. All intellectual contributions, analyses, interpretations, and conclusions are solely those of the authors.

Conflict of Interest: The authors declare no conflict of interest.

References

Alhasnawi, B. N., Hashim, H. K., Zanker, M., & Bureš, V. (2025). The rising, applications, challenges, and future prospects of energy in smart grids and smart cities systems. Energy Conversion and Management: X, 101162.

Ali, M. (2025). Classification of imbalanced travel mode choice dataset with SMOTE and prediction using interpretable machine learning. Sustainable Futures, 10, 101119. https://doi.org/10.1016/j.sftr.2025.101119

Asher, M., Oswald, Y., & Malleson, N. (2023). Predicting pedestrian counts using machine learning. AGILE: GIScience Series, 4, 18.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Dabiri, S., & Heaslip, K. (2018). Inferring transportation modes from GPS trajectories using a convolutional neural network. Transportation Research Part C: Emerging Technologies, 86, 360–371.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv Preprint arXiv:1702.08608.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? The Journal of Machine Learning Research, 15(1), 3133–3181.

Gil, I. A., Ferreira, F. A., Ferreira, N. C., Smarandache, F., Khanam, M., & Daim, T. (2026). Artificial intelligence, society 5.0 and smart city adaptation initiatives for businesses: An integrated approach. Technovation, 150, 103377.

Haghshenas, S. S., Guido, G., Haghshenas, S. S., Astarita, V., & Simic, V. (2026). Artificial intelligence techniques in road safety for sustainable mobility: An investigation into applications, challenges, and opportunities. Reliable Decision-Making for Sustainable Transportation, 203–223.

Jamal, S., Siddiqui, F., Alam, M. A., Ayman-Mursaleen, M., Zafar, S., & Naaz, S. (2026). Deep Reinforcement Learning for Sustainable Urban Mobility: A Bibliometric and Empirical Review. Sensors (Basel, Switzerland), 26(2), 376.

Kim, S.-K., & Chan, I. C. (2025). Novel Machine Learning-Based Smart City Pedestrian Road Crossing Alerts. Smart Cities, 8(4), 114.

Koech, M. J., & Sobhani, N. (2026). Predicting Future Pedestrian Movement Patterns: An LSTM Approach Using Historical Footfall Data from Dublin. Journal of Mathematics and Data Science (JMDS), 4(1), 15–22.

Lingani, G. M., Rawat, D. B., & Garuba, M. (2019). Smart traffic management system using deep learning for smart city applications. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), 0101–0106.

Murcio, R., & Wang, Y. (2025). Identify optimal pedestrian flow forecasting methods in Great Britain retail areas: A comparative study of time series forecasting on a footfall dataset. ISPRS International Journal of Geo-Information, 14(2), 50.

Ng, Y., Pei, Y., & Jin, Y. (2017). Footfall count estimation techniques using mobile data. 2017 18th IEEE International Conference on Mobile Data Management (MDM), 307–314.

Peráček, T., & Kaššaj, M. (2026). Strategic Management of Urban Services Using Artificial Intelligence in the Development of Sustainable Smart Cities—Managerial and Legal Challenges. Sustainability, 18(2), 582.

Pulugurta, S., Arun, A., & Errampalli, M. (2013). Use of artificial intelligence for mode choice analysis and comparison with the traditional multinomial logit model. Procedia-Social and Behavioral Sciences, 104, 583–592.

Schläpfer, M., Dong, L., O'Keeffe, K., Santi, P., Szell, M., Salat, H., Anklesaria, S., Vazifeh, M., Ratti, C., & West, G. B. (2021). The universal visitation law of human mobility. Nature, 593(7860), 522–527.

Wang, C., Huang, S., & Zhang, C. (2025). Short-term traffic flow prediction considering weather factors based on optimized deep learning neural networks: Bo-GRA-CNN-BiLSTM. Sustainability, 17(6), 2576.

Wolniak, R., Stecula, K., & Aydın, B. (2024). Digital transformation of grocery in-store shopping-scanners, artificial intelligence, augmented reality and beyond: A review. Foods, 13(18), 2948.

Wu, C., Ding, H., Fu, Z., & Sun, N. (2024). Multi-Scale Spatio-Temporal Attention Networks for Network-Scale Traffic Learning and Forecasting. Sensors, 24(17), 5543. https://doi.org/10.3390/s24175543

About the authors

Katayoon Pourmahdi is a PhD researcher at the School of Business and Economics at ÅAU. Her work focuses on integrating analytical approaches into decision-making, particularly in pricing, to bridge theory and practice.

Dr **Bahareh Naseri** is a postdoctoral researcher at the School of Business and Economics at ÅAU. Her research focuses on cost–benefit analysis within value chains.

Dr **Ilia Gugenishvili** is an Assistant Professor of Marketing at the School of Business and Economics at ÅAU. His research interests include sustainability, consumer behaviour, and technology.

Dr **Jozsef Mezei** is a Professor of Information Systems at the School of Business and Economics at ÅAU. His expertise includes ML in business contexts, decision-making under uncertainty, decision-support systems, analytics, soft computing, and fuzzy logic.

Dr **Anna-Greta Nyström** is a Professor of International Business at the School of Business and Economics at ÅAU. She specializes in innovation management, business models, and market shaping in ICT and emerging technologies.

It has been proofread and complies with the language standards used in the publication.

Summary

Smart cities rely on intelligent tools to manage mobility and reduce congestion, yet context-aware prediction of pedestrian and traffic flows remains limited. This study applies five ML models to a large urban dataset from Turku, Finland, incorporating temporal, spatial, meteorological, and socio-economic factors. The results show that XGBoost and neural networks perform best, but model accuracy varies across locations. These findings highlight the need for context-aware models and offer practical insights for urban planning and real-time traffic management.

Towards Agentic Personal Health and Organizational Domain Coordination in Occupational Nutrition

Kevin Nils Röhl,[1] Rainer Alt,[2] Jan Wirsam[1]
[1] HTW Berlin, University of Applied Science, Berlin, Germany
roehl@htw-berlin.de, wirsam@htw-berlin.de
[2] Leipzig University, Leipzig, Germany
rainer.alt@uni-leipzig.de

Workplace meal choices shape dietary behavior and health outcomes, while individuals increasingly generate health-related data through wearables. Although both domains relate to the same meal event, they remain digitally separate because they operate under different governance conditions and data sensitivities. This creates a coordination problem: individuals lack structured access to organizational meal information, while meal-service providers should not access raw personal health data. This paper develops a coordination model for cross-domain interaction between the personal and organizational domains. Drawing on coordination theory, the model specifies how locally computed personal indicators and organizational meal-related information can interact within a governing coordination infrastructure centered on the meal event as the unit of coordination. A prototype instantiation illustrates the design through bounded coordination artifacts and tool-mediated interaction. An initial controlled benchmark indicates the technical feasibility of the approach, while questions of real-world adoption and organizational impact remain for future research.

DOI
https://doi.org/10.18690/um.fov.4.2026.39

ISBN
978-961-299-152-4

Keywords:
digital health, coordination, agentic AI, wearables, cross-domain

University of Maribor Press

1 Introduction

Cardiovascular disease, diabetes and related metabolic risk make everyday food choices increasingly relevant for prevention and self-management. In occupational settings, canteens shape these choices because they determine which meals are offered and how dishes are composed (Friedman & Mozaffarian, 2016). Reviews of workplace cafeteria interventions show that changes in the food environment can influence dietary behavior and health-related outcomes (Naicker et al., 2021). Nutrition research further shows that meal composition and preparation matter for postprandial glycemic response. Carbohydrate amount, fiber, protein, fat and cooking can all change how a meal affects glucose after eating (Murillo et al., 2022; Pasmans et al., 2022). At the same time, nutrition-related data are generated on the personal side. Individuals use meal logs and wearable devices to observe related health outcomes (George et al., 2023). This also includes continuous glucose monitoring, especially in the context of prediabetes, gestational diabetes and type 2 diabetes. Research on personalized nutrition has shown that people can react very differently to the same meal and that glucose meal responses can support individualized dietary guidance (Zeevi et al., 2015). Although both domains are relevant to the same meal event, they rarely interact digitally. Privacy regulation, data-management challenges and the sensitivity of health-related information constrain direct sharing of personal insights in organizational settings (Donovan et al., 2025). Individuals usually do not have digital access to recipe, portion or preparation data from the canteen, while food-service actors cannot and should not interpret raw personal health data (Pateraki et al., 2020; Salehzadeh Niksirat et al., 2024; Virgillito et al., 2025). This creates a coordination problem between two domains that each hold only part of the relevant knowledge (Caccamo et al., 2023). In this context, coordination theory provides a useful lens because the challenge is the management of dependencies among activities distributed across multiple actors (Malone & Crowston, 1994). In multi-agent settings, these dependencies are shaped also by infrastructures that mediate and govern the interaction space (Omicini et al., 2004; Rodriguez-Aguilar et al., 2015).

Developments in edge AI, as well as LLM-based tool selection, make it feasible to process sensitive indicators locally and connect them to organizational services through bounded interfaces (Krishnan, 2025; Piccialli et al., 2025; Singh & Gill, 2023; Zhao et al., 2026). Prior coordination-infrastructure research has advanced the

structuring of interaction, but it still identifies agent decision support and decentralized coordination design as open challenges (Rodriguez-Aguilar et al., 2015). Related governance research likewise calls for further study of data governance across multi-actor settings, particularly where personal data are involved (Paparova et al., 2023). In occupational nutrition, this gap is particularly salient because personal health-related indicators and organizational meal-service data are jointly relevant to the same meal event but remain separated by different governance regimes and data sensitivities (Donovan et al., 2025). What is missing is a governed coordination model for interaction between personal and organizational domains around shared meal-related and health-related indicators. This paper addresses that gap by developing the Agentic Cross-Domain Coordination Model (ACCM). Its technical feasibility is examined through a prototype instantiation of selected components and evaluated with bounded LLM-based tool calls via a natural-language interface.

RQ: How can a model be conceptualized for cross-domain coordination between the personal health domain and the organizational domain in occupational nutrition settings?

2 Literature Review

2.1 Coordination Infrastructure

"How will the widespread use of information technology change the ways people work together?" (Malone & Crowston, 1994). This question motivates Malone and Crowston's research. They argue that the most significant effect of information technology is its ability to make coordination between individuals and groups both cheaper and more effective. As coordination costs decline, organizations can move beyond traditional hierarchies and adopt more flexible structures. In this way, information technology reshapes the fundamental ways work is organized within and across organizations. This study adopts Malone and Crowston's definition of coordination as "managing dependencies between activities" (Malone & Crowston, 1994). In Computer-Supported Cooperative Work (CSCW), this concept is extended through the incorporation of models of coordinative practices into computational artifacts that support actors in dealing with contingencies in cooperative work (Ciolfi et al., 2023). In the multi-agent system (MAS) literature, this becomes a matter of

governing the interaction space through coordination infrastructures. In MAS, the environment is conceived as an explicit part of the system, with the dual role of providing the surrounding conditions for agents to exist and an exploitable design abstraction for building applications (Weyns et al., 2006). Artifact-oriented approaches in MAS distinguish between agents and coordination artifacts or media. Agents exploit coordination artifacts at the coordination stage, while coordination media first enable interactions among agents and then mediate them in a more automated manner. In this sense, coordination artifacts externalize coordination knowledge into the interaction environment or infrastructure. This corresponds to the distinction between subjective and objective coordination: subjective coordination concerns the reasoning processes of agents, whereas objective coordination is enacted through coordination media that embed and enforce social laws and interaction constraints (Omicini et al., 2004). Agent Coordination Contexts (ACCs) refine this further by defining the controlled layer through which an agent can interact with its environment, making coordination and security constraints explicit at the agent-environment boundary (Ricci et al., 2006). This study therefore uses coordination theory as a lens to examine how meal-related dependencies are mediated across domains.

2.2 Large Language Models and Agentic AI

In cross-domain settings, another challenge is that users and services operate with different interpretive contexts. NLP-based interfaces can reduce knowledge barriers between users and information systems by enabling access through natural language (Sawant & Sonawane, 2024). LLMs extend this by offering more flexible natural-language understanding and by supporting semantic interoperability tasks such as schema alignment and knowledge integration across heterogeneous domains (Abu-Salih et al., 2025; Zhou et al., 2025). However, these capabilities do not remove governance concerns. Instead, they increase the need for secure and governed interaction (Abu-Salih et al., 2025). Agentic AI adds a capability beyond language understanding. AI agents can be categorized by their autonomy, capabilities and underlying technologies (Krishnan, 2025). Traditional agents include rule-based systems and agents using machine learning or reinforcement learning for adaptive decision-making (Xi et al., 2025). More advanced forms leverage LLMs, enabling complex reasoning, tool integration and diverse tasks such as scheduling and data summarization (Sapkota et al., 2026). Recent developments introduce Agentic AI,

referring to highly autonomous agents capable of dynamic task decomposition and multi-agent coordination, suited for complex domains such as medical decision support (Murugesan, 2025). Emerging standards such as the Model Context Protocol (MCP) enable advanced agents to interact and access external tools in multi-agent systems (Li & Xie, 2025). Their integration also raises challenges, including semantic interoperability and the need for robust governance (Huang & Hughes, 2025). Overall, MCP forms foundational infrastructure for collaborative multi-agent applications (Liao et al., 2025).

2.3 Interoperability & Data Spaces

In Europe, the International Data Spaces (IDS) and Gaia-X promote federated data ecosystems where organizations exchange data under common standards and sovereignty assurances (von Scherenberg et al., 2024). These frameworks emphasize legal agreements, identity management and trusted execution environments but focus mainly on B2B interoperability rather than individuals as distinct stakeholders (Scheider et al., 2023). In parallel, personal information management systems (PIMS) have emerged, exemplified by Solid pods and the MyData movement, which give individuals control over storage and selective sharing (Janssen & Singh, 2022). Despite progress in inter-organizational data spaces and PIMS, most frameworks privilege one over the other. This gap calls for a cross-domain interaction that bridges user-centric control with organization-level interoperability (Delacroix & Lawrence, 2019).

2.4 Synthesis

In the present problem space, a single meal event spans two domains that remain institutionally and technically separated. The core challenge is therefore one of coordination across distributed actors, data sources and governance regimes. Coordination theory defines this challenge as the management of dependencies between activities, while CSCW and MAS show that such dependencies can be managed through computational artifacts and governed interaction infrastructures. However, these perspectives do not address the semantic and knowledge barriers that arise when users, services and data sources operate with different vocabularies and structures across domains. Here, NLP-based interfaces and LLMs add an important capability by enabling semantic mediation and the translation of user

intent into structured requests across heterogeneous systems. Agentic AI adds the possibility of actors that can interpret intent and act across services, but data space and governance research make clear that such action must remain bounded by sovereignty and policy requirements. Taken together, these perspectives suggest that the design problem is to create a governed coordination model that links personal and organizational domains through controlled layers. The synthesis informs the formulation of four design requirements concerning local resource control, governed cross-domain interaction, bounded exchange and boundary-specific admissibility. Table 1 summarizes these requirements and their realization in ACCM.

Table 1: Literature-derived requirements and their realization in ACCM

Requirement	Literature	ACCM
R1 Local control of domain resources	PIMS / MyData; data governance	Personal and organizational resources remain local; domain-specific tools compute outputs locally
R2 Governed cross-domain interaction	Coordination-infrastructure theory	A Governing Coordination Infrastructure organizes objective coordination and exchange conditions
R3 Bounded and selective exchange	Coordination-infrastructure theory; governance literature	Cross-domain interaction occurs through bounded coordination artifacts
R4 Boundary-specific admissibility control	ACC theory	Agent Coordination Contexts define admissible actions and perceptions at the domain boundary

3 Methodology

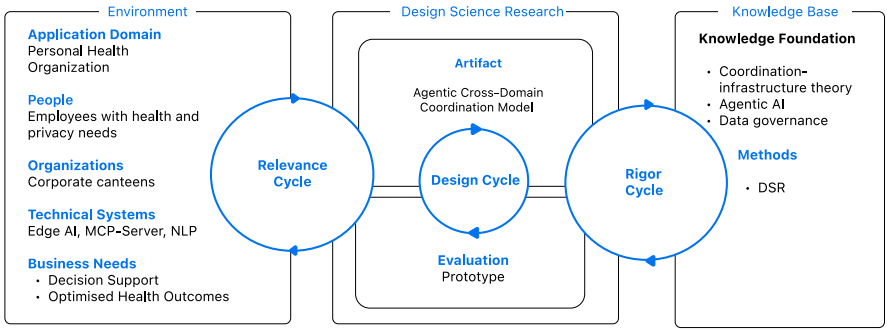

Figure 1: Three Cycle View adapted from
Source: (Hevner, 2007)

This study adopts design science research (DSR) to develop and evaluate an artifact for governed cross-domain interaction in occupational nutrition settings (Hevner, 2007).

Table 2: Illustrative indicators, organizational uses and tools

Domain	Measure	Use & Literature	Related Tools
Organization	Carb & portion size	Portion and starch adjustment (Vlachos et al., 2020)	find_menu_items (T5) get_dish_carbohydrates (T4) find_lowest_carb_dish (T2)
	Fiber, Fat, Protein	Increase fiber-rich components, higher-protein meals, and adjust recipe composition (Kim et al., 2019)	find_menu_items (T6)
	Preparation	Change processing, cooking, or food order (Cesbron-Lavau et al., 2021)	get_dish_preparation_info (T7)
Personal	Pre-meal baseline	Interpret Glucose starting state (Ehlert et al., 2026)	get_premeal_baseline_glucose (T8)
	Rise to peak	Glucose excursion magnitude (Ehlert et al., 2026)	find_highest_postmeal_peak (T1)
	2-h iAUC	Glucose total exposure (Ehlert et al., 2026)	get_postmeal_indicator_bundle
	Time to peak	Glucose response speed (Ehlert et al., 2026)	get_postmeal_indicator_bundle (T3)

Following Hevner's three-cycle view, the study combines relevance, rigor and design activities as show in Figure 1. The relevance cycle is grounded in the challenge of linking personal nutrition-related data with organizational meal-service data while preserving privacy. The rigor cycle draws on the theoretical synthesis from which the design requirements for ACCM were derived and is shown in Table 1. In the design cycle an initial ACCM model was derived from the requirements, implemented through a prototype instantiation of selected components and refined based on the prototype findings. The prototype was developed with real-world data collected during the study intervention by meala (meala GmbH, 2025), a food-tracking diary app that integrates glucose data from FreeStyle Libre and Dexcom and connects to Apple Health and Health Connect for contextual data such as activity and sleep. It contained more than 30,000 glucose values and 955 tracked meals from a company canteen and private settings. The study further used the menu data from a company canteen in Berlin, Germany, including 687 menus with nutritional information and portion sizes collected between August and November 2024. In total, 10 users aged 25 to 55 participated. Seven pursued healthier eating,

two aimed at performance improvement and one had no specific goal. To instantiate ACCM, a prototype was implemented with FastMCP (*Fastmcp*, 2026). It operationalized selected elements of the coordination model and exposed a restricted set of indicator tools derived from the Glucose360 platform (Ehlert et al., 2026) and glucose-related nutrition studies listed in Table 2.

Table 3: Benchmark queries and expected tools

ID	User query	Tool	Arguments
Q1	What was my highest post-meal peak between 2024-08-12 and 2024-08-18?	T1	date_range_start=2024-08-12, date_range_end=2024-08-18
Q2	What is the dish with the lowest carbs in all available menu data?	T2	none
Q3	For meal event ME1, how did I react?	T3	meal_event_id=ME1
Q4	For meal event ME1, what was the time to peak in the first 120 minutes?	T3	meal_event_id=ME1
Q5	For meal event ME1, what is the 2-hour incremental AUC (iAUC 0–120)?	T3	meal_event_id=ME1
Q6	How did I react to meal event ME1?	T3	meal_event_id=ME1
Q7	What is the rise to peak for meal id ME1?	T3	meal_event_id=ME1

ME1 – meal id from food diary

The prototype provided a technical instantiation of ACCM to examine the feasibility of governed LLM-based tool selection and local calculated indicators. Evaluation was conducted as a feasibility-oriented ex post assessment using a benchmark of prompt-expectation test cases. For each test case, labels were defined manually before execution by specifying the expected tool call and the expected arguments as shown in Table 3. The benchmark queries were executed three times per mode using the same tool descriptions and instruction setup across models. The evaluated models were qwen3.5:0.8b, granite4:latest, llama3.1:8b and gpt-4o. They were selected to compare smaller models suitable for edge AI settings with a closed cloud-based model. The local open models were run via Ollama on an Apple M1 Pro and 32 GB RAM, while GPT-4o was accessed via API. The evaluation examined tool-selection accuracy, parameter correctness, false positives, false negatives, overcall behavior and latency. Tool-selection accuracy measured whether the labeled tool was selected. False positives captured incorrect or unnecessary tool calls, whereas false negatives captured missed required tool calls. Overcall behavior referred specifically to cases in which a model invoked more tool use than required. Latency was recorded descriptively as end-to-end response time and is interpreted cautiously

because execution environments differed across models. This evaluation was used to assess the technical plausibility of the prototype and to derive implications for the design of the governed layer.

4 Model

Figure 2 specifies ACCM as a federated coordination model for workplace meal events. The model draws on coordination theory by treating cross-domain interaction as the management of dependencies between two operationally separate domains: a personal health domain and an organizational meal-service domain. In line with coordination-infrastructure theory, ACCM separates subjective coordination from objective coordination, which is organized through a shared governing infrastructure. Each domain contains local resources, local tools, local agents and user-facing applications or services. In the personal domain, resources consist of employee-held health-related data, while tools derive bounded indicators from these data for local use. In the organizational domain, resources consist of meal-service and nutrition-related data, while tools process menu, recipe and ingredient information. In both domains, agents coordinate locally over these resources and tools. Raw domain data are therefore not exposed to the other side as part of the coordination arrangement. Cross-domain interaction is mediated through Agent Coordination Contexts (ACCs) positioned at the boundary between each domain and the shared infrastructure. The ACC defines the admissible actions and perceptions available to a given agent at the boundary between its domain and the governing coordination infrastructure. It preserves the distinction between local agent autonomy and infrastructural governance: the agent remains responsible for local interpretation and action selection, while the infrastructure regulates the terms under which cross-domain interaction becomes possible. The central coordination device in ACCM is the meal event. The model represents it through bounded coordination artifacts within the governing infrastructure. These artifacts link organizational descriptions of meals with locally computed personal indicators related to the same event. In this way, the meal event becomes the shared coordination object without requiring either domain to relinquish control over its underlying resources. Cross-domain exchange is limited to bounded, decision-relevant outputs produced locally and made available under infrastructural conditions. The participating domains remain distinct, maintain their own resources and interact through a governed coordination infrastructure. ACCM therefore

specifies a form of selective interoperability in which cross-domain coordination is achieved through bounded coordination artifacts, governed exchange conditions, and boundary admissibility controls.

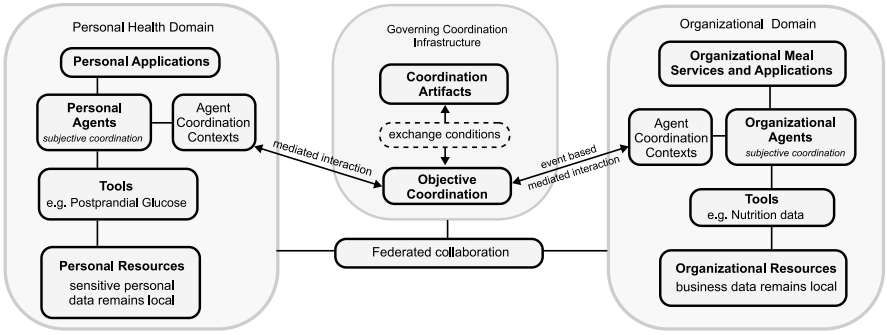

Figure 2: Conceptual Agentic Cross-Domain Coordination Model
Source: Own

5 Results

The prototype was used to generate 21 evaluations per model (7 cases, each repeated three times), gpt-4o achieved 100.0% true positives and the lowest mean latency (1,970 ms). Among the open models, granite4:latest obtained the highest performance (61.9% TP), followed by llama3.1:8b (52.4% TP) and qwen3.5:0.8b (38.1% TP). The complete results are shown in Table 4. At the case level, the cases listed in Table 3 showed different levels of performance across the open models. Q1 ("What was my highest post-meal peak between 2024-08-12 and 2024-08-18?") and Q5 ("For meal event ME1, what is the 2-hour incremental AUC?") each recorded 7/9 correct runs, while Q2 ("What is the dish with the lowest carbs in all available menu data?") recorded 6/9 correct runs. The date-bounded case Q1 achieved a 100.0% date-pass rate across all models. Q4 ("For meal event ME1, what was the time to peak in the first 120 minutes?") recorded 5/9 correct runs. Performance was lower for the semantically similar reaction queries in Q3 ("For meal event ME1, how did I react?"), recorded 3/9 and Q6 ("How did I react to meal event ME1?"), recorded 2/9 correct runs. The same results for Q7 ("What is the rise to peak for meal id ME1?").

Table 2: Tool-invocation benchmark across models

Model	TP (%)	FP	FN	Overcall (%)	Avg. latency (ms)
llama3.1:8b	52.4	0	10	0.0	9,288
granite4:latest	61.9	3	5	14.3	5,608
qwen3.5:0.8b	38.1	11	1	52.4	63,866
gpt-4o	100.0	0	0	0.0	1,970

TP = true positive rate; FP = false positives; FN = false negatives; Overcall = unnecessary tool invocation rate

6 Discussion

This study frames workplace nutrition as a cross-domain coordination challenge in which the relevant dependency is distributed across two operationally separate domains. The design-knowledge contribution of the study lies in an event-centered coordination pattern for selective interoperability in sensitive settings. In ACCM, the meal event functions as a bounded coordination artifact that links organizational meal descriptors with locally computed personal indicators without requiring either domain to expose its underlying resources. The contribution lies in their integration through a governed coordination model in which a shared event structures the scope of coordination, the conditions of exchange, and admissibility at the domain boundary. In this sense, the meal event serves both as a semantic coordination unit and as a practical privacy boundary for cross-domain interaction.

The prototype and benchmark provide preliminary support for this pattern. They show that natural-language requests can be translated into bounded tool calls and cross-domain responses, thereby reducing semantic and knowledge barriers for users. At the same time, the findings indicate that privacy-preserving coordination remains reliable only when AI operates within explicit governance boundaries and restricted action spaces. In the evaluated benchmark, the cloud-based model achieved the strongest tool-selection performance, whereas the tested local models were less reliable, especially for semantically broader formulations. The implication is the continued importance of a governed coordination layer based on local computation, bounded tool mediation and precise event-bounded exchange. Edge AI appears promising for such settings, but further progress in local model quality and edge-specific fine-tuning is needed before it can support this coordination pattern with comparable robustness. The evaluation was limited to a prototype and a controlled benchmark assessing technical feasibility. It did not validate downstream effects. On the personal side, the study did not examine whether the

proposed coordination pattern improves health-related outcomes. On the organizational side, it did not evaluate whether the generated indicators improve menu planning or other forms of decision support. The prototype should therefore be understood as a preparatory step toward a future intervention study in which the coordination model can be instantiated in real organizational settings to assess adoption, decision processes and measurable effects on nutrition-related health outcomes

7 Conclusion

This paper presented ACCM as a coordination model for governed cross-domain interaction between personal health platforms and organizational meal-service platforms. The model specifies how locally computed personal indicators and organizational meal-related information can be connected through a governed layer while preserving the local control of sensitive personal data. Beyond assessing technical feasibility, the initial prototype and controlled benchmark also informed the iterative refinement of ACCM by providing preliminary design knowledge about the governed layer. The findings suggest that coordination artifacts can support this interaction pattern and that the meal event serves as a useful coordination and privacy boundary for cross-domain interaction. The paper therefore contributes a conceptual design-oriented model for selective cross-domain interoperability in occupational nutrition. Future research should examine the model in real organizational settings and assess adoption and practical outcome effects.

References

Abu-Salih, B., Alotaibi, S., Alanazi, A. L., Abu Khurma, R., Al-Shboul, B., Khouri, A., & Aljaafari, M. (2025). Using large language models for semantic interoperability: A systematic literature review. *ICT Express*, *11*(4), 819–837. https://doi.org/10.1016/j.icte.2025.06.011

Caccamo, M., Pittino, D., & Tell, F. (2023). Boundary objects, knowledge integration, and innovation management: A systematic review of the literature. *Technovation*, *122*, 102645. https://doi.org/10.1016/j.technovation.2022.102645

Cesbron-Lavau, G., Goux, A., Atkinson, F., Meynier, A., & Vinoy, S. (2021). Deep Dive Into the Effects of Food Processing on Limiting Starch Digestibility and Lowering the Glycemic Response. *Nutrients*, *13*(2), 381. https://doi.org/10.3390/nu13020381

Ciolfi, L., Lewkowicz, M., & Schmidt, K. (2023). Computer-Supported Cooperative Work. *Handbook of Human Computer Interaction*, 1–26. https://doi.org/10.1007/978-3-319-27648-9_30-1

Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law*, *9*(4), 236–252. https://doi.org/10.1093/IDPL/IPZ014

Donovan, S. M., Abrahams, M., Anthony, J. C., Bao, Y., Barragan, M., Bermingham, K. M., Blander, G., Keck, A. S., Lee, B. Y., Nieman, K. M., Ordovas, J. M., Penev, V., Reinders, M. J., Sollid, K., Thosar, S., & Winters, B. L. (2025). Personalized nutrition: perspectives on challenges, opportunities, and guiding principles for data use and fusion. *Critical Reviews in Food Science and Nutrition*, *65*(30), 7151–7169. https://doi.org/10.1080/10408398.2025.2461237

Ehlert, B., Aron, D., Perelman, D., Wu, Y., & Snyder, M. P. (2026). Glucose360: An Open-Source Python Platform with Event-Based Integration for Continuous Glucose Monitoring Data Analysis. *Diabetes Technology & Therapeutics*, *28*(2), 130. https://doi.org/10.1177/15209156251374711

fastmcp. (2026). https://github.com/PrefectHQ/fastmcp

Friedman, F., & Mozaffarian, D. (2016). Dietary and Policy Priorities for Cardiovascular Disease, Diabetes, and Obesity. *Circulation*, *133*(2), 187–225. https://doi.org/10.1161/CIRCULATIONAHA.115.018585

George, A. S. H., Shahul, A., & George, Dr. A. S. (2023). Wearable Sensors: A New Way to Track Health and Wellness. *Partners Universal International Innovation Journal (PUIIJ)*, *01*(04), 15–34. https://doi.org/10.5281/ZENODO.8260879

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, *19*(2). https://aisel.aisnet.org/sjis/vol19/iss2/4

Huang, K., & Hughes, C. (2025). *Agentic AI Communication Protocols and Security*. 81–110. https://doi.org/10.1007/978-3-032-02130-4_4

Janssen, H., & Singh, J. (2022). Personal Information Management Systems. *Internet Policy Review*, *11*(2). https://doi.org/10.14763/2022.2.1659

Kim, J. S., Nam, K., & Chung, S. J. (2019). Effect of nutrient composition in a mixed meal on the postprandial glycemic response in healthy people: a preliminary study. *Nutrition Research and Practice*, *13*(2), 126. https://doi.org/10.4162/nrp.2019.13.2.126

Krishnan, N. (2025). AI Agents: Evolution, Architecture, and Real-World Applications. *ArXiv*. https://doi.org/10.48550/arXiv.2503.12687

Li, Q., & Xie, Y. (2025). From Glue-Code to Protocols: A Critical Analysis of A2A and MCP Integration for Scalable Agent Systems. *ArXiv*, (abs/2505.03864). https://doi.org/10.48550/arXiv.2505.03864

Liao, C. C., Liao, D., & Surya Gadiraju, S. (2025). AgentMaster: A Multi-Agent Conversational Framework Using A2A and MCP Protocols for Multimodal Information Retrieval and Analysis. *ArXiv*, *abs/2507.21105*. https://doi.org/10.48550/arXiv.2507.21105

Malone, T. W., & Crowston, K. (1994). The Interdisciplinary Study of Coordination. *ACM Computing Surveys (CSUR)*, *26*(1), 87–119. https://doi.org/10.1145/174666.174668

meala GmbH. (2025). *meala*. https://www.heymeala.com/en/

Murillo, S., Mallol, A., Adot, A., Juárez, F., Coll, A., Gastaldo, I., & Roura, E. (2022). Culinary strategies to manage glycemic response in people with type 2 diabetes: A narrative review. *Frontiers in Nutrition*, *9*, 1025993. https://doi.org/10.3389/fnut.2022.1025993

Murugesan, S. (2025). The Rise of Agentic AI: Implications, Concerns, and the Path Forward. *IEEE Intelligent Systems*, *40*(2), 8–14. https://doi.org/10.1109/MIS.2025.3544940

Naicker, A., Shrestha, A., Joshi, C., Willett, W., & Spiegelman, D. (2021). Workplace cafeteria and other multicomponent interventions to promote healthy eating among adults: A systematic review. *Preventive Medicine Reports*, *22*, 101333. https://doi.org/10.1016/j.pmedr.2021.101333

Omicini, A., Ossowski, S., & Ricci, A. (2004). Coordination Infrastructures in the Engineering of Multiagent Systems. *Methodologies and Software Engineering for Agent Systems*, 273–296. https://doi.org/10.1007/1-4020-8058-1_17

Paparova, D., Aanestad, M., Vassilakopoulou, P., & Bahus, M. K. (2023). Data governance spaces: The case of a national digital service for personal health data. *Information and Organization*, *33*(1), 100451. https://doi.org/10.1016/j.infoandorg.2023.100451

Pasmans, K., Meex, R. C. R., van Loon, L. J. C., & Blaak, E. E. (2022). Nutritional strategies to attenuate postprandial glycemic response. *Obesity Reviews*, *23*(9), e13486. https://doi.org/10.1111/obr.13486

Pateraki, M., Fysarakis, K., Sakkalis, V., Spanoudakis, G., Varlamis, I., Maniadakis, M., Lourakis, M., Ioannidis, S., Cummins, N., Schuller, B., Loutsetis, E., & Koutsouris, D. (2020). Biosensors and Internet of Things in smart healthcare applications: challenges and opportunities. *Wearable and Implantable Medical Devices: Applications and Challenges*, 25–53. https://doi.org/10.1016/B978-0-12-815369-7.00002-1

Piccialli, F., Chiaro, D., Sarwar, S., Cerciello, D., Qi, P., & Mele, V. (2025). AgentAI: A comprehensive survey on autonomous agents in distributed AI for industry 4.0. *Expert Systems with Applications*, *291*, 128404. https://doi.org/10.1016/J.ESWA.2025.128404

Ricci, A., Viroli, M., & Omicini, A. (2006). Agent coordination contexts in a MAS coordination infrastructure. *Applied Artificial Intelligence*, *20*(2–4), 179–202. https://doi.org/10.1080/08839510500484207

Rodriguez-Aguilar, J. A., Sierra, C., Arcos, J. L., Lopez-Sanchez, M., & Rodriguez, I. (2015). Towards next generation coordination infrastructures. *The Knowledge Engineering Review*, *30*(4), 435–453. https://doi.org/10.1017/S0269888915000090

Salehzadeh Niksirat, K., Velykoivanenko, L., Zufferey, N., Cherubini, M., Huguenin, K., Humbert, M., Velykoivanenko, L., Zufferey, N., Cherubini, M., Huguenin, K., & Humbert, M. (2024). Wearable Activity Trackers: A Survey on Utility, Privacy, and Security. *ACM Computing Surveys*, *56*(7), 183. https://doi.org/10.1145/3645091

Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2026). AI Agents vs. Agentic AI: A Conceptual taxonomy, applications and challenges. *Information Fusion*, *126*, 103599. https://doi.org/10.1016/J.INFFUS.2025.103599

Sawant, P., & Sonawane, K. (2024). NLP-based smart decision making for business and academics. *Natural Language Processing Journal*, *8*, 100090. https://doi.org/10.1016/j.nlp.2024.100090

Scheider, S., Lauf, F., Möller, F., & Otto, B. (2023). A Reference System Architecture with Data Sovereignty for Human-Centric Data Ecosystems. *Business and Information Systems Engineering*, *65*(5), 577–595. https://doi.org/10.1007/s12599-023-00816-9

Singh, R., & Gill, S. S. (2023). Edge AI: A survey. *Internet of Things and Cyber-Physical Systems*, *3*, 71–92. https://doi.org/10.1016/J.IOTCPS.2023.02.004

Virgillito, D., Catalfo, P., & Ledda, C. (2025). Wearables in Healthcare Organizations: Implications for Occupational Health, Organizational Performance, and Economic Outcomes. *Healthcare*, *13*(18), 2289. https://doi.org/10.3390/HEALTHCARE13182289

Vlachos, D., Malisova, S., Lindberg, F. A., & Karaniki, G. (2020). Glycemic Index (GI) or Glycemic Load (GL) and Dietary Interventions for Optimizing Postprandial Hyperglycemia in Patients with T2 Diabetes: A Review. *Nutrients*, *12*(6), 1561. https://doi.org/10.3390/nu12061561

von Scherenberg, F., Hellmeier, M., & Otto, B. (2024). Data Sovereignty in Information Systems. *Electronic Markets 2024 34:1*, *34*(1), 1–11. https://doi.org/10.1007/S12525-024-00693-4

Weyns, D., Omicini, A., Odell, J., Weyns, D., Omicini, A., & Odell, J. (2006). Environment as a first class abstraction in multiagent systems. *Autonomous Agents and Multi-Agent Systems 2006 14:1*, *14*(1), 5–30. https://doi.org/10.1007/s10458-006-0012-0

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang, W., Jiang, C., Zou, Y., Liu, X., … Gui, T. (2025). The rise and potential of large language model based agents: a survey. *Science China Information Sciences*, *68*(2), 1–44. https://doi.org/10.1007/S11432-024-4222-0

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J. A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalová, L., Pevsner-Fischer, M., … Segal, E. (2015). Personalized Nutrition by Prediction of Glycemic Responses. *Cell*, *163*(5), 1079–1094. https://doi.org/10.1016/j.cell.2015.11.001

Zhao, L., Liu, S., Xin, T., Tan, J., Wang, X., Li, Y., Bian, Z., Chen, Y., Kong, F., Bian, J., Qian, C., & Zhang, Z. (2026). AI agent in healthcare: applications, evaluations, and future directions. *Npj Artificial Intelligence 2026 2:1*, *2*(1), 31-. https://doi.org/10.1038/s44387-026-00076-4

Zhou, J., Li, H., Chen, S., Chen, Z., Han, Z., & Gao, X. (2025). Large language models in biomedicine and healthcare. *Npj Artificial Intelligence 2025 1:1*, *1*(1), 44-. https://doi.org/10.1038/s44387-025-00047-1

SME Production Resilience: A Four-Layer AI Use Case Selection Approach

Klaus Müller, Michael Schöppe, Horst Wildemann
Technical University Munich, Research Institute for Corporate Governance, Logistics and Production, Munich, Germany
klaus1.mueller@tum.de, michael.schoeppe@tum.de, sek.bwl@mgt.tum.de

DOI
https://doi.org/10.18690/um.fov.4.2026.40

ISBN
978-961-299-152-4

Small and medium sized manufacturers (SME) face recurring disruptions which can be addressed by an increase of resilience. Current solutions based on Artificial Intelligence (AI) offer a possibility to meet requirements in terms of resilience while offering high business efficiency. However, the identification of suitable use cases shows a high complexity due to its innovative character and multiple solutions being introduced in a short time. This paper proposes an approach for early-stage use case identification that combines corporate domains to frame the application context, resilience dimensions to specify the intended effect, and a compact self-assessment of technical and organizational prerequisites. The approach derived to identify use cases is implemented in a web-based IT-demonstrator that guides users through four decision layers and delivers an individualized selection of AI use cases aligned with the SME context and resilience objective. To test comprehensibility, practical plausibility, and usefulness across heterogeneous SME conditions the demonstrator and selection logic were refined and validated with industry experts.

Keywords:
AI use case selection, production resilience, SMEs, implementation readiness, IT-demonstrator

University of Maribor Press

1 Introduction

Resilience of production systems has gained importance recently due to crises such as COVID-19, international conflicts, and tariffs. Such events can significantly affect businesses and must be assessed by crisis type and duration (Lund et al., 2020). In a crisis, production materials can show volatile prices or even be unprocurable (Bardt, Grömling, & Schmitz, 2022). Resilience addresses these challenges and can be defined as the ability to withstand and recover from disruptions with potential impact on the production processes (Lerch, Jäger, & Horvat, 2022; Luo et al., 2024; Mwangola, 2018). To do so, companies must identify, evaluate and mitigate production risks (Romero, Stahre, Larsson, & Ronnback, 2021).

Resilience originated in material science as the ability to return to an original state after a load. By extension, a resilient production system can anticipate, repel, and recover from disruptions (Lerch et al., 2022; Mwangola, 2018). In contrast to mere robustness, a resilient system goes beyond recovering to its original state by actively adapting to new conditions and transforming underlying processes (Bauer, Böhm, Bauernhansl, & Sauer, 2021). A resilient response unfolds in five phases: preparing, preventing, protecting, responding, and recovering (Kohl et al., 2021; Maier, Puppala, & Oberle, 2024). It thus requires anticipation, robustness, agility, flexibility, adaptivity, and learning (Kohl et al., 2021; Wildemann, 2025).

A resilient production system must therefore fulfil two functions: maintaining production despite disturbances and recovering to a state equal or better than before (Mockenhaupt & Schlagenhauf, 2024; Wildemann, 2025). Disturbances can arise from events such as the pandemic named above. Internal disturbances stem from operational risks such as machine breakdowns, employee illness, or quality defects. Resilience can be actively increased through management approaches such as resilience engineering (Romero et al., 2021; Wenig, Amann, & Wölflick, 2024). Traditional measures focus on redundancies, overcapacities, and safety stocks. These approaches are typically associated with a loss of efficiency in terms of the profitability of a production system by increasing the capital required for a similar output (Ganin et al., 2017). In general, resilience investments should be allocated by expected impact. Artificial Intelligence (AI) opens new possibilities to target measures to increase the resilience of production systems and to overcome the apparent contradiction of resilience and profitability.

AI describes digital systems which can imitate human behavior. Unlike classical digital solutions AI systems can handle unfamiliar situations by relating new tasks to prior knowledge (Seifert et al., 2018). These capabilities allow AI to predict risks and disturbances from available data, strengthening anticipation. Their capacity to process large datasets also can improve transparency over the state of the production system. Building on this, AI can rapidly adapt production plans, increasing agility, flexibility and adaptivity.

Successful AI implementation requires understanding both how AI works and which applications fit a given company's structure. In fact, many companies have already invested heavily in AI but struggle with a lack of returns, data availability or the lack of expertise in their companies (Challapally, Pease, Raskar, & Chari, 2025; IDC, 2025). This paper therefore addresses the research question:

How can companies with internal production systems efficiently identify use cases to increase their production resilience without negatively impacting their profitability?

This paper contributes to the existing literature by providing an approach to implementing AI based on empirical study results particularly focusing on the identification of specific AI use cases. It also offers practitioners a guideline for assessing their own capabilities regarding the implementation of AI use cases building up on case studies from other companies.

2 Theoretical and Technological Foundation

AI relies on statistical models that employ different methods according to the task. AI methods can be differentiated in supervised learning, in which a model is trained to fulfil a specific task on given data, in semi-supervised learning in which the decisions are not fully previously specified and unsupervised learning in which the AI derives conclusions from given data by itself. A leap forward in the application and public reception of AI was achieved by the introduction of generative AI in combination with foundational models. This combination enabled new content to be created through natural language interaction between user and AI in form of human speech. At the same time, the barrier to adopting AI in a business has decreased as the IT expertise required of users has decreased.

2.1 AI in the context of resilience

Building resilience in a company requires addressing multiple interacting dimensions in managing risks. These dimensions include financial, operative, technological, organizational and reputational resilience (Wildemann, 2022). Financial resilience requires sufficient liquidity to cover unexpected debt defaults or sales drops. Operative resilience maintains production stability during supply or demand disruptions, technological resilience covers the ability to adapt digital systems and infrastructure, organizational resilience addresses strategic governance and decision-making capacity in crises, and reputational resilience concerns public perception and ESG compliance. These goals can be operationalized in the concepts of absorption, adaptability, and transformation (Linkov & Trump, 2019; Rudloff, 2022).

Wildemann (2022) describes eight design principles for company resilience (figure 1), which can be related to the capabilities of AI. In a production setting, AI can forecast demand, volumes and market prices to optimize supply chains and pricing (Iseri, Iseri, Chrisandina, Iakovou, & Pistikopoulos, 2025). AI can augment human decision-making by surfacing and evaluating relevant data. It enhances transparency by measuring and analyzing real-time data, increasing resilience by operationalizing internal and external information (Iseri et al., 2025; Rožanec et al., 2023). AI can also predict machine failures. This predictive-maintenance approach identifies actions needed to keep production running while increasing efficiency through the avoidance of unnecessary repairs. Automated quality control aids in receiving direct feedback for production planning, making human controls obsolete. Combining predictive maintenance, quality control and production planning, improves production planning by enriching the accuracy of the approach.

By analyzing large datasets and automated decision-making, AI can identify atypical business deviations (Baryannis, Validi, Dani, & Antoniou, 2019). Given their low cost, AI-based prediction models help to achieve resilience without compromising profitability. The examples show how AI can strategically strengthen resilience of manufacturing companies. Based on expert-interviews, Merhi and Harfouche (2024) evaluated critical success factors to implementing AI based on the Technology-Organisation-Environment framework, resulting in a critical relevance of top management support for successful implementation (Tornatzky & Fleischer, 1990). Improving resilience through AI can therefore be seen as a strategic concern.

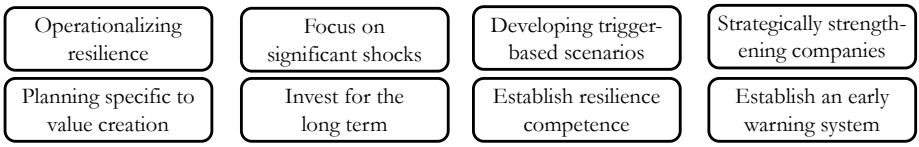

Figure 1: Design principles for increasing resilience
Source: Own depiction related to Wildemann (2022)

The 'invest for the long term' principle reflects that resilience by itself produces value only when a shock occurs. Companies must therefore prepare the required structures before a shock arrives which demands a high degree of foresight and management commitment to invest accordingly. Besides top management support, factors such as perceived compatibility, perceived benefits, data quality, technical experience and complexity also drive successful AI implementation (Merhi & Harfouche, 2024; Schwaeke, Peters, Kanbach, Kraus, & Jones, 2025). Building competence in resilience and AI requires clearer guidance about actual compatibility of possible approaches and the necessary means to achieve a running system. By analyzing external data such as news reports, AI can support the identification of events that may threaten business continuity (Kohl et al., 2021). AI thus enables automated early warning systems. Automating external data evaluation reduces the necessary manual time invest, reducing the cost to increase resilience.

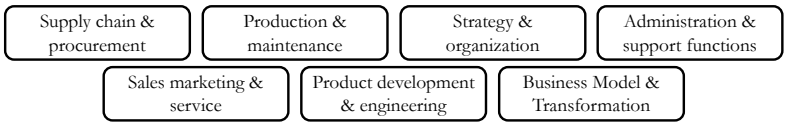

Figure 2: Corporate areas for increasing resilience
Source: Own depiction related to Porter and Millar (1985)

To make resilience actionable, this paper differentiates distinct corporate domains. This functional decomposition follows Porter's value chain logic, which structures firm activities by area of value creation (Porter & Millar, 1985). Based on this, this paper distinguishes the domains: Supply Chain and Procurement; Production and Maintenance; Strategy and Organization; Administration and Support Functions; Sales, Marketing and Service; Product Development and Engineering; Business Model and Transformation. The domains capture the levers of resilience across

operational and strategic time frames, ranging from short-term absorption and operational recovery to long-term adaptation and transformation.

Supply Chain and Procurement are critical for production resilience, since disruptions directly affect material availability and production continuity (Nugroho, Santosa, & Khoa, 2025). Production and Maintenance shape resilience through the system's ability to absorb and recover from disruptions. Effective maintenance strategies, especially those that integrate preventive and predictive approaches, reduce unplanned downtime and improve the availability of a system. This enables production to continue during disturbances and recover faster afterwards (Wachter, Beckschulte, Hinrichs, Sohnius, & Schmitt, 2024). Strategy and Organization enhance production resilience by embedding resilience goals into strategic planning, governance, and decision-making. This strengthens the company's ability to anticipate disruptions, allocate resources, and coordinate cross-functional responses (Toro-Gallego, Sapena-Bolufer, Plaza-Navas, & Torres-Pruñonosa, 2025). Administrative and support related processes can improve resilience through reliable management of data, structured failure documentation, and continuous information exchange across organizational levels. Robust digital infrastructure and systematic failure management reduce downtime and support the adaptability during disruptions (Wachter et al., 2024). Sales, Marketing and Service affect operational stability through demand patterns, customer behavior, and market communication, which determine planning reliability and resource utilization. Demand shifts and volatility can destabilize schedules when information is delayed or inaccurate (Ivanov, 2021). Product Development and Engineering play a preparatory role, shaping the product and process data that later determines production stability. Well-prepared engineering data enables production systems to handle variants, process changes, and resource constraints more effectively, building resilience before disturbances occur (Meyer, Hinrichsen, & Niggemann, 2023). Business Model and Transformation improve resilience by establishing adaptability and recovery as principles of the value-creation system. This enables production systems to adapt and recover, rather than merely withstand failures (Reifert et al., 2022).

2.2 Implementation Requirements of AI Use Cases

AI can increase the profitability and resilience of production systems, but companies are limited in the resources they can invest. Many firms also face barriers related to data, capabilities, and management of data, that limit the AI deployment (Kolev-Schaefer et al., 2025). Technical requirements are therefore critical success factors, since they determine which use cases are feasible. In context of production, data availability and data quality are central prerequisites, since model outputs depend directly on the underlying datasets and their preparation (Sánchez, Calderón, & Herrera, 2025). Technical and data readiness are essential for sustained operational success, since AI applications require a solid data and technology foundation (Uren & Edwards, 2023). In manufacturing, quality assurance for AI relevant datasets should address completeness, correctness, consistency, uniqueness, timeliness, and accuracy, because missing or incomplete data can make ML applications unusable or highly error prone (Wiemer, Dementyev, & Ihlenfeldt, 2021).

Beyond technical requirements, assessing organizational prerequisites is equally critical since AI initiatives depend on coordinated decisions about priorities, responsibilities, compliance and change management across the firm (Papagiannidis, Enholm, Dremel, Mikalef, & Krogstie, 2023). AI governance is therefore a central enabling condition. First, explicit leadership support enables prioritization, funding and conflict resolution during implementation (Alzahrani, 2025). Second, AI initiatives must be strategically anchored in alignment with the organization's overarching goals (Prasetyo, Peranginangin, Martinovic, Ichsan, & Wicaksono, 2025). Third, legal and governance readiness requires awareness of regulatory expectations and the ability to translate them into internal rules, responsibilities and control mechanisms for trustworthy implementation (Papagiannidis et al., 2023). Fourth, culture and workforce readiness matter, since adoption can stall when employees perceive AI as threatening, making transparent communication and education essential (Alzahrani, 2025).

3 Methodology

This research is part of the publicly funded research project ResProKI (IGF 01IF23407 N / 1, October 2024-September 2026). This paper is passed on a sequential multi-method research design to identify, structure, and validate AI use

cases for strengthening production resilience in SMEs. The work follows the Design Science Research Methodology, which proceeds through problem identification, definition of solution objectives, design and development, demonstration and evaluation (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007). First, an evidence base was built by analyzing documented case material and conducting a structured literature review (vom Brocke et al., 2009). The review combined the search topics artificial intelligence, machine learning, production resilience and AI use case identification. In parallel, a market screening of available AI solutions and vendors was conducted in August 2025 to identify practice-oriented application clusters and implementation patterns. These inputs were consolidated into an initial use-case inventory and a preliminary mapping logic linking use cases to corporate domains and resilience dimensions. Second, qualitative expert interviews with SME representatives and AI practitioners were conducted as problem centered interviews to identify resilience related pain points and constraints. The findings were consolidated into a structured problem perspective and served to derive and specify candidate use cases and assessment criteria. Third, the design and development phase involved implementing the mapping logic and selection procedure in an IT-demonstrator, thereby translating the conceptual approach into an operational artefact. Demonstration was embedded in this phase by applying the demonstrator to representative inputs and use case scenarios to verify the internal consistency of the decision layers and to check whether the generated shortlists were plausible before conducting the formal evaluation.

The demonstrator and its underlying selection logic were validated within three sessions of focus group interviews with SME representatives (n=5 on 08.07.2025; n=8 on 15.01.2026; n=5 on 16.03.2026). They were complemented by individual expert interviews (n=6) conducted from August 2025 to March 2026. The discussions focused on plausibility, clarity, and applicability across heterogeneous SME contexts.

4 Design and Development

The artefact developed is a web-based IT-demonstrator that operationalizes the selection logic into a guided procedure for SMEs. Its purpose is to reduce ambiguity in early-stage use case exploration by combining contextual scoping with a structured assessment of prerequisites. Instead of producing a generic catalogue, the

demonstrator links use cases to the user's operational context and filters them against the organization's current ability to implement data driven applications in a stable manner. The selection logic follows a four-step procedure. Users first select the corporate domain in which resilience shall be strengthened, namely Supply Chain and Procurement, Production and Maintenance, Strategy and Organization, Administration and Support Functions, Sales Marketing and Service, Product Development and Engineering, or Business Model and Transformation.

They then select the intended resilience dimension, namely financial, operative, technological, organizational, or reputational resilience. Based on this contextual framing, the demonstrator captures technical and organizational prerequisites through a self-assessment that reflects the company's internal maturity for the specified criteria. Among the technical prerequisites, users have to evaluate the availability of data including the evaluation of whether data is collected in central hubs. As the second criterion, the availability of a digital infrastructure in the company is evaluated. The last technical prerequisite is the availability of digital technologies in a company and its use. On the side of the organizational prerequisites, users assess: management support for introducing digital technologies, the strategic anchoring of AI to increase resilience; knowledge of applicable legal frameworks and employee and cultural support for adopting digital technologies. As the presented evaluation is based on a qualitative assessment, each criterion is rated on a three-level scale from scarcely fulfilled to fully fulfilled and mapped to numeric values from 1 to 3, which are aggregated across criteria to form technical and organizational prerequisite scores. The three-point scale was chosen for ease of use: it lets users benchmark internal prerequisites against peers easily as 'better,' 'average,' or 'weaker without requiring finer-grained distinctions. The combined scores position the company in a 2x2 matrix that classifies use cases as low- or high-requirement on each axis enabling a prerequisite-based preselection of candidates. In the assessment, the technical layer focuses on whether the company has the data basis and digital infrastructure required for AI use cases, in particular systematic data capture, accessible and consistent information flows across systems, and a sufficiently mature baseline of digital tooling in daily operations. The organizational layer captures whether a technically feasible use case can be implemented and sustained. It assesses whether digital initiatives are supported by organizational policy and governance, whether AI is anchored as a strategic topic rather than treated as isolated experimentation, whether legal and compliance implications are

sufficiently understood to manage data protection and regulatory requirements, and whether the organizational culture supports adoption of new tools and changes in established work routines. Both prerequisite assessments use the same compact ordinal scale, supporting structured reflection, comparability across companies, and the derivation of an overall readiness profile. From the combined profile, the demonstrator produces a shortlist of use cases that match the selected context and are consistent with the assessed prerequisites.

5 Evaluation

The logic as a basis for the web-based IT-demonstrator implementing the selection logic, was validated based on a static mock-up in an application oriented evaluation following the DSRM to examine its applicability and practical usefulness in industrial settings (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007; Ulrich, 1981). Accordingly, the evaluation was conducted with SME representatives from the fashion and food industries, metalworking industry, the packaging industry, and supplier industries for mechanical engineering and production technology, using focus group interviews and individual interviews (Mayer, 2013). The focus groups reviewed the overall logic of the approach, including the selection procedure and the clarity of criteria and scales, with a focus on completeness and practical plausibility. Within the workshop, the participants applied the demonstrator's prerequisite criteria to their companies, providing application-oriented data on their own technical and organizational readiness. Regarding the technical requirements, data capture was rated predominantly mid-scale, with companies describing production and inventory data as distributed across multiple sources and partially maintained in parallel spreadsheets. At the same time, a smaller subset reported a consolidated ERP environment with structured cross-departmental access. On the organizational side, awareness of legal frameworks showed the widest spread, ranging from unclear internal responsibilities and missing AI guidelines (with the EU AI Act and data protection named as central uncertainties) to established processes supported by external advisors. Cultural acceptance ranged similarly from resistance to digital change to high openness, with adoption depending strongly on visible benefits in daily work and accompanying training formats. In a second step, the logic for the demonstrator was assessed in semi structured individual interviews with SME experts, focusing on usability, the clarity and perceived transparency of the guided procedure, and the usefulness of the generated shortlist (Dicicco-Bloom & Crabtree,

2006). After applying the feedback from the initial logic validation, the demonstrator was implemented and evaluated with a focus group of SME experts to validate the practical feasibility of implemented decision layers under realistic user interaction.

The use case catalogue behind the IT-Demonstrator helps SMEs to efficiently identify suitable use cases by pre-selecting candidates and summarizing each one's functionality, goals and benefits, and technical and organizational requirements. By comparing the achievable benefits with the requirements, SME-representatives can easily identify the viability of the use case including the potential impact regarding the efficiency and resilience impact. During the pre-evaluation of the AI use-cases it was identified that most of these cases were aimed at an efficiency increase while positively impacting the resilience factors at the same time. These effects are shown as use case summary presented to SME representatives in the IT-Demonstrator.

6 Discussion

In the literature, multiple papers were identified which give an overview over potential use cases of AI in the production (e.g. Toorajipour, Sohrabpour, Nazarpour, Oghazi, & Fischl, 2021) grouping AI use cases according to implementation area or AI methodologies. Early-stage exploration of AI use cases for production resilience benefits from a structured decision logic that includes application context, intended resilience effect, and implementation prerequisites. By combining corporate domains and resilience dimensions, the approach specifies the solution space in a targeted way by surfacing use case candidates that fit the company's individual context and resilience objective, rather than presenting generic AI options. The collected use cases were grouped according to the prerequisites and the capability assessment or potential users. The prerequisite assessment structures the solution space by indicating which candidates are plausible under the organization's current technical and organizational conditions, thereby supporting follow up feasibility work. In practice, the IT-demonstrator provides a reusable entry point to offer SMEs additional context specific use cases and to generate an individualized shortlist before committing resources to detailed evaluation and implementation planning. The central benefit lies in structured approach to finding feasible AI applications: the catalogue leads employees to comparable, context-specific examples and provides reference points from which further applications can be derived through analogical transfer in their own operational environment. By

replacing abstract technology discussions with concrete use cases, the demonstrator lowers the threshold for initiating AI projects. For research, the paper contributes an operational mapping logic between domains, resilience dimensions, and readiness criteria implemented in a usable demonstrator.

Limitations of the presented work can be found in the fact that no practical implementation of the collected use cases was executed during the project to verify the suitability of the proposed solutions for the company representative using the IT demonstrator. However, it was mentioned by the participants of the validation focus group that the use cases in the current state can be used as an inspiration for employees identifying potential use cases. At the same time, the participants mentioned an additional limitation that an enrichment of the available information would be desirable to include company internal information regarding the success of a use case and suppliers which can be re-used by different departments.

The evaluation focused on usability and practical plausibility and did not measure implementation outcomes in operational deployments. Future work should involve longitudinal testing of the demonstrator in SMEs and quantify effects on the quality of decisions for use cases and the following implementation success. Furthermore, it should extend the approach from guidance at the identification of use cases to quantified evaluation by integrating cost-benefit logic. Additionally, the mapping logic should be refined through broader empirical applications, including cross-industry validation and updates to reflect emerging AI solution classes.

7 Conclusion

This paper developed a four-layer approach for identifying AI use cases that can support production resilience in SMEs. The approach combines corporate domains, resilience dimensions and technical as well as organizational requirements to narrow a set of possible AI applications to a context specific shortlist. By implementing this logic in a web-based IT demonstrator, the study provides a structured entry point for SMEs that need to explore AI opportunities but face limited resources and different implementation conditions.

The main contribution lies in shifting early use case exploration from a generic technology search to a guided selection process based on resilience objectives and readiness criteria. The evaluation indicates that the demonstrator is understandable and practically plausible as a first step before detailed investment assessments.

This article was written as a part of the project "Concept of a resilient and efficient production system by using Artificial Intelligence in SME", which was funded by the Federal Ministry for Economic Affairs and Energy under the reference IGF 01IF23407 N / 1

References

Alzahrani, S. (2025). Assessing artificial intelligence and advanced analytics adoption: A technological, organizational, and environmental perspective. *Edelweiss Applied Science and Technology*, *9*(7), 312–325. https://doi.org/10.55214/25768484.v9i7.8567

Bardt, H., Grömling, M., & Schmitz, E. (2022). Wirtschaftliche Folgen des Ukraine-Krieges: Zunehmende Belastungen für die deutsche Wirtschaft. *IW-Report*.

Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2019). Supply chain risk management and artificial intelligence: state of the art and future research directions. *International Journal of Production Research*, *57*(7), 2179–2202. https://doi.org/10.1080/00207543.2018.1530476

Bauer, D., Böhm, M., Bauernhansl, T., & Sauer, A. (2021). Increased resilience for manufacturing systems in supply networks through data-based turbulence mitigation. *Production Engineering*, *15*(3-4), 385–395. https://doi.org/10.1007/s11740-021-01036-4

Challapally, A., Pease, C., Raskar, R., & Chari, P. (2025). *The GenAI Divide: State of AI in Business 2025*.

Dicicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, *40*(4), 314–321. https://doi.org/10.1111/j.1365-2929.2006.02418.x

Ganin, A. A., Kitsak, M., Marchese, D., Keisler, J. M., Seager, T., & Linkov, I. (2017). Resilience and efficiency in transportation networks. *Science Advances*, *3*(12), e1701079.

IDC (2025). *CIO Playbook 2025: It's Time for AI-nomics*.

Iseri, F., Iseri, H., Chrisandina, N. J., Iakovou, E., & Pistikopoulos, E. N. (2025). AI-based predictive analytics for enhancing data-driven supply chain optimization. *Journal of Global Optimization*. Advance online publication. https://doi.org/10.1007/s10898-025-01509-1

Ivanov, D. (2021). Supply Chain Viability and the COVID-19 pandemic: a conceptual and formal generalisation of four major adaptation strategies. *International Journal of Production Research*, *59*(12), 3535–3552. https://doi.org/10.1080/00207543.2021.1890852

Kohl, H., Buß, D., Gebauer, H., Glawar, R., Heller, T., Klan, S., . . . Wilms, M. (2021). *White Paper »RESYST« – Resiliente Wertschöpfung in der produzierenden Industrie – innovativ, erfolgreich, krisenfest*. München: Fraunhofer-Gesellschaft e. V.

Kolev-Schaefer, G., Matthes, J., Schaefer, T., Schmitz, E., Weber, B., & Schmitz-Brieber, J. (2025). *Resilienz der deutschen Lieferketten nach der Zeitenwende: Studie des EPICO KlimaInnovation e.V. in Zusammenarbeit mit dem Institut der deutschen Wirtschaft*. Berlin/Köln.

Lerch, C., Jäger, A., & Horvat, D. (2022). *Resilient in der Krise: Bedeutung von Industrie 4.0 und Lean Management während der COVID-19 Pandemie*. Fraunhofer ISI. https://doi.org/10.24406/PUBLICA-341

Linkov, I., & Trump, B. D. (2019). *The science and practice of resilience*. Springer.

Lund, S., Manyika, J., Woetzel, J., Barriball, E., Krishnan, M., Alicke, K., . . . Hutzler, K. (2020). Risk, resilience, and rebalancing in global value chains. Retrieved from https://www.mckinsey.com/capabilities/operations/our-insights/risk-resilience-and-rebalancing-in-global-value-chains, last accessed 18.02.2026

Luo, X., Kang, K., Lu, L., Yu, C., Li, C., Li, B., . . . Zhou, Y. (2024). Resilience evaluation of low-carbon supply chain based on improved matter-element extension model. *Plos One*, *19*(4), e0301390.

Maier, F. A., Puppala, S., & Oberle, M. (2024). Artificial Intelligence Applications for Resilience in Manufacturing — A Systematic Literature Review, 778–784. https://doi.org/10.1109/ICAIIC60209.2024.10463234

Mayer, H. O. (2013). *Interview und schriftliche Befragung*. Oldenbourg Wissenschaftsverlag. https://doi.org/10.1524/9783486717624

Merhi, M. I., & Harfouche, A. (2024). Enablers of artificial intelligence adoption and implementation in production systems. *International Journal of Production Research*, *62*(15), 5457–5471. https://doi.org/10.1080/00207543.2023.2167014

Meyer, F., Hinrichsen, S., & Niggemann, O. (2023). How to Generate Assembly Instructions with Robotic Process Automation. In AHFE International (Ed.), *AHFE International, Human Interaction & Emerging Technologies (IHIET 2023): Artificial Intelligence & Future Applications* (pp. 629–638). AHFE International. https://doi.org/10.54941/ahfe1004070

Mockenhaupt, A., & Schlagenhauf, T. (2024). *Digitalisierung und Künstliche Intelligenz in der Produktion*. Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-41935-6

Mwangola, W. (2018). Conceptualizing Supply Chain Resilience: An Alternative Dynamic Capabilities Perspective. *American Journal of Management*, *18*(4), 76–88. https://doi.org/10.33423/ajm.v18i4.186

Nugroho, A., Santosa, B., & Khoa, T. M. (2025). The Role of Supply Chain Resilience in Enhancing Operational Efficiency among Manufacturing SMEs. *Journal of Economics and Management*, *3*(2), 34–41. https://doi.org/10.70716/ecoma.v3i2.234

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI Governance: Identifying Best Practices and Potential Barriers and Outcomes. *Information Systems Frontiers : A Journal of Research and Innovation*, *25*(1), 123–141. https://doi.org/10.1007/s10796-022-10251-y

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Porter, M. E., & Millar, V. A. (1985). How Information Gives You Competitive Advantage. *Harvard Business Review*, *63*(4), 149–160.

Prasetyo, M. L., Peranginangin, R. A., Martinovic, N., Ichsan, M., & Wicaksono, H. (2025). Artificial intelligence in open innovation project management: A systematic literature review on technologies, applications, and integration requirements. *Journal of Open Innovation: Technology, Market, and Complexity*, *11*(1), 100445. https://doi.org/10.1016/j.joitmc.2024.100445

Reifert, R.-J., Krawczyk-Becker, M., Prenzel, L., Pavlichkov, S., Khatib, M. A., Hiremath, S. A., . . . Sezgin, A. (2022). Toward Resilience in Mixed Critical Industrial Control Systems: A Multi-Disciplinary View. *IEEE Access*, *10*, 124563–124581. https://doi.org/10.1109/ACCESS.2022.3224425

Romero, D., Stahre, J., Larsson, L., & Ronnback, A. O. (2021). Building Manufacturing Resilience through Production Innovation. In *2021 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)* (pp. 1–9). IEEE. https://doi.org/10.1109/ICE/ITMC52061.2021.9570204

Rožanec, J. M., Novalija, I., Zajec, P., Kenda, K., Tavakoli Ghinani, H., Suh, S., . . . Soldatos, J. (2023). Human-centric artificial intelligence architecture for industry 5.0 applications. *International Journal of Production Research*, *61*(20), 6847–6872. https://doi.org/10.1080/00207543.2022.2138611

Rudloff, B. (2022). Wirtschaftliche Resilienz: Kompass oder Catchword? Welche Fallstricke und Folgeeffekte die EU im Krisenmanagement beachten muss. *2747-5115*.

Sánchez, E., Calderón, R., & Herrera, F. (2025). Artificial Intelligence Adoption in SMEs: Survey Based on TOE–DOI Framework, Primary Methodology and Challenges. *Applied Sciences*, *15*(12), 6465. https://doi.org/10.3390/app15126465

Schwaeke, J., Peters, A., Kanbach, D. K., Kraus, S., & Jones, P. (2025). The new normal: The status quo of AI adoption in SMEs. *Journal of Small Business Management*, *63*(3), 1297–1331. https://doi.org/10.1080/00472778.2024.2379999

Seifert, I., Bürger, M., Wangler, L., Christmann-Budian, S., Rohde, M., Gabriel, P., & Zinke, G. (2018). Potenziale der künstlichen Intelligenz im produzierenden Gewerbe in Deutschland. Retrieved from https://www.bmwk.de/Redaktion/DE/Publikationen/Studien/potenziale-kuenstlichen-intelligenz-im-produzierenden-gewerbe-in-deutschland.pdf?__blob=publicationFile&v=19

Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. *Journal of Business Research*, *122*, 502–517. https://doi.org/10.1016/j.jbusres.2020.09.009

Tornatzky, L. G., & Fleischer, M. (1990). *The processes of technological innovation. Issues in organization and management series*. Lexington, Mass.: Lexington Books.

Toro-Gallego, C., Sapena-Bolufer, J., Plaza-Navas, M.-A., & Torres-Pruñonosa, J. (2025). The Science of Organisational Resilience: Decoding Its Intellectual Structure to Understand Foundations and Future. *Administrative Sciences*, *15*(10), 404. https://doi.org/10.3390/admsci15100404

Ulrich, H. (1981). Die Betriebswirtschaftslehre als anwendungsorientierte Sozialwissenschaft. In M. Geist & R. Köhler (Eds.), *Die Führung des Betriebes: Curt Sandig zu seinem 80. Geburtstag gewidmet* (1-23). Stuttgart: Poeschel.

Uren, V., & Edwards, J. S. (2023). Technology readiness and the organizational journey towards AI adoption: An empirical study. *International Journal of Information Management*, *68*, 102588. https://doi.org/10.1016/j.ijinfomgt.2022.102588

Vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. *Proceedings of the European Conference on Information Systems*. (17).

Wachter, C., Beckschulte, S., Hinrichs, M. P., Sohnius, F., & Schmitt, R. H. (2024). Strategies for Resilient Manufacturing: A Systematic Literature Review of Failure Management in Production. *Procedia CIRP*, *130*, 1393–1402. https://doi.org/10.1016/j.procir.2024.10.257

Wenig, C., Amann, A., & Wölflick, M. (2024). Resilienz in globalen Produktionsnetzwerken. *Zeitschrift Für Wirtschaftlichen Fabrikbetrieb*, pp. 171–175.

Wiemer, H., Dementyev, A., & Ihlenfeldt, S. (2021). A Holistic Quality Assurance Approach for Machine Learning Applications in Cyber-Physical Production Systems. *Applied Sciences*, *11*(20), 9590. https://doi.org/10.3390/app11209590

Wildemann, H. (2022, May 16). Resilienz wird zur Kernkompetenz. *Frankfurter Allgemeine Zeitung*, p. 18.

Wildemann, H. (2025). *Resiliente Unternehmensführung: Leitfaden zur Krisenbewältigung und Transformation von Unternehmen* (4th ed.). München: TCW Transfer-Centrum für Produktions-Logistik und Technologie-Management GmbH & Co. KG.

Big Five Personality Traits and Security Awareness Training Preferences

Isabel K. Glückert, Tobias Fertig, Ricard Bibaj, Sebastian Biedermann

TTZ-WUE Cyber Security - Technical University of Applied Sciences Würzburg-Schweinfurt, Würzburg Germany
isabel.glueckert@thws.de, tobias.fertig@thws.de, ricard.bibaj@thws.de, sebastian.biedermann@thws.de

Individual behavior plays a central role in information security, as many incidents result from user behavior. Accordingly, Information Security Awareness training has become a central measure to support individual security-related decision-making. However, standardized ISA training often shows limited effectiveness, as it rarely accounts for individual differences in learning preferences and motivation. This study examines how personality traits relate to ISA training preferences. A quantitative survey assessed personality using the BFI-2 and collected preference ratings for 18 training methods. Significant relationships were found for Agreeableness and Conscientiousness. Higher Agreeableness was associated with lower preference for socially exposed methods, while individuals high in Conscientiousness preferred structured and practice-oriented methods. The findings suggest that personality shapes how ISA training is perceived and highlight the importance of considering individual differences in the design of training methods. Providing a variety of training methods, rather than a single standardized approach, may improve perceived fit and acceptance.

DOI
https://doi.org/10.18690/um.fov.4.2026.41

ISBN
978-961-299-152-4

Keywords:
information security awareness, big five personality model, training preferences, human factors, personalization

University of Maribor Press

1 Introduction

Information security has become a central organizational challenge in light of increasing digitalization and attack sophistication (BSI, 2024; ENISA, 2024). A substantial share of cybersecurity incidents is attributable to human factors (IBM, 2025; Ugbebor et al., 2024), including compromised legitimate accounts and social engineering attacks exploiting stress, trust, or curiosity (Proofpoint, 2023; Ugbebor et al., 2024). Accordingly, human behavior is a critical determinant of organizational security (Helisch & Pokoyski, 2009; Khadka & Ullah, 2025; Parrish et al., 2009; Schütz, 2018; Weber et al., 2020). To mitigate these risks, organizations rely on Information Security Awareness (ISA) trainings (Bada et al., 2019; BSI, 2024; ENISA, 2024). However, traditional awareness trainings focus on knowledge transfer and often show limited impact on security behavior (Bada et al., 2019; Fertig & Schütz, 2020; Weber et al., 2020). Knowledge alone is insufficient, as individuals behave differently despite comparable knowledge or intentions (Ajzen, 1991). Standardized ISA trainings rarely account for individual differences in motivation and learning styles (Shappie et al., 2020). Personality is a central factor for individual behavioral differences (DeYoung, 2015). Empirical research shows that Big Five personality traits are associated with security-relevant attitudes and behaviors (McCormac et al., 2017; Shappie et al., 2020). Consequently, psychological factors must be considered in awareness training design to achieve long-term improvements in security-relevant behavior (Bada et al., 2019; Fertig & Schütz, 2020; Proofpoint, 2023; Weber et al., 2020). However, it remains unclear how personality traits are reflected in preferences for different ISA training methods. Addressing this gap, this study examines the relationship between the Big Five personality traits and ISA training preferences using a quantitative survey. The focus is on perceived preferences rather than on behavioral outcomes or training effectiveness. This leads to the following research question and hypothesis:

RQ1: How are the Big Five personality traits related to individual preferences for different Information Security Awareness (ISA) training methods in academic and work contexts?

H1: If individuals differ in their Big Five personality traits, they also differ in their preferences for different Information Security Awareness (ISA) training methods.

The paper is structured as follows: Section 2 presents the theoretical foundations and reviews prior research on information security awareness and personality. Section 3 outlines the methodology of the quantitative study. Section 4 reports the results, which are discussed in Section 5. Finally, Section 6 summarizes the key findings and highlights implications for future research.

2 Related Work

Information security is commonly conceptualized through the CIA triad, encompassing confidentiality, integrity, and availability (ISO, 2018). ENISA (2024) defines information security as the interplay of technical, organizational, and human measures to protect information assets. However, security programs remain predominantly technology-oriented and insufficiently consider behavioral factors (ENISA, 2024; Khadka & Ullah, 2025; SANS Institute, 2025). Approximately 85% of cybersecurity incidents are attributable to human errors (Ugbebor et al., 2024). Consequently, end-user behavior plays a central role in preventing cyber-attacks (Prümmer et al., 2024) and ISA has emerged as a distinct research field (Fertig & Schütz, 2020). ISA refers to employees' awareness of security-related risks and their understanding of security rules (Fertig et al., 2022; Richter et al., 2018). The relevance of awareness is particularly evident in phishing, one of the most common attack methods that exploits human errors and cognitive vulnerabilities (2025 Data Breach Investigations Report, 2025; Bada et al., 2019; ENISA, 2024; Schütz, 2018).

2.1 Limitations of Standardized Awareness Training Approaches

Increasing threat complexity places a burden on security professionals, 66% of whom perceive their work as more stressful than in 2020 (ISACA, 2025). Accordingly, organizations invest substantial resources in ISA trainings, recognizing them as a central measure (ENISA, 2024). This growth is reflected in a global ISA training market of USD 4.30 billion in 2023, projected to exceed USD 21 billion by 2032 (Zion Market Research, 2024). Despite these investments, the behavior-changing impact of ISA trainings remains limited (Bada et al., 2019). This is confirmed by large-sample studies reporting only marginal reductions in misclicks and high dropout rates (Ho et al., 2025). This limited effectiveness contributes to ISA trainings being perceived as a formal obligation rather than supportive assistance (Bada et al., 2019; Hornetsecurity GmbH, 2024; SANS Institute, 2025).

Consequently, knowledge-based training approaches have limited influence on security behavior (Bada et al., 2019; Bilal Khan, 2011; Fertig & Schütz, 2020; Proofpoint, 2023; Schütz, 2018; Weber et al., 2020). The failure of classical training approaches to account for individual adaptation has long been recognized (Eighteen, 1999). Accordingly, behavioral change depends on multiple individual factors. A central framework in this context is the Integrated Behavioral Model (IBM), which conceptualizes human behavior as the result of multiple factors (Montaño & Kasprzyk, 2008). In security contexts, the IBM shows that security-compliant actions often fail despite knowledge, while standardized training inadequately addresses employee needs (Schütz, 2018). These findings highlight the need to consider individual differences for sustainable behavioral change (Proofpoint, 2023).

2.2 Learning-Theoretical Foundations of ISA Training

Learning is commonly understood as an active process in which individuals select, organize, and integrate information with prior knowledge to construct new understanding (Jonassen & Rohrer-Murphy, 1999; Mayer, 2002). Information processing occurs through different channels, such as verbal and non-verbal systems (Clark & Paivio, 1991). At the same time, learning is constrained by the limited capacity of working memory. The design of instruction significantly influences the cognitive load of learners, and inappropriate teaching methods can impose additional cognitive burden (Sweller, 2010). In the context of information security, these challenges become particularly evident. Although ISA training methods can improve users' knowledge and attitudes, their impact on actual security-related behavior remains limited (Prümmer et al., 2024) and behavioral outcomes are influenced by individual factors (Schütz, 2018). Against this background, it becomes evident that the selection of appropriate ISA training methods is crucial for learning success and depends on individual characteristics. In information security, where behavioral change often remains limited despite training, it is relevant to examine how personality traits relate to preferences for different ISA training methods.

2.3 Personality as an Influencing Factor on Security Behavior

Personality influences how individuals respond to security-relevant requirements and represents a central leverage point for designing effective ISA training methods (DeYoung, 2015; Shappie et al., 2020). This is partly because personality influences

how individuals learn and process information. People perceive identical situations differently, meaning that the same training experiences are not equally effective for all learners (Chamorro-Premuzic, 2005). Established personality models such as the Big Five have long been used to capture individual differences (Rammstedt & John, 2007). The Big Five model describes five stable traits, Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness, and serves as a framework for explaining individual differences in security-relevant behavior (McCrae & John, 1992; Roccas et al., 2002; Shappie et al., 2020). Empirical findings in the context of learning processes show that Big Five personality traits are related to different learning styles and influence how individuals process information and apply learning strategies (Komarraju et al., 2011). These differences are also reflected in security contexts, where personality traits are associated with security-related attitudes, behavioral intentions, and actual behavior (Grandhi & Still, 2025; McCormac et al., 2017; Shappie et al., 2020; Uffen & Breitner, 2013). Early studies indicate that personality traits, particularly Conscientiousness and Agreeableness, often represent stronger predictors of security-compliant behavior than technology-related attitudes (Shropshire et al., 2006). More recent research considers all five personality traits (Grandhi & Still, 2025; Shappie et al., 2020). These findings indicate that higher Conscientiousness, Agreeableness, and emotional stability are associated with greater security awareness, whereas higher risk propensity shows opposite effects (McCormac et al., 2017). Accordingly, the Big Five model provides a theoretical foundation for explaining individual differences in security-relevant behavior and designing target-group-specific ISA training methods. These individual differences imply that preferences for training methods vary across individuals (Chamorro-Premuzic, 2005) and are reflected in the design of ISA training methods, as different methods provide distinct learning experiences.

2.4 Types and Effectiveness of ISA Training Methods

As personality traits influence learning processes and responses to security requirements (Shappie et al., 2020), the selection of ISA training methods becomes critical. Text-based ISA methods are widely used but show limited effects on security behavior and sustainable change (Bada et al., 2019). Given the increasing importance of digital training methods, e-learning has become a central instrument in organizational training and development (Overton, 2023). Around 91% of organizations rely on technology-supported learning formats, including e-learning

courses, webinars, and learning platforms. These technologies enable flexible, accessible learning independent of time and location (Overton, 2023) and are widely used in organizational information security awareness programs (Fertig et al., 2019). Beyond e-learning, game-based methods foster active participation and perspective-taking, enhancing understanding of social manipulation and phishing (Fatima et al., 2019; Prümmer et al., 2024). Presentation-based methods support structured knowledge transfer (Prümmer et al., 2024; Sykosch et al., 2020). The literature identifies video-based, text-based, and discussion-oriented formats (Prümmer et al., 2024). The evaluation of ISA training preferences forms the study's empirical core.

3 Methodology

This study employed a quantitative research design (Creswell & Creswell, 2018) to examine relationships between Big Five personality traits and ISA training preferences. While qualitative approaches explore individual experiences, this study adopts a quantitative design to identify relationships between personality traits and ISA training preferences across a larger sample. Individual differences are captured through standardized personality measures, enabling comparison of personality profiles. This approach ensures sufficient variability across personality traits and allows identifying generalizable patterns across individuals, which would be difficult to capture using purely qualitative methods. Data were collected via an online survey including the Big Five Inventory-2 (BFI-2) (Soto & John, 2017) and ratings for 18 ISA training methods. Personality was assessed before evaluating ISA training methods to avoid priming effects (Schwarz & Strack, 1991). Big Five traits were treated as independent variables, and training preferences as dependent variables.

3.1 Measurement Instruments and Variables

Participants' personality traits were assessed using the Big Five Inventory-2 (BFI-2), a validated instrument measuring the five traits Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness (Soto & John, 2017). The questionnaire comprises 60 items rated on a five-point Likert scale (Likert, 1932). The analysis was conducted at the trait level, focusing on higher-order personality traits and their relationship with ISA training preferences. Subsequently, ISA training preferences were assessed using a literature-based categorization system distinguishing seven training categories (Prümmer et al., 2024). Based on this

framework and complemented by informal expert consultations, 18 ISA training methods were identified. The consultations involved three practitioners (e.g., IT security managers and awareness specialists) and served to validate the practical relevance and completeness of the selected methods. They were conducted in an exploratory manner without a formal qualitative design. Accordingly, no structured interview protocol or systematic coding procedure was applied. These methods include passive methods such as videos, interactive approaches such as workshops and simulations, and game-based methods such as gamification. The selection of training methods aimed to capture diverse personality profiles and preference patterns. To ensure uniform understanding, each of the 18 ISA training methods was accompanied by a brief standardized description outlining its procedure to minimize bias from differing interpretations. Participants rated their preference for each ISA training method on a five-point Likert scale (1 = not preferred at all to 5 = strongly preferred). The full instrument is provided in Appendix A.

Table 1: Assignment of ISA training methods to categories

Training category	ISA Training methods
Game-based Training	Security competitions, quizzes, gamified learning
Presentation-based Training	Web seminars, expert lectures, awareness-themed days
Simulation-based Training	Simulations, hands-on learning, role plays, case studies
Information-based Training	E-learning platforms, podcasts, mobile learning applications
Video-based Training	Videos
Text-based Training	Documentation
Discussion-based Training	Group work, discussion rounds, mentoring programs

Source: (Prümmer et al., 2024)

3.2 Data Collection and Sample

Data were collected via an online LimeSurvey survey. Participants were recruited via institutional mailing lists, professional networks, and digital channels. Eligible participants had basic experience or regular exposure to IT- and information-security-related topics in academic or work contexts. To ensure meaningful evaluation, a screening question assessed prior experience with learning or training methods in academic or professional contexts. Only participants confirming such

experience proceeded with the survey. This open formulation avoided restricting participation to a specific context and enabled a broader assessment of training preferences. In total, 173 individuals participated. Data cleaning ensured data quality. Incomplete responses were removed, and only fully completed questionnaires without missing values were included. Responses were also screened for plausibility, and datasets with indications of inattentive responding (e.g., uniform response patterns across items or straight-lining) were removed. After data cleaning, a total of 134 valid responses remained for analysis. The sample consisted predominantly of students (80%) and employees from organizations with established ISA training programs (20%). The focus on students is justified, as higher education institutions are more frequently affected by cyberattacks than companies, with 92% reporting incidents compared to 39% of companies (GOV.UK, 2022). Students were enrolled in IT-related degree programs and had experience with IT systems, for example through internships or part-time employment.

4 Results

This section presents empirical results, focusing on relationships between Big Five personality traits and ISA training preferences, including reliability and correlations.

4.1 Reliability and Data Assessment

Internal consistency was assessed using Cronbach's alpha (Cronbach, 1951) for the ISA training items. For the BFI-2 (Soto & John, 2017), established reliability coefficients (.60–.88) were adopted, indicating adequate to good internal consistency (Cronbach, 1951; Danner et al., 2019). For the 18 ISA training items, Cronbach's alpha was .909, indicating excellent internal consistency (Cronbach, 1951). However, this value should be interpreted with caution, as the instrument captures preferences for diverse ISA training methods rather than a unidimensional construct. The high internal consistency may reflect general preference tendencies across training methods and the large number of items. Variable distributions were examined using the Shapiro–Wilk test (Shapiro & Wilk, 1965); significant deviations from normality ($p < .05$) led to the use of Spearman's ϱ for correlation analyses (Spearman, 1904).

4.2 Analysis of Relationships

In total, 90 associations were examined by linking the five Big Five personality traits with the 18 ISA training methods. As the study focuses on preferences for distinct ISA training methods, each method was treated as an individual variable and analyzed using item-level correlation analysis. To examine whether the ISA training methods form underlying dimensions, an exploratory factor analysis was conducted. The results indicated no clear multi-factor structure, as most items loaded on a single dominant factor. This suggests that ISA training preferences reflect general preference tendencies rather than distinct training categories. Therefore, item-level analysis was considered more appropriate to capture specific preference patterns across different training formats. Four significant relationships ($p < .05$) were identified, all relating to Agreeableness and Conscientiousness. Given the large number of tests and the absence of correction for multiple comparisons, some findings may reflect chance. Table 2 summarizes these relationships. No statistically significant associations were identified for Openness, Extraversion, or Neuroticism ($p \geq .05$). Agreeableness was negatively associated with preferences for web seminars ($r = -.177$, $p = .041$) and role plays ($r = -.186$, $p = .031$). In contrast, Conscientiousness was positively associated with preferences for hands-on learning ($r = .246$, $p = .004$) and documentation ($r = .190$, $p = .028$). These results suggest that structured and practice-oriented ISA training methods are preferred by conscientious individuals. With regard to H1, relationships between personality traits and ISA training preferences were observed only for Agreeableness and Conscientiousness, with no effects for Openness, Extraversion, or Neuroticism. The results do not support H1 but indicate selective personality–preference relationships.

Table 2: Significant ISA training–personality correlations

	Agreeableness (mean)	Conscientiousness (mean)
Web seminar	r = -.177; p = .041	-
Role play	r = -.186; p = .031	-
Hands-on learning	-	r = .246; p = .004
Documentation	-	r = .190; p = .028

5 Discussion

Overall, four examined relationships are empirically relevant and associated with preferences for specific ISA training methods, exclusively concerning Agreeableness and Conscientiousness. The discussion therefore focuses on subjective preferences rather than observed behavior or the effectiveness of the ISA training methods.

5.1 Personality-Specific Preference Patterns for ISA Training Methods

For Agreeableness, negative relationships were observed with web seminars and role plays. In this study, web seminars are characterized by online content delivery with limited interaction. In contrast, role plays require active participation, social interaction, and enactment of typical situations (Martin et al., 2014; Sheets, 1998). The lower preference of agreeable individuals suggests that training methods involving mandatory participation or increased social visibility are perceived as less suitable. At first glance, this pattern contradicts the definition of Agreeableness (John & Srivastava, 1999), but it can be explained by considering different levels of analysis. Shappie et al. (Shappie et al., 2020) show that Agreeableness is associated with security-relevant behavior. Accordingly, agreeable individuals tend to comply with rules and display protective behavior toward policies, data, and systems (Shappie et al., 2020). In contrast to these behavior-based findings (Shappie et al., 2020), the present study indicates that agreeable individuals perceive ISA training methods with high social exposure, such as role plays or interactive live training, as less suitable. This lower preference can be interpreted as a desire for harmony and reduced social visibility rather than as a lack of willingness to engage in security-compliant behavior. For Conscientiousness, positive relationships emerge with practice-oriented and structuring ISA training methods, particularly hands-on learning and documentation. In the context of this study, hands-on learning refers to training approaches in which security-relevant content is conveyed through direct application of tasks. Documentation refers to written informational materials that present content in a structured, comprehensible, and accessible manner. Within the Big Five model, Conscientiousness is associated with organization, sense of duty, self-discipline, and goal-directed behavior (John & Srivastava, 1999). Moreover, current research shows that Conscientiousness plays a central role in relation to security-relevant behavior (Shappie et al., 2020). Such behavior manifests in rule compliance, regular software updates, and secure password use (Gratian et al., 2018).

Against this background, it is plausible that individuals with high Conscientiousness prefer ISA training methods that are clearly structured, provide concrete guidance for action, and enable direct implementation. These findings suggest that structured and practice-oriented ISA training methods align well with conscientious individuals. These methods cater to a preference for order and clarity. They also facilitate the sustainable integration of security-relevant content into everyday work practices.

5.2 Non-significant Findings, Limitations, and Implications

No significant relationships were found for Openness, Extraversion, and Neuroticism, although these personality traits are commonly associated with social interaction, emotional reactivity, and openness to new experiences (John & Srivastava, 1999). Accordingly, effects on the evaluation of interactive training methods might have been expected. However, the absence of such effects can be interpreted in light of prior research. The Theory of Planned Behavior highlights the gap between behavioral intentions and actual behavior (Ajzen, 1991; Shappie et al., 2020). Personality may influence security-related behavior without being reflected in subjective training preferences. The exploratory factor analysis further refines this interpretation, as ISA training preferences did not form clearly distinguishable categories. This suggests that participants evaluate training formats more holistically. Consequently, personality-based adaptation should focus on specific training methods rather than relying solely on broad training categories. Several limitations must be considered. First, the study relies on self-reported preferences, and not all participants had comparable experience with the evaluated training methods. As a result, some assessments may be based on expectations rather than experience, and preferences do not allow conclusions about training effectiveness. Second, a large number of relationships were tested, while only a few reached statistical significance, indicating that some findings may be due to chance. Third, the sample consisted predominantly of students, limiting generalizability to organizational contexts. Future studies should include working professionals. These limitations underline the need to complement preference-based measurements with behavior-oriented research designs. This study contributes to information security research by linking personality traits to training preferences, thereby extending prior research beyond behavioral outcomes. While prior research primarily focuses on behavioral outcomes, this study takes an earlier step by examining which training methods are preferred by individuals with different personality profiles. This provides a basis for

selecting and designing ISA training methods. From an educational perspective, these findings suggest that considering individual differences allows ISA training methods to be more closely aligned with learners' preferences and needs. This may increase perceived fit, acceptance of training, and motivation. Given the gap between knowledge, attitudes, and behavior in information security (Schütz, 2018), individualized training methods may help bridge this gap and support more targeted ISA training. Based on these findings, two directions for future research emerge. First, qualitative studies could provide deeper insights into the underlying evaluation processes by examining which characteristics are perceived as suitable or unsuitable by different personality types. These insights can inform the design of more effective ISA training. Second, future research should move beyond preference-based analyses and examine the actual effectiveness of ISA training methods by comparing personality-aligned and non-aligned training methods. This comparison could assess whether aligned formats lead to more positive perceptions, higher engagement, improved knowledge retention, and more stable security-related behavior.

6 Conclusion

This study examined how Big Five personality differences are reflected in preferences for ISA training methods. The results show that preference patterns are meaningful only for Agreeableness and Conscientiousness. Higher Agreeableness was associated with lower preference for web seminars and role plays, while Conscientiousness was associated with structured and practice-oriented ISA training methods such as hands-on learning and documentation. These findings indicate that personality traits shape how ISA training methods are evaluated. The results should be interpreted as exploratory and hypothesis-generating. No significant relationships were found for Openness, Extraversion, and Neuroticism, suggesting that personality differences do not consistently translate into preferences. Overall, this study contributes to information security research by examining the link between personality and ISA training preferences, showing that effects are selective rather than uniform. For practice, the findings suggest that ISA training methods should not be designed as a single standardized format, but rather that a variety of ISA training methods should be provided. Taking personality-specific preferences into account may improve perceived fit and acceptance of trainings, thereby supporting motivation and participation. At the same time, it must be emphasized that preferences do not allow conclusions about the effectiveness of a given method, but

indicate which ISA training methods are perceived as suitable by specific target groups. Future research should examine whether these differences are reflected in the effectiveness of ISA training methods, for example through field studies. Studies capturing knowledge acquisition, engagement and security-relevant behavior would be valuable. Overall, sustainable information security cannot be achieved solely through technical measures or standardized training. Instead, it requires considering individual differences in how ISA training methods are perceived and evaluated.

References

2025 Data Breach Investigations Report. (2025). Verizon Business. https://www.verizon.com/business/resources/reports/dbir/

Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, Theories of Cognitive Self-Regulation, 50, 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Bada, M., Sasse, A. M., & Nurse, J. R. C. (2019). Cyber Security Awareness Campaigns: Why do they fail to change behaviour? (arXiv:1901.02672). arXiv. https://doi.org/10.48550/arXiv.1901.02672

Bilal Khan. (2011). Effectiveness of information security awareness methods based on psychological theories. AFRICAN JOURNAL OF BUSINESS MANAGEMENT, 5(26). https://doi.org/10.5897/AJBM11.067

BSI. (2024, November 12). Die Lage der IT-Sicherheit in Deutschland 2024. Bundesamt für Sicherheit in der Informationstechnik. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2024.html?nn=132646

Chamorro-Premuzic, T. (with Furnham, A.). (2005). Personality and intellectual competence. L. Erlbaum Associates. https://doi.org/10.4324/9781410612649

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. Educational Psychology Review, 3(3), 149–210. https://doi.org/10.1007/BF01320076

Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (fifth Edition). Sage Publications.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297–334. https://doi.org/10.1007/BF02310555

Danner, D., Rammstedt, B., Bluemke, M., Lechner, C., Berres, S., Knopf, T., Soto, C. J., & John, O. P. (2019). Das Big Five Inventar 2: Validierung eines Persönlichkeitsinventars zur Erfassung von 5 Persönlichkeitsdomänen und 15 Facetten. Diagnostica, 65(3), 121–132. https://doi.org/10.1026/0012-1924/a000218

DeYoung, C. G. (2015). Cybernetic Big Five Theory. Journal of Research in Personality, Integrative Theories of Personality, 56, 33–58. https://doi.org/10.1016/j.jrp.2014.07.004

Eighteen, R. (1999). Training needs analysis for IT training. Industrial and Commercial Training, 31(4), 149–153. https://doi.org/10.1108/00197859910275773

ENISA. (2024, September 19). ENISA Threat Landscape 2024. European Union Agency for Cybersecurity. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024

Fatima, R., Yasin, A., Liu, L., & Wang, J. (2019). How persuasive is a phishing email? A phishing game for phishing awareness. Journal of Computer Security, 27(6), 581–612. https://doi.org/10.3233/JCS-181253

Fertig, T., & Schütz, A. (2020, Januar 7). About the Measuring of Information Security Awareness: A Systematic Literature Review. https://doi.org/10.24251/HICSS.2020.798

Fertig, T., Schütz, A. E., Weber, K., & Müller, N. H. (2019). Measuring the Impact of E-Learning Platforms on Information Security Awareness. In P. Zaphiris & A. Ioannou (Hrsg.), Learning and Collaboration Technologies. Designing Learning Experiences (S. 26–37). Springer International Publishing. https://doi.org/10.1007/978-3-030-21814-0_3

Fertig, T., Schütz, A., & Weber, K. (2022, Januar 4). Automated Measuring of Information Security Related Habits. Hawaii International Conference on System Sciences (HICSS). https://doi.org/10.24251/HICSS.2022.925

GOV.UK. (2022, Juli 11). Educational institutions findings annex—Cyber Security Breaches Survey 2022. GOV.UK. https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2022/educational-institutions-findings-annex-cyber-security-breaches-survey-2022

Grandhi, S. R., & Still, J. D. (2025). The Big Five in Action: A Systematic Review of Personality, Cyber Awareness, and Behaviors. In A. Moallem (Hrsg.), HCI for Cybersecurity, Privacy and Trust (S. 17–34). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-92833-8_2

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., & Ginther, A. (2018). Correlating human traits and cyber security behavior intentions. Computers & Security, 73, 345–358. https://doi.org/10.1016/j.cose.2017.11.015

Helisch, M., & Pokoyski, D. (Hrsg.). (2009). Security Awareness. Vieweg+Teubner. https://doi.org/10.1007/978-3-8348-9594-3

Ho, G., Mirian, A., Luo, E., Tong, K., Lee, E., Liu, L., Longhurst, C. A., Dameff, C., Savage, S., & Voelker, G. M. (2025). Understanding the Efficacy of Phishing Training in Practice. 2025 IEEE Symposium on Security and Privacy (SP), 37–54. https://doi.org/10.1109/SP61157.2025.00076

Hornetsecurity GmbH. (2024, Juni 3). Umfrage zur Security Awareness: 1 von 4 (25,7 %) Unternehmen bietet kein Security Awareness Training an. Hornetsecurity. https://www.hornetsecurity.com/de/blog/security-awareness-umfrage-2024/

IBM. (2025). X-Force 2025 Threat Intelligence Index | IBM. https://www.ibm.com/reports/threat-intelligence

ISACA. (2025, September 29). State of Cybersecurity 2025. ISACA. https://www.isaca.org/resources/reports/state-of-cybersecurity-2025

ISO. (2018, Februar). ISO/IEC 27000:2018. ISO. https://www.iso.org/standard/73906.html

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In Handbook of personality: Theory and research, 2nd ed (2nd ed., S. 102–138). Guilford Press.

Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. Educational Technology Research and Development, 47(1), 61–79. https://doi.org/10.1007/BF02299477

Khadka, K., & Ullah, A. B. (2025). Human factors in cybersecurity: An interdisciplinary review and framework proposal. International Journal of Information Security, 24(3), 119. https://doi.org/10.1007/s10207-025-01032-0

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. Personality and Individual Differences, 51(4), 472–477. https://doi.org/10.1016/j.paid.2011.04.019

Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology, 22 140, 1–55.

Martin, B. O., Kolomitro, K., & Lam, T. C. M. (2014). Training Methods: A Review and Analysis. Human Resource Development Review, 13(1), 11–35. https://doi.org/10.1177/1534484313497947

Mayer, R. E. (2002). Multimedia learning. In Psychology of Learning and Motivation (Bd. 41, S. 85–139). Academic Press. https://doi.org/10.1016/S0079-7421(02)80005-6

McCormac, A., Zwaans, T., Parsons, K., Calic, D., Butavicius, M., & Pattinson, M. (2017). Individual differences and Information Security Awareness. Computers in Human Behavior, 69, 151–156. https://doi.org/10.1016/j.chb.2016.11.065

McCrae, R. R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. Journal of Personality, 60(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Montaño, D. E., & Kasprzyk, D. (2008). Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. In Health behavior and health education: Theory, research, and practice, 4th ed (S. 67–96). Jossey-Bass/Wiley.

Overton, L. (2023). Learning at work 2023 survey report. https://www.cipd.org/globalassets/media/knowledge/knowledge-hub/reports/2023-pdfs/2023-learning-at-work-survey-report-8378.pdf

Parrish, J. L., Bailey, J. L., & Courtney, J. F. (2009). A Personality Based Model for Determining Susceptibility to Phishing Attacks. 285–296.

Proofpoint. (2023, Juni 14). Der Faktor Mensch 2025. Proofpoint. https://www.proofpoint.com/de/resources/threat-reports/human-factor-social-engineering

Prümmer, J., van Steen, T., & van den Berg, B. (2024). A systematic review of current cybersecurity training methods. Computers & Security, 136, 103585. https://doi.org/10.1016/j.cose.2023.103585

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality, 41(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Richter, S., Straub, T., & Lucke, C. (2018, September 6). Information Security Awareness – eine konzeptionelle Neubetrachtung. Multikonferenz Wirtschaftsinformatik 2018 (MKWI).

Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The Big Five Personality Factors and Personal Values. https://doi.org/10.1177/0146167202289008

SANS Institute. (2025). 2025 Security Awareness Report. SANS Institute. https://www.sans.org/mlp/ssa-security-awareness-report

Schütz, A. (2018, Mai 17). Information Security Awareness: It's Time to Change Minds! ICDDAt: Sibiu, Romania. International Conference on Applied Informatics (ICAI).

Schwarz, N., & Strack, F. (1991). Context Effects in Attitude Surveys: Applying Cognitive Theory to Social Research. European Review of Social Psychology, 2(1), 31–50. https://doi.org/10.1080/14792779143000015

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples)†. Biometrika, 52(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Shappie, A. T., Dawson, C. A., & Debb, S. M. (2020). Personality as a predictor of cybersecurity behavior. Psychology of Popular Media, 9(4), 475–480. https://doi.org/10.1037/ppm0000247

Sheets, J. (1998). Role-Playing as a Training Tool for Reference Student Assistants. Reference Services Review, 26(1), 37–41. https://doi.org/10.1108/00907329810307425

Shropshire, J., Warkentin, M., Johnston, A., & Schmidt, M. (2006). Personality and IT security: An application of the five-factor model. 6, 415.

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. Journal of Personality and Social Psychology, 113(1), 117–143. https://doi.org/10.1037/pspp0000096

Spearman, C. (1904). The proof and measurement of association between two things. The American Journal of Psychology, 15(1), 72–101. https://doi.org/10.2307/1412159

Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. Educational Psychology Review, 22(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5

Sykosch, A., Doll, C., Wübbeling, M., & Meier, M. (2020). Generalizing the phishing principle: Analyzing user behavior in response to controlled stimuli for IT security awareness assessment. Proceedings of the 15th International Conference on Availability, Reliability and Security, ARES '20, 1–10. https://doi.org/10.1145/3407023.3409205

Uffen, J., & Breitner, M. H. (2013). Management of Technical Security Measures: An Empirical Examination of Personality Traits and Behavioral Intentions. International Journal of Social and Organizational Dynamics in IT (IJSODIT), 3(1), 14–31. https://doi.org/10.4018/ijsodit.2013010102

Ugbebor, F., Aina, O., Abass, M., & Kushanu, D. (2024). EMPLOYEE CYBERSECURITY AWARENESS TRAINING PROGRAMS CUSTOMIZED FOR SME CONTEXTS TO REDUCE HUMAN-ERROR RELATED SECURITY INCIDENTS. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (Online), 3(3), 382–409. https://doi.org/10.60087/jklst.vol3.n3.p382-409

Weber, K., Schütz, A. E., Fertig, T., & Müller, N. H. (2020). Exploiting the Human Factor: Social Engineering Attacks on Cryptocurrency Users. In P. Zaphiris & A. Ioannou (Hrsg.), Learning and Collaboration Technologies. Human and Technology Ecosystems (S. 650–668). Springer International Publishing. https://doi.org/10.1007/978-3-030-50506-6_45

Zion Market Research. (2024). Cybersecurity Awareness Training Market Size, Share, Analysis, Trends, Growth, Forecasts, 2032. Zion Market Research. https://www.zionmarketresearch.com/report/cybersecurity-awareness-training-market?utm.com

Appendix A: Preferences for ISA Training Methods Instrument

In the following, different training methods in the field of information security are presented. Please indicate for each method how much you would personally prefer to participate in this type of training.

Please try to imagine that you are required to complete such a training as part of your studies or work.

Please evaluate each training method using the following scale:

1 = Not preferred at all
2 = Rather not preferred
3 = Neutral
4 = Rather preferred
5 = Strongly preferred

Table 3: Preferences for ISA Training Methods Instrument

	Description
Security competitions	You participate in a competition where you solve tasks related to IT security, such as identifying vulnerabilities or detecting attacks. You compete with other participants and compare your results.
Quizzes	You answer questions about IT security, for example in the form of short tests or multiple-choice questions. You receive immediate feedback on whether your answers are correct or incorrect.
Gamified learning	You learn IT security content in a game-like environment, for example by collecting points, unlocking levels, or completing small challenges within a learning platform.
Web seminars	You participate in an online session where IT security topics are explained. You mainly listen but may also ask questions or interact via chat.
Expert lectures	You listen to an expert presenting IT security topics, such as current threats or common mistakes in handling data.
Awareness-themed days	You participate in an event dedicated to IT security, which may include different formats such as presentations, short exercises, or informational sessions throughout the day.
Simulations	You are placed in a realistic situation, for example receiving a suspicious email, and must decide how to respond.
Hands-on learning	You actively work on tasks yourself, such as identifying phishing emails, configuring security settings, or applying specific security measures.
Role plays	You act out a situation together with others, for example a conversation with someone trying to obtain confidential information.
Case studies	You analyze a real or realistic security incident and reflect on what happened and how it could have been handled.
E-learning platforms	You use an online platform with structured IT security courses that you can complete independently.
Podcasts	You listen to audio content about IT security, such as interviews or discussions on relevant topics.
Mobile learning applications	You use a mobile app to learn IT security content in short, flexible learning units.
Videos	You watch videos that explain or demonstrate IT security topics.
Documentation	You read written materials such as guidelines, manuals, or wiki pages related to IT security.
Group work	You work together with others on tasks or topics related to IT security and exchange ideas within a team.
Discussion rounds	You participate in discussions where experiences and opinions about IT security are shared.
Mentoring programs	You are guided by an experienced person who provides individual support and shares knowledge about IT security.

Analysis of the Impact of Advanced Technologies on Business Processes

Eva Krhač Andrašec, Tilen Medved, Tomaž Kern
University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
eva.krhac1@um.si, tilen.medved2@um.si, tomaz.kern@um.si

This study examines how advanced manufacturing technologies influence the performance of business processes through a Business Process Management approach. A basic product manufacturing process was selected as a case study and modelled in its current (AS-IS) and redesigned (TO-BE) states using the EPC method in the ARIS tool. The redesigned process integrates 3D scanning, CAD modelling, additive manufacturing, and collaborative robotic assembly while removing manual fabrication, paper documentation, and repetitive handling activities. Both process models were compared using performance indicators, including the number of activities, documents, and the share of IT-supported tasks. A comparative analysis shows that technology integration allows mass customization and increases digital support and automation. The findings suggest that although advanced technologies enable long-term process standardization and data transparency, organizations may experience short-term structural complexity and implementation challenges. The study contributes empirical insight into how advanced technologies reshape process structure and performance measurement.

DOI
https://doi.org/10.18690/um.fov.4.2026.42

ISBN
978-961-299-152-4

Keywords:
business process management, advanced technologies, process modelling, digital transformation, manufacturing process

University of Maribor Press

1 Introduction

Modern organizations face unprecedented pressure to improve process efficiency in response to dynamic market demands and shorter product lifecycles. Industry 4.0 and 5.0 paradigms emphasize aligning production processes with advanced technologies to meet customer needs quickly and flexibly (Sachdeva et al., 2023). This article explores the impact of these technologies on key process performance indicators. In the following, we first define business process management and key performance indicators. This is followed by a chapter that discusses various modern technologies and their impact on business process execution.

However, existing studies often discuss advanced technologies such as additive manufacturing, 3D scanning, CAD/CAM, and collaborative robots separately, with less attention given to how their combined implementation changes the structure and measurable indicators of a concrete business process. In particular, there is limited process-level evidence showing how such technologies affect process models, documentation flows, role allocation, and IT-supported activities. This study addresses this gap by comparing a real badge manufacturing process before and after the integration of selected advanced technologies.

1.1 Business process management

Business Process Management (BPM) is a structured, holistic approach that enhances organizational performance by systematically managing and continuously improving business processes (Lahajnar & Rožanec, 2015). It encompasses the identification, modelling, execution, monitoring, and optimization of processes, aligning operational activities with strategic objectives (Dumas et al., 2013). The BPM life cycle is typically divided into six core phases: preparation for improvement, process mapping, analysis, improvement, solution implementation, and process monitoring (Dumas et al., 2018). Among these, the most critical phases for achieving sustainable process enhancement are analysis, improvement, and monitoring.

The analysis phase forms the cornerstone of informed decision-making in BPM. It entails a thorough examination of existing processes to identify inefficiencies, redundancies, bottlenecks, and deviations from expected performance levels. Analytical activities are typically divided into two key dimensions: operational and

structural analysis of process indicators (Krhač Andrašec, 2022). Operational analysis focuses on dynamic performance indicators, including time, cost, quality, and flexibility (Franz et al., 2012; van der Aalst, 2013; Urh et al., 2018). Structural analysis, on the other hand, examines the underlying architecture of processes, including task sequencing, role allocation, communication flows, and interdependencies among process elements (Sarnikar & Deokar, 2017). Both dimensions are essential for uncovering the root causes of performance issues and developing data-driven strategies for process improvement.

The improvement phase aims to modify and optimize business processes based on insights gained during analysis. A wide range of measures, methods, and techniques (Krhač Andrašec, 2022) may be employed, depending on the specific context and identified performance gaps. Common approaches include Lean Management (Sari et al., 2017), Six Sigma (Tosuner et al., 2016), Total Quality Management (Rusev & Salonitis, 2016), and Business Process Reengineering (Botha et al., 2012), each offering distinct methods and techniques for process optimization. Methods and techniques such as root cause analysis (Pinney et al., 2015), simulation (Dumas et al., 2018), and benchmarking (Boutros & Cardella, 2017) further support the systematic development and evaluation of improvement scenarios. Interventions may target changes in workflows, information systems, or organizational roles.

The monitoring phase involves continuously tracking key performance indicators (KPIs) to assess the effectiveness of implemented improvements and ensure long-term alignment with organizational objectives (Dumas et al., 2018). This phase not only verifies the success of transformation initiatives but also facilitates early detection of emerging issues, thereby sustaining a cycle of continuous improvement. Feedback obtained during monitoring is typically reintegrated into the BPM life cycle, promoting adaptability and responsiveness in dynamic business environments.

In summary, BPM provides a comprehensive framework for managing and improving organizational processes. The integration of operational and structural analyses, targeted improvement methods and techniques, and rigorous performance monitoring enables organizations to achieve higher levels of efficiency, quality, and agility.

1.2 Advanced technologies and their impacts

In the context of Industry 4.0 and 5.0, several advanced technologies (especially collaborative robots, 3D scanning, and additive manufacturing) are changing how business processes are executed. We examine how these technologies are impacting different key process performance indicators. These technologies also serve as key enablers that make new process designs and performance improvements possible.

Collaborative robots (cobots) are designed to work safely alongside humans and support repetitive, strenuous, or precise tasks. By automating repetitive assembly work, cobots can increase production rates and allow workers to focus on more complex activities (Borboni et al., 2023). Wang et al. (2019) also report productivity gains from cobot integration in a real assembly process, while human-robot collaboration can improve efficiency, quality, and safety (Gualtieri et al., 2020). Cobots also provide ergonomic benefits by reducing exposure to unfavourable postures and repetitive strain (Realyvásquez-Vargas et al., 2019). In addition, they improve consistency, reduce human error, and increase flexibility, as they can be reprogrammed for new tasks more easily than conventional stationary automation. This makes small-batch and customized production more cost-efficient by reducing the need for retooling or specialized equipment (Ford & Despeisse, 2016; Vranić, 2017).

Three-dimensional scanning improves process efficiency, especially in design, prototyping, and quality control. It captures physical objects as high-resolution digital data, which can then support CAD modelling and reduce the need for manual measurement. Modern 3D scanning can speed up product development in design, testing, and production, while reducing development time, costs, and design errors (Haleem et al., 2022; Javaid et al., 2021). In some cases, accurate scan data has reduced design cycle costs by up to 75% compared with traditional methods (Muminović et al., 2024). 3D scanning also supports quality control by enabling fast comparison of manufactured parts with CAD specifications, helping to detect defects or deviations more quickly than traditional inspection methods (Yao, 2005). This can reduce reject rates and rework, improve cost efficiency, and support consistent quality (Klemenčić & Skala, 2013), while also enabling greater flexibility in engineering changes and mass customization.

Additive manufacturing (AM), commonly referred to as 3D printing, streamlines prototyping and production by building products layer by layer from digital models. This reduces the need for specialized tools and enables greater design freedom. AM can shorten development cycles and reduce manufacturing time, leading to faster time-to-market for new designs (Attaran, 2017). It also improves cost efficiency because no hard tooling or moulding is required, making small production runs and customized parts more economically viable. In addition, AM uses only the material needed for production, which reduces waste, costs, and environmental impact compared with traditional cutting processes (Duraković, 2018). Companies can also print spare parts or customized products on demand, reducing inventory needs and the risk of overproduction (Shivananda Devi & Nilanjana, 2022).

Computer-aided design and manufacturing (CAD/CAM) systems support modern product and process development by linking digital design with production. The introduction of 3D CAD modelling can shorten development time from concept to production while also reducing costs (Vranić et al., 2017), as potential problems can be identified virtually and the need for physical mock-ups is reduced. CAD-based rapid prototyping also enables complex product geometries with minimal additional costs or delays (Jorgić et al., 2025). Once the design is finalized, CAM software can optimize manufacturing by creating efficient toolpaths, reducing machining time and scrap, and supporting consistent quality through automation (Iwata, 2003). CAD/CAM integration also embeds specifications and tolerances into manufacturing instructions, reducing interpretation errors and improving consistency across production runs (Sivakumar & Dhanalakshmi, 2013; Sun, Xia, & Mao, 2010). Finally, customizable CAD models support product customization and quick engineering changes (Shin & Kwak, 2003), enabling mass customization when combined with additive manufacturing or high-speed CNC (Shui-liang, 2007; Barros, Duarte, & Chaparro, 2014).

The reviewed literature shows that each technology can improve specific aspects of process performance, such as speed, quality, flexibility, ergonomics, or cost efficiency. Nevertheless, most studies focus on individual technologies or general implementation benefits. Fewer studies examine how several advanced technologies interact within one redesigned business process and how this affects structural process indicators. This paper focuses on the combined process impact of 3D scanning, CAD modelling, additive manufacturing, and cobot-assisted assembly.

2 Methods

This study analyses the impact of advanced manufacturing technologies on business process performance indicators through process modelling and comparison. We selected a badge manufacturing process as an example case study. A single-case design was selected because it enables a detailed comparison of process structure, roles, documents, and system support before and after redesign. The analysed process represents a real small-batch badge production workflow used as a case study for modelling and redesign. This current process (AS-IS) was modelled using the EPC method in the ARIS tool based on documented workflow steps, roles, and information flows. The model captured all activities, involved employee roles, documents, and system support.

Next, a redesigned process (TO-BE) was developed by introducing selected Industry 4.0 and 5.0 technologies, namely 3D scanning, CAD modelling, AM, and cobot assembly. Manual fabrication steps, paper documentation, and repetitive handling activities were mostly removed or automated. Both process models were analysed using process indicators defined by Krhač Andrašec (2022).

Finally, a SWOT analysis was performed to evaluate the TO-BE model and assess the overall impact of technology implementation. The study addresses the following research question: "How does the integration of advanced manufacturing technologies affect performance indicators of a production process?".

The contribution of the study is twofold. First, it provides a process-level comparison of the same manufacturing workflow in its AS-IS and TO-BE states, rather than discussing technology impacts only at a general organizational level. Second, it shows how the integration of multiple technologies changes structural indicators such as the number of activities, documents, and IT-supported tasks. In this way, the study links the theoretical discussion of Industry 4.0/5.0 technologies with concrete BPM-based process redesign evidence.

3 Results

3.1 Current badge-making process

The current (AS-IS) process of badge manufacturing begins when, upon receiving a customer order, the availability of products in stock is checked, the data is entered, and planning changes are monitored. This is followed by the calculation of material requirements, after which it is determined that either:

- the sticker with the logo is not available,
- the sticker carrier is not available,
- the materials or components are available.

In the first case, material requirements for the first semi-finished product are calculated. If the production material is unavailable, the process proceeds with procurement; if it is available, the material issue process follows. After the material and components are issued from the warehouse, the 2D printer operator prints the labels, and the warehouse worker and production worker transfer the components to stock.

In the second case, material requirements for the second semi-finished product are calculated. Depending on material availability, the process continues with procurement or material issue. After the material and components are issued from the warehouse, the production worker cuts and shapes the base material, performs surface finishing and cleaning, and glues the clip to the carrier. The warehouse worker and production worker then transfer the components to stock.

Once all materials and components are available, capacity planning follows. If capacity is unavailable, the process is repeated. Otherwise, the planner prepares the working documentation, and the material issue process is carried out. The warehouse worker prints the product labels, after which the production worker manually affixes the labels to the carriers and packages the products in bags. Finally, product transfer to the warehouse and production reporting are carried out simultaneously. The process ends when the products are in stock and ready for shipment.

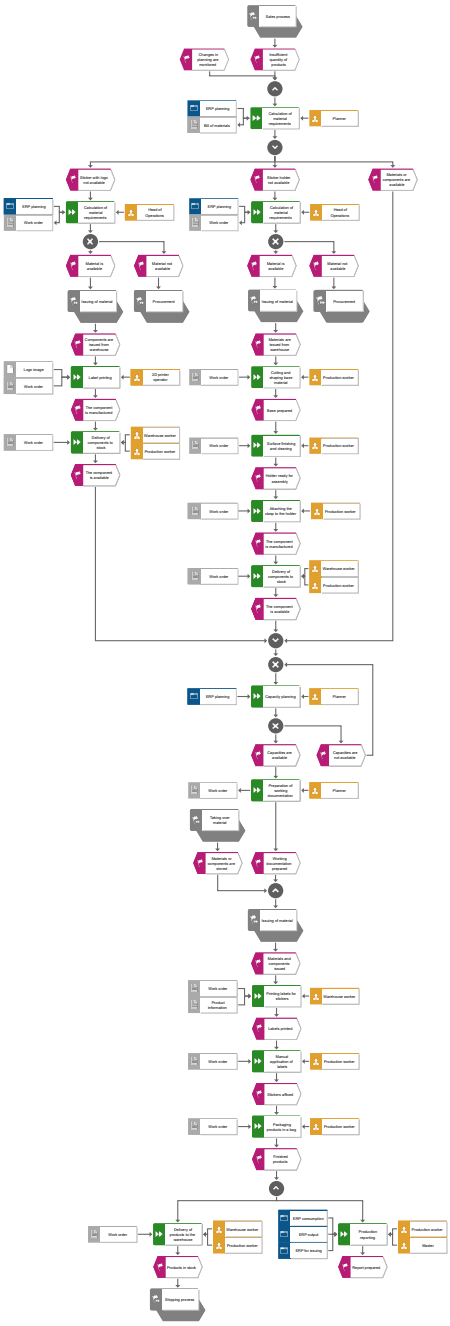

Figure 1: Process model in initial (AS-IS) state

Source: Own

The ARIS tool was used to prepare the current state of the selected business process. The model of the selected process execution is shown in Figure 1.

3.2 Redesigned badge-making process

In the TO-BE process, to introduce mass customization, the production flow is redesigned to integrate digital technologies and eliminate manual, low-value activities. When the product is not in stock, the new process introduces 3D scanning of customer samples and 3D CAD modelling. These activities run in parallel with a simplified material availability check to ensure that printing materials are ready for use. The 3D scan and CAD model together form a digital job packet that includes the CAD file, 3D printer settings, and label design. This replaces the paper-based drawings and work instructions used in the AS-IS process and provides a single, standardized input for production.

In the manufacturing phase, manual activities such as cutting, shaping, and surface finishing are removed entirely. The 3D printer now fabricates the badge holder or semi-finished component directly from the digital model. Because of on-demand AM, there is no longer a need to plan or issue semi-finished holders from stock. Only the filament or resin used for printing must be monitored and replenished. This reduces waiting times and eliminates delays caused by workshop availability or external suppliers. The result is a shorter and more flexible production cycle that can efficiently handle both standard and custom orders.

During assembly and packaging, a cobot replaces several manual, repetitive tasks. The cobot applies the label with vision guidance, performs an automatic visual inspection to verify label position and quality, and helps attach the clasp to the holder and package the finished component. These changes enhance process precision, minimize human error, and substantially reduce the physical strain on workers. The cobot cell ensures a more stable and predictable cycle time while freeing operators to focus on supervision and quality monitoring.

Finally, the information flow is fully digitalized. The paperwork, orders, and manual confirmations in the AS-IS process are replaced by automatic data exchange with the ERP. The system now records printing status, inspection results, and completion

confirmations directly from the cobot cell. This enables real-time visibility of work in progress and quality data, removing late or missing postings.

The changes are marked with red circles in Figure 2 below.

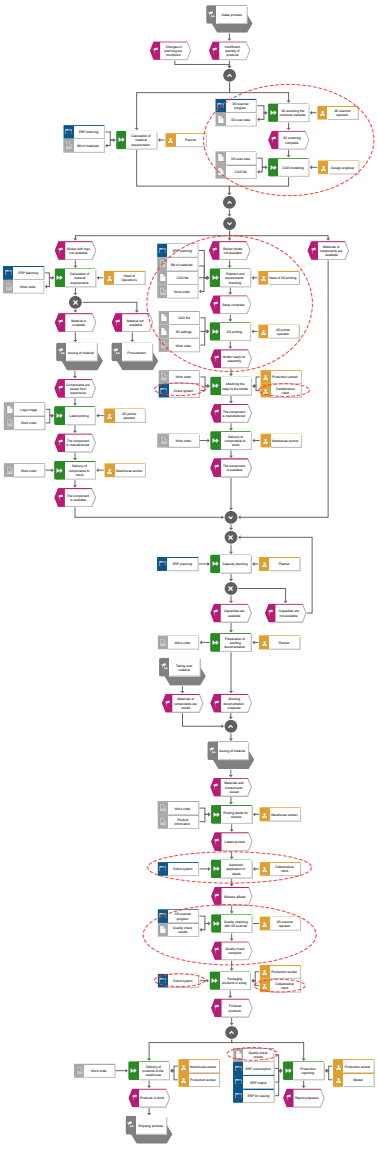

Figure 2: Process model in redesigned (TO-BE) state

Source: Own

4 Discussion

4.1 Analysis and comparison of process models

The TO-BE process introduces 3D scanning, CAD modelling, 3D printing, and robotic assistance, while removing manual fabrication, paper documentation, and repetitive handling. One of the main benefits is mass customization, which combines the high-volume efficiency of mass production with the individualized tailoring of custom-made products (Meglič & Jeraj, 2004). The expected results include shorter lead times, reduced rework, higher first-pass yields, improved ergonomics, lower inventory levels, and simplified planning that focuses solely on printing materials (Attaran, 2017; Realyvásquez-Vargas et al., 2019; Shivananda Devi & Nilanjana, 2022).

In accordance with the established research purpose, an analysis of the processes in both their current and redesigned states has been conducted. The analysis encompasses the fundamental indicators identified by Krhač Andrašec (2022). The number of activities increases (from 16 to 18) due to introduction of machine vision quality control integrated with the help of cobot, which increases the quality of products. The number of documents rises, but they are mainly replaced by electronic forms that are now stored within the system. At the same time, the share of IT-supported activities grows from 31.25% to 55.56%, reflecting the integration of digital modelling, AM and robotic assistance.

The comparison shows that technology integration initially increases structural complexity despite higher automation levels. Therefore, it is possible to conclude that the process after the implementation of advanced technologies initially deteriorates primarily in terms of time (e.g. due to adaptation to a new method of execution) and cost (e.g. due to a significant initial investment). Given the initially poorer understanding of the process and of advanced technologies, it is expected that at least a temporary deterioration in process execution quality will occur. Overall, despite short-term structural challenges, the redesigned process is expected to support long-term improvements through more IT-supported activities, greater use of electronic documentation, more stable execution, and a gradual shift of employees from repetitive manual work toward supervision, quality monitoring, and exception handling.

4.2 SWOT analysis of the integrated digital production workflow

The SWOT analysis in Figure 3 summarizes the main operational factors linked to the integrated 3D scan–CAD–AM–cobot workflow. It highlights key internal strengths (Bhuiyan, 2011; Liu et al., 2022) and weaknesses (Quanjin et al., 2020; Jonuzakov, 2023) together with external opportunities (Gruska and Cherry, 2005; Walter and Marcham, 2020) and threats that may affect implementation, performance, and long-term sustainability.

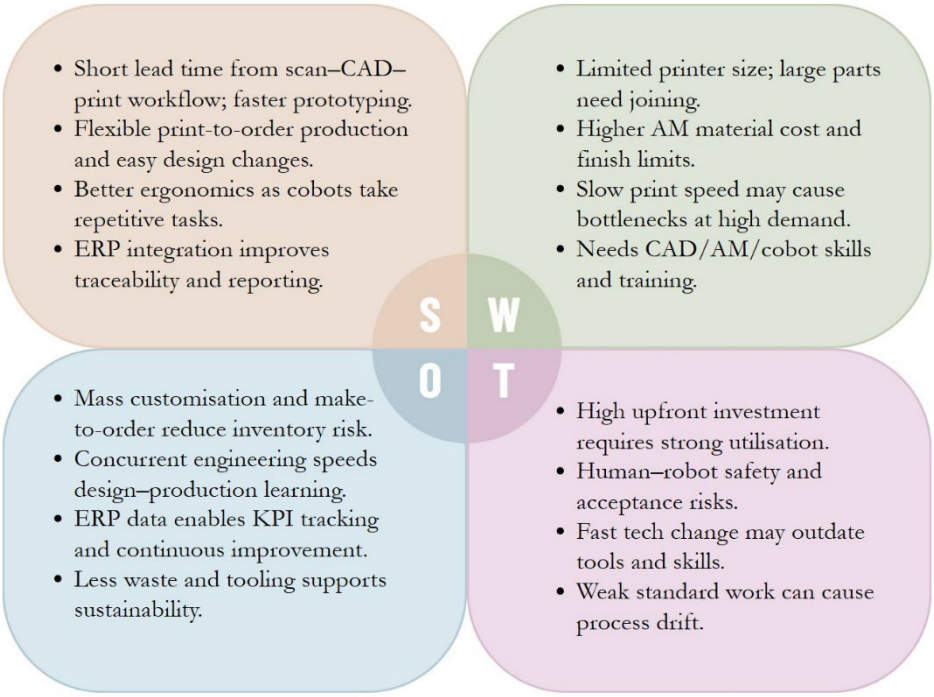

Figure 3: SWOT analysis of the redesigned process
Source: Adapted from Jorgić et al. (2025)

Overall, the redesigned process replaces manual and paper-based steps with a digital workflow integrating 3D scanning, CAD modelling, AM, and robotic assistance linked to the ERP system. The comparison shows higher digital integration, automated handling, and consolidated information flows through a unified digital job packet. These effects align with established findings that AM shortens development cycles and reduces setup effort (Attaran, 2017; Vranić et al., 2017;

Wong & Hernandez, 2012), CAD-based workflows reduce design errors and iteration time, 3D scanning supports faster modelling and inspection (Ebrahim, 2015; Haleem et al., 2022), and cobots stabilize execution while reducing manual strain (Liu et al., 2022; Pearce et al., 2018; Villani et al., 2018). Consistent with product-development research, the results suggest that digital technologies improve process consistency and flexibility, although implementation may initially increase structural complexity and require organizational adaptation (Agrawal & Bhuiyan, 2014; Guzik, 2023).

Despite these benefits, several practical limits must be considered. Printer size, material constraints, and high production volumes may limit the suitability of the redesigned process, while investments in equipment, training, and safety require careful planning. These factors should be included in capacity planning, workforce preparation, and operational control.

This study also has limitations. First, it is based on a single case process, which limits the generalizability of the findings. Second, the TO-BE process is a redesigned model rather than a fully implemented and measured process. Therefore, the expected effects on time, costs, quality, and learning curves should be validated in practice. Third, the analysis does not include detailed investment costs, training costs, equipment utilization, or organizational readiness, all of which may affect whether the redesigned process delivers the expected benefits.

To support long-term improvement, organizations should monitor key performance indicators such as lead time, first-pass yield, and labour time per unit. Regular KPI review links process analysis, improvement, and monitoring. Overall, Industry 4.0 and 5.0 technologies can stabilize execution, improve transparency, and reduce manual strain, while supporting future optimization and scaling. The main novelty of this study is not the analysis of one technology, but the BPM-based demonstration of how several technologies jointly reshape the structure, documentation, automation level, and control logic of a concrete manufacturing process. Future work could extend the EPC-based information flow model with material flow analysis to better assess inventory movement, resource utilization, and the efficiency of mass customization processes.

References

Agrawal, A., & Bhuiyan, N. (2014). Achieving success in NPD projects. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, *8*(2), 476-481.

Attaran, M. (2017). The rise of 3-D printing: The advantages of additive manufacturing over traditional manufacturing. *Business horizons*, *60*(5), 677-688.

Barros, M., Duarte, J. P., & Chaparro, B. M. (2014). Integrated generative design tools for the mass customization of furniture. In *Design Computing and Cognition'12* (pp. 285-300). Dordrecht: Springer Netherlands.

Bhuiyan, N. (2011). A framework for successful new product development. *Journal of Industrial Engineering and Management (JIEM)*, *4*(4), 746-770.

Borboni, A., Reddy, K. V. V., Elamvazuthi, I., Al-Quraishi, M. S., Natarajan, E., & Azhar Ali, S. S. (2023). The expanding role of artificial intelligence in collaborative robots for industrial applications: A systematic review of recent works. *Machines*, *11*(1), 111.

Botha, G. J., Kruger, P. S., & De Vries, M. (2012). Enhancing customer experience through business process improvement: an application of the Enhanced Customer Experience Framework (ECEF). *South African Journal of Industrial Engineering*, *23*(1), 39-56.

Boutros, T., & Cardella, J. (2017). *The basics of process improvement*. CRC Press.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). *Fundamentals of business process management* (Vol. 1, p. 2). Heidelberg: Springer.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). *Fundamentals of business process management* (2nd ed.). Heidelberg: Springer.

Duraković, B. (2018). *Design for additive manufacturing: benefits, trends and challenges. Period Eng Nat Sci 6: 179–191.*

Ebrahim, M. A. B. (2015). 3D laser scanners' techniques overview. *Int J Sci Res*, *4*(10), 323-331.

Ford, S., & Despeisse, M. (2016). Additive manufacturing and sustainability: an exploratory study of the advantages and challenges. *Journal of cleaner Production*, *137*, 1573-1587.

Franz, P., & Kirchmer, M. (2012). *Value-driven business process management: The value-switch for lasting competitive advantage*. New York: McGraw-Hill.

Gruska, G. F., & Cherry, D. (2005). APQP: Not just for document creation. *Quality*, *44*(2), 32.

Gualtieri, L., Palomba, I., Merati, F. A., Rauch, E., & Vidoni, R. (2020). Design of human-centered collaborative assembly workstations for the improvement of operators' physical ergonomics and production efficiency: A case study. *Sustainability*, *12*(9), 3606.

Guzik, A. (2023). Overview of new product development strategies and models. *Catallaxy*, *8*(1), 21-34.

Haleem, A., Javaid, M., Singh, R. P., Rab, S., Suman, R., Kumar, L., & Khan, I. H. (2022). Exploring the potential of 3D scanning in Industry 4.0: An overview. *International Journal of Cognitive Computing in Engineering*, *3*, 161-171.

Iwata, K. (2003). Computer-aided manufacturing (CAM). In J. Webster (Ed.), *Wiley encyclopedia of electrical and electronics engineering* (pp. 1–7). New York: Wiley.

Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2021). Industrial perspectives of 3D scanning: features, roles and it's analytical applications. *Sensors International*, *2*, 100114.

Jonuzakov, S. (2023). *Economic aspects of additive manufacturing* (Doctoral dissertation, Politecnico di Torino).

Jorgić, D., Medved, T., Urh, B. (2025). *Uporaba sodobnih tehnologij v procesu razvoja proizvoda/proizvodnega procesa*. University of Maribor, Faculty of Organizational Sciences, Maribor.

Klemenčić, M., & Skala, K. (2013, May). 3D visual analytics for quality control in engineering. In *2013 36th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 258-262). IEEE.

Krhač Andrašec, E. (2022). Business process improvement methods and techniques and their impact on the efficiency of organizational systems. Doctoral dissertation, University of Maribor

Lahajnar, S., & Rožanec, A. (2015). Primerjava metodologij za menedžment poslovnih procesov. *Uporabna informatika*, *23*(4), 226-238.

Liu, L., Guo, F., Zou, Z., & Duffy, V. G. (2022). Application, development and future opportunities of collaborative robots in manufacturing: A literature review. *International Journal of Human–Computer Interaction, 38*(18), 1801–1824.

Meglič, J., & Jeraj, M. (2004). Information support of product development process for mass customization. V Informatics and management: selected topics (str. 169–187). P. Lang.

Muminović, A. J., Gierz, Ł., Rebihić, H., Smajić, J., Pervan, N., Hadžiabdić, V., ... & Sydor, M. (2024). Enhancing furniture manufacturing with 3d scanning. *Applied Sciences*, *14*(10), 4112.

Pearce, M., Mutlu, B., Shah, J., & Radwin, R. (2018). Optimizing makespan and ergonomics in integrating collaborative robots into manufacturing processes. *IEEE transactions on automation science and engineering*, *15*(4), 1772-1784.

Pinney, S. J., Page, A. E., Jevsevar, D. S., & Bozic, K. J. (2015). Current concept review: quality and process improvement in orthopedics. *Orthopedic research and reviews*, 1-11.

Quanjin, M., Rejab, M. R. M., Idris, M. S., Kumar, N. M., Abdullah, M. H., & Reddy, G. R. (2020). Recent 3D and 4D intelligent printing technologies: A comparative review and future perspective. *Procedia Computer Science*, *167*, 1210-1219.

Realyvásquez-Vargas, A., Arredondo-Soto, K. C., García-Alcaraz, J. L., Márquez-Lobato, B. Y., & Cruz-García, J. (2019). Introduction and configuration of a collaborative robot in an assembly task as a means to decrease occupational risks and increase efficiency in a manufacturing company. *Robotics and Computer-Integrated Manufacturing*, *57*, 315-328.

Rusev, S. J., & Salonitis, K. (2016). Operational excellence assessment framework for manufacturing companies. *Procedia CIRP*, *55*, 272-277.

Sachdeva, A., Agrawal, R., Chaudhary, C., Siddhpuria, D., Kashyap, D., & Timung, S. (2023). Sustainability of 3D printing in industry 4.0: A brief review. *3D Printing Technology for Water Treatment Applications*, 229-251.

Sari, N., Rotter, T., Goodridge, D., Harrison, L., & Kinsman, L. (2017). An economic analysis of a system wide Lean approach: cost estimations for the implementation of Lean in the Saskatchewan healthcare system for 2012–2014. *BMC health services research*, *17*(1), 523.

Sarnikar, S., & Deokar, A. V. (2017). A design approach for process-based knowledge management systems. *Journal of Knowledge Management*, *21*(4), 693-717.

Shin, J. H., & Kwak, B. M. (2003). Customizing CAD system and its Application for Engineering Design. In *Proceedings of the KSME Conference* (pp. 1187-1192). The Korean Society of Mechanical Engineers.

Shivananda Devi, K., & Nilanjana, R. (2022). *Additive manufacturing*. Laxmi Publications.

Shui-liang, F. (2007). Computer aided manufacturing system oriented to mass customization. Computer Integrated Manufacturing Systems. 13(7), 1317–1322.

Sivakumar, S., & Dhanalakshmi, V. (2013). An approach towards the integration of CAD/CAM/CAI through STEP file using feature extraction for cylindrical parts. *International Journal of Computer Integrated Manufacturing*, *26*(6), 561–570.

Sun, Y., Xia, X., & Mao, Y. (2010). *Modeling based on structural characteristics integrated with product engineering quality information*. In *Proceedings of the Third International Workshop on Advanced Computational Intelligence (IWACI 2010)*. IEEE.

Tosuner, Z., Gücin, Z., Kiran, T., Büyükpinarbaşili, N., Turna, S., Taşkiran, O., & Arici, D. S. (2016). A Six Sigma Trial For Reduction of Error Rates in Pathology Laboratory. A Six Sigma Trial For Reduction of Error Rates in Pathology Laboratory. *Turk patoloji dergisi*, *32*(3), 171–177.

Urh, B., Krhač, E., Roblek, M., & Kern, T. (2018). Ocena učinkovitosti prenove procesa na podlagi strukture procesa. Uporabna informatika, 26(1), 12-12.

van der Aalst, W. M. P. (2013). *Business process management: A comprehensive survey*. *International Scholarly Research Notices*, 2013, Article 507984.

Villani, V., Pini, F., Leali, F., & Secchi, C. (2018). *Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications*. *Mechatronics, 55*, 248–266.

Vranić, A., Bogojević, N., Ćirić-Kostić, S., Croccolo, D., & Olmi, G. (2017). *Advantages and drawbacks of additive manufacturing*. *IMK-14 – Istraživanje i razvoj, 23*(2), 57–62.

Walter, A., & Marcham, C. L. (2020). *Environmental advantages in additive manufacturing. Professional Safety, 65*(1), 34–38.

Wang, K. B., Dailami, F., & Matthews, J. (2019). *Towards collaborative robotic polishing of mould and die sets. Procedia Manufacturing, 38*, 1499–1507.

Wong, K. V., & Hernandez, A. (2012). *A review of additive manufacturing. International Scholarly Research Notices*, 2012, Article 208760.

Yao, A. (2005). Applications of 3D scanning and reverse engineering techniques for quality control of quick response products. *The International Journal of Advanced Manufacturing Technology, 26*, 1284–1288.

Efficient Explainable and Evidence Based Precision and Personalisation of Treatment

Nilmini Wickramasinghe,[1] Nalika Ulapane,[1] Anamika Ranaut,[1] Neale Cohen[2]

[1] La Trobe University, School of Computing, Engineering and Mathematical Sciences, Melbourne, Australia
n.wickramasinghe@latrobe.edu.au, n.ulapane@latrobe.edu.au, anamika.rannout03@gmail.com
[2] Baker Heart and Diabetes Institute, Melbourne, Australia
neale.cohen@baker.edu.au

Modern advances in computation enable the use of complex machine learning algorithms and artificial intelligence to assist human decision-making. However, the lack of explainability entailing from the black box nature of complex machine learning algorithms, inhibit their adoption in real-world applications especially in fields like healthcare. To address this challenge, we explored using the Odds Ratio (OR)—a clinically well-known measure of evidence—coupled with a sorting algorithm to prototype an explainable clinical decision support system (CDSS). This CDSS intakes relevant patient information such as demographic variables, clinical variables, medical history, and so on, and ranks treatment options personalised for patients, based on OR evidence. We present in this work-in-progress paper how our algorithm performs personalised ranking of therapies, taking Type-2 diabetes as a case study. As future work, we endeavour to codesign this further with clinicians to produce a primary care CDSS and assess long-term clinical outcomes.

DOI
https://doi.org/10.18690/um.fov.4.2026.43

ISBN
978-961-299-152-4

Keywords:
artificial intelligence, clinical decision support system, explainable AI, odds ratio, personalisation of treatment, precision medicine, type-2 diabetes

University of Maribor Press

1 Introduction

Computational power has advanced greatly over the recent decades. This has enabled integrating advanced machine learning and deep learning algorithms to assist and augment complex decision-making processes. Along this trajectory, artificial intelligence (AI) has made inroads to assist clinical decision-making as well (Adlung et al., 2021). AI can now assist with various tasks such as pre-clinical data processing, diagnosis, patient stratification, treatment planning, early warning, and more (Adlung et al., 2021). Although AI systems offer such capabilities, many such systems function as "black box models," providing predictions without adequate reasoning or explanation, thereby limiting the interpretability of system outputs (Ennab & Mcheick, 2024). This lack of explainability undermines the trust healthcare professionals can have on AI systems, as it is essential for clinicians to be able to understand and justify support-system outputs when making crucial clinical decisions (Amann et al., 2020). Black box AI systems have therefore faced significant barriers up to date when it comes to implementation in real-world healthcare settings (Frasca et al., 2024). The repeated emergence of the 'explainability' concern emphasises the importance of making AI systems more explainable if they are to be widely adopted in healthcare settings. As an attempt to address this challenge, we attempt to answer the following overarching research question: How might efficient and explainable AI-assisted clinical decision-support systems be developed?

In an attempt to answer the aforesaid research question, we explored making use of the Odds Ratio (OR) (Higgins, Li, & Deeks, 2023)-a clinically well-known and an easily explainable measure of clinical evidence-coupled with a sorting algorithm, to produce an explainable clinical decision support system (CDSS) to assist treatment planning in primary care. This CDSS takes relevant patient information as inputs and provides a personalised ranking of treatment options based on historical clinical evidence. The CDSS intakes variables such as demographic variables, clinical variables, medical history, and so on, and uses OR evidence to rank treatment options personalised for patients.

We present in this work-in-progress paper how personalised ranking of therapies is performed through our algorithm, taking Type-2 diabetes as a case study. Future work will involve codesigning this further with clinicians to produce a CDSS for primary care and assess its usefulness, usability and long-term clinical outcomes.

2 Methodology

2.1 Overarching Methodology

The overarching methodology is based on our previous works on digital twins of patients in healthcare (Wickramasinghe et al., 2025) and is depicted in Figure 1. When a new patient arrives, a digital twin of the patient is constructed using information about the patient. The information considered include the patient's demographic information (like age, sex, and ethnicity), disease information, medical records, patient history, tests and results, and multi-omics (if available). This digital twin is compared with records of historical patients, and a cohort of historical patients that best matches the present patient is selected based certain pre-defined criteria. Evidence of treatment options and health outcomes within the matching cohort are then evaluated and ranked based on Odds ratio with respect to a target health outcome.

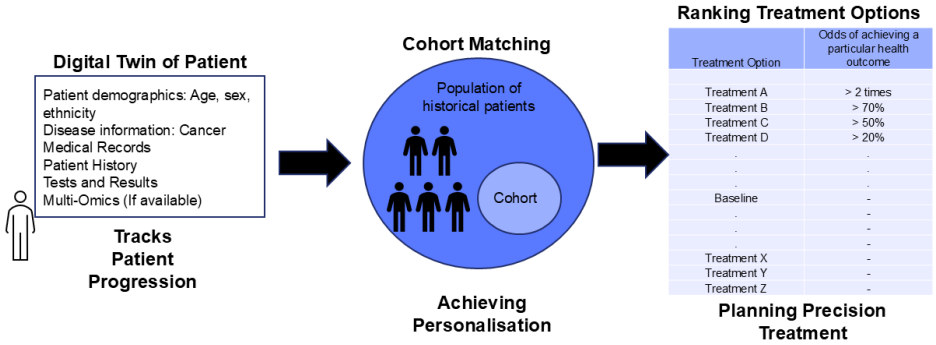

Figure 1: Proposed methodology: Personalised ranking of treatment options based on Odds ratio

Source: Own

2.1 Ranking Treatment Options based on Odds Ratio

The odds ratio (OR) is a commonly used and well-established metric to measure the effectiveness of pharmacological interventions (Higgins, Li, & Deeks, 2023). The way we use OR is depicted in Figure 2. Following the aforementioned approach of cohort matching, we consider the patients within the cohort. We compare all the treatment options against one another and calculate the OR and the Standard Error (SE) as depicted in Figure 1. We then select the statistically significant OR values, i.e., p-value ≤ 0.05, and rank the treatment options from highest odds of success to least odds of success, thereby, presenting clinicians an evidence-based ranking of treatment options personalised to a patient.

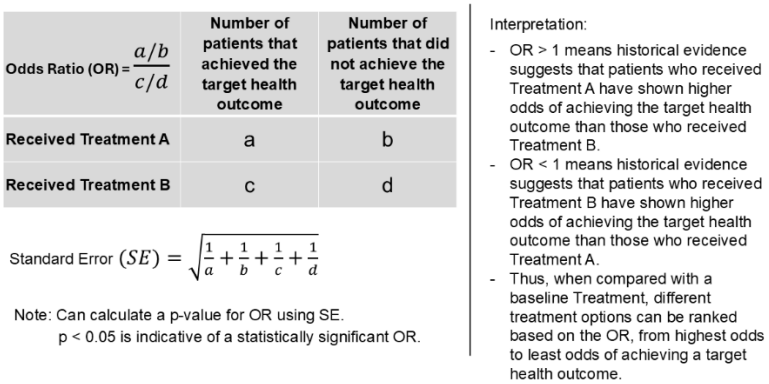

Figure 2: Proposed methodology: Odds ratio-based ranking of treatment options.
Source: Own

2.2 Case Study on Treatment Planning for Type-2 Diabetes

We selected a case of planning Type-2 diabetes therapies using deidentified data from a dedicated Australian diabetes clinic to demonstrate how our approach works. In the diabetes case, when a new patient arrives, historical records of a matching cohort (patient's age $\pm$ 10 years and matching sex) is selected. All available therapies from the cohort are compared with the performance of Metformin. Metformin is a commonly used diabetes therapy, and we consider it the baseline therapy against which other therapies are compared against. The odds ratio for each therapy is calculated (with p-value) against Metformin for the target health outcome of post-

therapy HbA1c ≤ 7%. We then rank the therapies based on statistically significant (i.e., $p \leq 0.05$) odds ratios within their respective categories of first-, second-, and later-line therapies.

3 Results

The ranked therapies resulting from our approach to two patients are depicted in Figures 3 and 4.

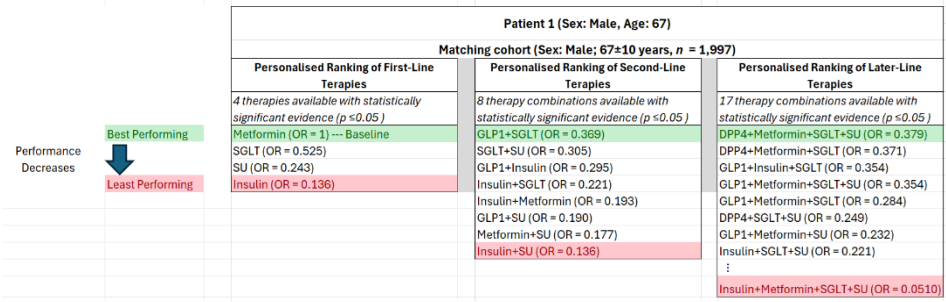

Figure 3: Ranked type-2 diabetes therapies for a male patient
Source: Own

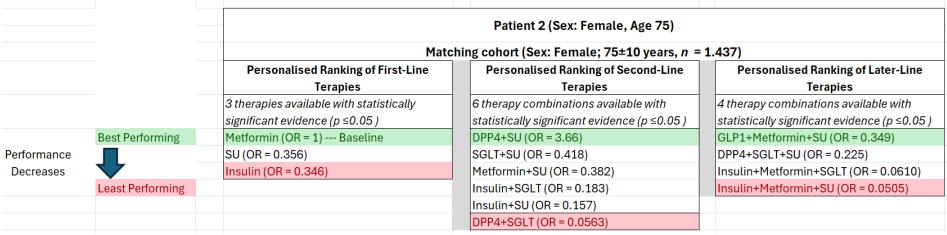

Figure 4: Ranked type-2 diabetes therapies for a female patient
Source: Own

4 Discussion and Conclusions

Answering the following research question was attempted: How might efficient and explainable AI-assisted clinical decision-support systems be developed? A CDSS for primary care was prototyped as an attempt to answer the research question by taking treatment planning for Type-2 diabetes as a case study.

The prototype ensured efficiency as the used algorithm was computationally simple. The system also ensured explainability as it is based on Odds ratio—a measure that is commonly known to clinicians. Since the system performs personalised ranking of therapies based on historical evidence, it enables delivering precise and personalised care through providing evidence-based clinical decision support. Internal assessment of this system performed through discussions held with senior clinicians revealed that a system like this could be useful in scenarios like diabetes treatment planning as there is a large list of available treatment options. It was also suggested that further advancements to enable monitoring factors like longitudinal variation of kidney function and body mass index (BMI) are desired.

As a contribution to theory, this work proposed an efficient and explainable system that can be used for evidence-based personalised treatment planning in primary care settings for diverse health conditions. This system not only can provide clinical decision support but can also serve as a passive tracker of clinical evidence.

As a contribution to practice, Type-2 diabetes was considered as a case study, and it was demonstrated how a large list of therapies can be ranked based on evidence personalised to a patient.

The main limitation of this study is the proposed system being developed only to a prototype stage and lacks clinical validation in terms of qualitative feedback from clinicians about the usefulness and usability of the system, and also as to how the use of a system like this might affect long term patient outcomes.

Future work will involve codesigning the proposed system further with clinicians to produce a CDSS for primary care and assess its usefulness, usability and long-term clinical outcomes.

The proposed system is relatively simple and inexpensive to develop and implement, easy to understand, and can be implemented to assist treatment planning for many health conditions subscribing by the value-based healthcare principles.

Acknowledgment

The support for this work by Optus, La Trobe University and the Baker Heart and Diabetes Institute, Australia, are acknowledged.

References

Adlung, L., Cohen, Y., Mor, U., & Elinav, E. (2021). Machine learning in clinical decision making. Med, 2(6), 642–665. https://www.cell.com/med/fulltext/S2666-6340(21)00155-0.

Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. BMC Medical Informatics and Decision Making, 20, 310. https://doi.org/10.1186/s12911-020-01332-6.

Ennab, M., & Mcheick, H. (2024). Enhancing interpretability and accuracy of AI models in healthcare: A comprehensive review on challenges and future directions. Frontiers in Robotics and AI, 11, Article 1444763. https://doi.org/10.3389/frobt.2024.1444763.

Frasca, M., La Torre, D., Pravettoni, G., & Cutica, I. (2024). Explainable and interpretable artificial intelligence in medicine: A systematic bibliometric review. Discover Artificial Intelligence, 4, 15. https://doi.org/10.1007/s44163-024-00114-7.

Higgins, J. P. T., Li, T., & Deeks, J. J. (Eds.). (2023). Cochrane handbook for systematic reviews of interventions (Version 6.4). Cochrane. https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-06

Wickramasinghe, N., Ulapane, N., & Raghunathan, K. (2025). Digital twins of patients: A cohort matching interpretation. In Studies in Health Technology and Informatics (Vol. 329, pp. 1099–1103). IOS Press. https://doi.org/10.3233/SHTI251009.

When Text Is Not Enough: Structural Limits of Text-Only Transformer-Based Emotion Classification

Szymon Chirowski,[1] Maciej Czerniak,[1]
Ondrej Mitas,[2] Maks Burchard[1]

[1] Breda University of Applied Sciences, Applied Data Science & AI, Breda, the Netherlands
242621@buas.nl, 243552@buas.nl, 240894@buas.nl
[2] Breda University of Applied Sciences, Tourism, Breda, the Netherlands
mitas.o@buas.nl

This study investigates whether limitations observed in text-only transformer-based emotion classification pipelines reflect implementation shortcomings or structural constraints inherent to unimodal modeling. A pipeline was constructed using unscripted dialogue from *MasterChef Polska*, incorporating automated speech-to-text transcription, neural machine translation, and benchmarking across SVM, Bi-LSTM, and RoBERTa architectures. While the fine-tuned RoBERTa model achieved substantially higher accuracy (0.755), confusion matrix analysis and explainable AI techniques revealed persistent structural asymmetries, including uneven performance across emotion categories, high-arousal anger-joy confusion, and translation-induced distortions. Evaluation against automated labels further exposed a "Ground Truth Paradox," where models are validating each other rather than a human-verified set of conclusions. Increased architectural capacity improves performance but does not resolve structural limitations of text-only emotion classification.

DOI
https://doi.org/10.18690/um.fov.4.2026.44

ISBN
978-961-299-152-4

Keywords:
emotion recognition, affective computing, transformer models, multimodal limitations, explainable AI

University of Maribor Press

1 Introduction

The growing integration of Artificial Intelligence systems into digital media environments has intensified interest in automated emotion recognition. Across domains such as marketing analytics, media monitoring, and audience engagement measurement, quantifying affect at scale is often presented as a cost-efficient alternative to manual interpretation. Text-based Natural Language Processing pipelines, particularly those built on transformer architectures such as Robustly Optimized BERT Approach (RoBERTa), are widely adopted due to their scalability and relatively low operational overhead (Schuller et al., 2026; Wu et al., 2025). However, reliance on unimodal, text-only systems raises fundamental questions about theoretical validity. Emotional meaning in real-world interactions rarely emerges from language alone; affective interpretation depends on vocal prosody, facial expression, interpersonal dynamics, and situational context (Picard, 2000; Wu et al., 2025). When classification pipelines rely exclusively on textual input, these multimodal cues are necessarily excluded.

The technical design of emotion recognition systems reflects underlying psychological assumptions. Many contemporary pipelines adopt discrete categorical taxonomies derived from Basic Emotion Theory, proposing universal categories such as joy, sadness, anger, fear, surprise, and disgust (Ekman, 1992; Schuller et al., 2026). These predefined classes align conveniently with supervised machine learning, where labelled examples are mapped to fixed outputs (Azari et al., 2020). From this perspective, emotions are treated as relatively stable patterns inferable from linguistic markers.

In contrast, constructionist accounts conceptualize affect as context-dependent and dynamically assembled rather than biologically fixed (Barrett, 2017). Emotional experience emerges through the integration of language, bodily regulation, prior knowledge, and environmental context. If affect is constructed through multimodal interaction, restricting analysis to textual signals may impose structural constraints on classification accuracy (Wu et al., 2025). This tension between categorical convenience and contextual complexity has direct implications for the reliability of text-only emotion recognition systems.

Rather than treating misclassification as an implementation defect, the analysis examines whether certain limitations stem from structural properties inherent to unimodal text modelling. In doing so, the paper contributes to ongoing discussions about the validity, reliability, and governance of automated affect recognition systems. As regulatory frameworks such as the EU AI Act (Regulation (EU) 2024/1689, 2024) increasingly scrutinize high-risk AI applications, understanding the boundaries of text-only emotion classification becomes both a technical and ethical concern. If incorrect AI-based coding of affect is a substantial risk, it is important to identify these errors and their boundary conditions.

2 Theoretical Background: The Conflict of Affective Frameworks

Automated emotion recognition systems operationalize specific assumptions about the nature of affect. In most computational implementations, emotion classification is framed as a supervised learning task in which observable signals are mapped to predefined categorical labels (Azari et al., 2020; Schuller et al., 2026). This structure reflects the basic emotion framework, which conceptualizes emotions as discrete, biologically grounded states distinguishable through identifiable expressive patterns (Ekman, 1992). Within this paradigm, categories such as anger, fear, joy, sadness, disgust, and surprise function as stable targets for classification.

This categorical alignment is computationally convenient, as fixed emotion labels translate directly into supervised training objectives. However, empirical findings in affective science complicate this assumption. Reviews of physiological and behavioral evidence indicate substantial variability within emotion categories and overlap across categories, challenging the claim that emotions possess distinct biological signatures (Barrett, 2006; Kreibig, 2010; Mauss & Robinson, 2009). Similarly, observational research suggests that prototypical facial configurations are infrequent in natural interaction and highly context-dependent (Barrett, 2011; Russell et al., 2003). Such variability weakens the assumption that discrete categories correspond to stable, universally encoded signals.

In contrast, constructionist accounts conceptualize emotions as dynamically assembled through the integration of bodily states, contextual information, conceptual knowledge, and language (Barrett, 2017). Within this framework, emotion categories function as interpretive constructs rather than biologically

discrete entities. Emotional meaning is therefore context-sensitive and multimodal, emerging from interaction rather than residing solely in linguistic form. Recently, robust empirical evidence has supported constructionist accounts of emotion modelling (Azari et al., 2020).

This theoretical divergence has direct implications for automated emotion recognition. If emotional experience is constructed through multimodal and contextual processes, unimodal systems - particularly those restricted to text - may capture lexical regularities without fully representing affective dynamics (Wu et al., 2025). Classification performance under such constraints may reflect alignment with labelling conventions rather than faithful modelling of emotional states. The tension between categorical convenience and contextual complexity thus forms the conceptual foundation of the present study.

3 Methodology: The Methodological Pipeline

3.1 Dataset and Signal Conversion

The dataset consists of unscripted dialogue extracted from the televised Polish culinary competition *MasterChef Polska*, a high-arousal conversational environment.

The training set consists of 28,395 sentences utilized for training and validation, sourced from a diverse corpus of unscripted television programming - including survival reality shows, competitive tournaments, and spontaneous dialogue - across multiple languages, including but not limited to Dutch, Arabic, Slovak, Polish, and Russian while the target language was English. Despite this multi-program training base, the study maintains a rigorous focus on *MasterChef Polska* as the primary case analysis for evaluation, utilizing a separate test set of 791 sentences (source: Polish; target: English). All instances were analyzed at the sentence level, with each sentence assigned a single dominant emotion label from seven classes: joy, sadness, anger, fear, surprise, disgust, and neutral.

Although the raw material exists as audio-visual recordings, the analytical pipeline intentionally retained only the textual modality to simulate standard text-mining constraints. Visual and acoustic features such as facial expression and prosody were excluded.

3.1.1 Automated Speech-to-Text (STT)

Audio recordings were transcribed using the AssemblyAI Universal 2 model (Khare et al., 2025). Model selection was based on benchmarking against OpenAI Whisper-Large (Radford et al., 2023) using Word Error Rate (WER) (Jurafsky & Martin, 2023). The WER analysis was drawn from a manually verified 15-minute subset of unscripted dialogue to ensure the benchmark reflected the "high-stress" environment of the show. Universal 2 achieved a lower overall WER and provided sentence-level segmentation compatible with downstream emotion classification. Consequently, it was adopted for the final pipeline.

$$WER = \frac{S+I+D}{N} * 100 \quad (1)$$

Equation 1. Word Error Rate (WER), where *S* denotes the total words replaced in transcription by a model, *I* words added by a model that did not match original transcription, *D* total words omitted by a model, *N* number of words in total.

Table 1: Word Error Rate (WER)

Model	Substitutions	Insertions	Deletions	Total Words	WER
Universal 2	88	12	27	1167	10.883%
Whisper-Large	86	25	11	1043	11.697%

3.1.2 Ground Truth Construction

The training set utilized labels inherited from the client's proprietary, third-party LLM-based emotion recognition system. While the specific model architecture remains undisclosed due to commercial sensitivity, the system was confirmed to translate the source Polish text to English to generate the training labels used in this study. The test set (N = 791) was manually annotated by native Polish-speaking students. To ensure consistency, the student team reviewed the manually prepared evaluation subset against the project's specific annotation guidelines provided by the client, cross-referencing utterances with original audio-visual recordings to account for facial expressions and prosodic cues. This methodology results in an imbalanced distribution across the seven emotional classes - joy, sadness, anger, fear, surprise, disgust, and neutral - as detailed in Table 2.

Table 2: Distribution of Emotional Classes Across Datasets

Emotion Class	Training (LLM)	Training (%)	Evaluation (Manual)	Evaluation (%)
Surprise	12,164	42.8	413	52.2
Joy	5,554	19.6	120	15.2
Anger	1,909	6.7	35	4.4
Neutral	3,301	11.6	102	12.9
Sadness	1,867	6.6	85	10.8
Fear	1,800	6.35	27	3.4
Disgust	1,800	6.35	9	1.1
Total	**28,395**	**100**	**791**	**100**

This introduces a labeling asymmetry: the training data reflects automated labeling conventions, whereas the evaluation phase relies on manually verified ground truth. As a result, the classifier is evaluated against a human-verified standard to determine how effectively a model trained on operational AI labels can align with human emotional interpretation. We define this dependency as a "Ground Truth Paradox": while the model is optimized to align with a commercial labeling framework, its ultimate validity is tested against the nuanced benchmarks of native human perception (Bender et al., 2021; Schuller et al., 2026). This limitation is explicitly considered in the interpretation of results.

3.2 Machine Translation

Because the classifier operates in English, Polish transcripts were translated using the MarianMT Polish–English neural machine translation model based on the OPUS-MT framework (Tiedemann & Thottingal, 2020). The pipeline follows a sequential structure:

Audio → Transcription → Translation → Emotion Classification

No multilingual joint training or multimodal fusion was implemented. Translation therefore constitutes an independent transformation stage prior to classification.

3.3 Model Architectures and Benchmarking

To assess the impact of architectural complexity, three model classes were evaluated using identical splits:

- Support Vector Machine (SVM) with linear kernel and term frequency–inverse document frequency (TF-IDF) features (Cortes & Vapnik, 1995; Jurafsky & Martin, 2023),
- Bidirectional LSTM (Hochreiter & Schmidhuber, 1997), the architecture consisted of a 300-dimension embedding layer, two bidirectional LSTM layers with a hidden size of 128, and a dropout rate of 0.3,
- Fine-tuned RoBERTa-base transformer (Liu et al., 2019; Wolf et al., 2020).

All models performed sentence-level single-label classification. Comparative benchmarking allowed assessment of whether performance differences arose from increased contextual modelling capacity. Individual sentences were analyzed as the primary unit of classification to simulate the constraints of real-time media monitoring and to minimize complexity in the multi-speaker environment of a televised cooking competition. While longer contexts might theoretically yield higher classification accuracy by resolving cross-speaker ambiguities, this study focuses on the structural limits of affect recognition from isolated linguistic markers.

3.4 Training Procedure

RoBERTa was fine-tuned using AdamW (Loshchilov & Hutter, 2019) with a class-weighted cross-entropy loss. This approach assigns higher penalties to misclassifications of minority classes, ensuring the model does not become biased toward the majority class. Back-translation augmentation (Sugiyama & Yoshinaga, 2019) was applied to increase exposure to minority classes. Performance was evaluated using accuracy, precision, recall, and F1-score metrics (Jurafsky & Martin, 2023).

3.5 Explainable AI (XAI) Audit

To examine model decision behavior beyond aggregate performance metrics, a post-hoc interpretability method was applied: Gradient × Input (Shrikumar et al., 2017). This method was used to analyze decision patterns in correctly classified and misclassified instances to determine which specific linguistic features drove the model's predictions.

4 Results

The following section presents the quantitative performance metrics of the methodological pipeline, comparing speech-to-text accuracy, baseline model benchmarks, and the specific failure modes identified through Explainable AI (XAI).

4.1 Performance Benchmarking Across Model Generations

The three model classes described in §3.3 were evaluated on the manually annotated human-verified test set (N=791). The SVM baseline, implemented with TF-IDF vectorization and a 50,000-token vocabulary, achieved an overall accuracy of 0.467. The model demonstrated moderate performance on high-frequency and high-arousal categories such as anger and joy, with an average F1-score of 0.547 across these classes. However, it failed to detect any instances of the disgust category during testing. This error indicates a substantial limitation in capturing nuanced affective distinctions using surface-level lexical representations.

The Bidirectional LSTM model, equipped with embedding, pooling, and fully connected layers, yielded an overall accuracy of 0.45. Although precision for the fear category was relatively high, recall remained low at 0.23. This imbalance suggests that while the model successfully identified prototypical expressions of fear, it struggled to generalize more subtle or context-dependent instances present in unscripted dialogue.

The transformer-based RoBERTa model substantially outperformed both baselines, achieving an overall accuracy of 0.755. The performance gain highlights the advantage of contextualized token representations for emotion classification in high-variance conversational text.

It is important to note that the dataset exhibits strong class imbalance, with the surprise category representing more than half of the evaluation instances (413 of 791). This distribution influences overall performance metrics and partially explains the higher weighted F1-score relative to the macro-average.

Table 3: Results Comparison Table (Evaluated on Manually Labeled Test Set, N=791)

Model	Accuracy	Precision	Recall	F1-score
SVM	0.467	0.478	0.343	0.358
Bi-LSTM	0.450	0.570	0.452	0.419
RoBERTa	0.755	0.758	0.755	0.755

4.1.1 The Role of Neural Machine Translation

The move from SVM and Bi-LSTM to RoBERTa improves accuracy to 0.755 on the test set, but the model's performance heavily depends on English-centric pre-training and the quality of translation. Polish transcripts are converted via the MarianMT NMT layer, which can introduce semantic distortions. About 9.7% of classification errors originate in this translation stage, often due to flattened Polish morphological markers or literal translations that misrepresent context. For instance, "Synu" can convey paternal authority or informal camaraderie depending on its suffix, but is usually translated as "Son," losing interpersonal nuance. Similarly, "Granat," meaning pomegranate in MasterChef Polska, is sometimes translated as "Grenade," producing false signals of anger or fear for the classifier. These examples illustrate that even a high-performing model like RoBERTa remains vulnerable to contextual distortions if the translation layer does not preserve cultural and pragmatic cues.

The 9.7 percent translation error was calculated through a manual forensic review of all 194 misclassified instances produced by the RoBERTa model. Two researchers from the student team cross-referenced the model's incorrect predictions with the original Polish audio and the translated English transcripts. An error was attributed to the translation layer if the English text contained a specific semantic distortion - such as the literal rendering of the culinary term "Granat" (pomegranate) as "Grenade" - that logically misled the classifier into a high-intensity emotional trigger like anger or fear.

4.2 Granular Failure Modes

Although the RoBERTa architecture substantially outperformed the SVM and Bi-LSTM baselines, confusion matrix analysis reveals uneven performance across emotional categories. Performance varied significantly depending on class frequency within the dataset.

The disgust category represents the most extreme example of class imbalance, with only nine instances in the evaluation set. Despite this limitation, the model achieved a recall of 0.667 and an F1-score of 0.632, indicating that lexical markers associated with disgust were detected in several cases. However, the small sample size limits the reliability of performance estimates and increases sensitivity to misclassification.

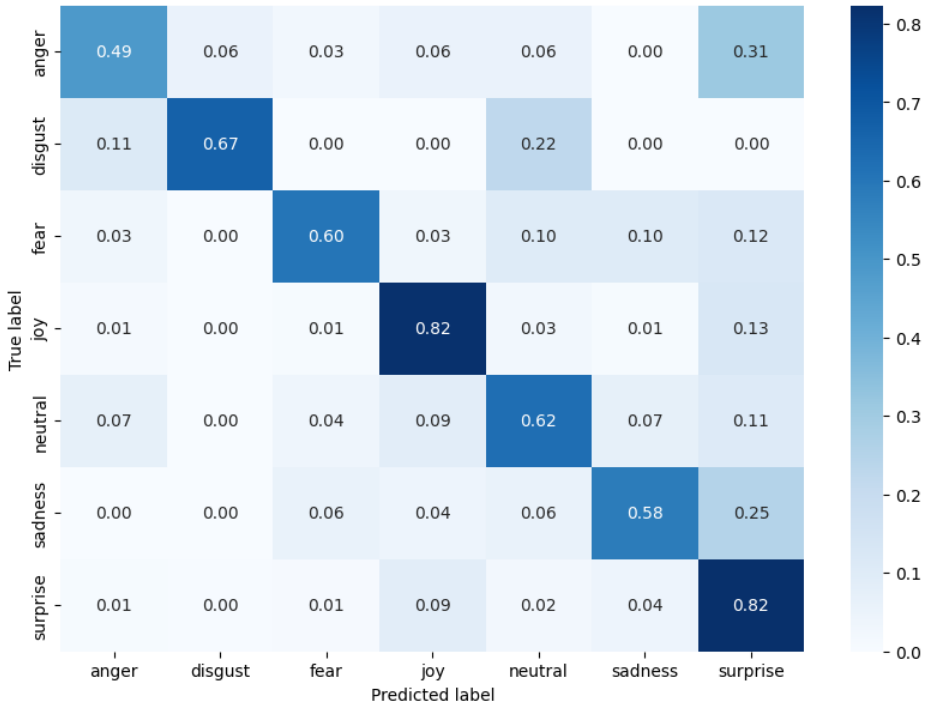

Figure 1: Confusion Matrix for RoBERTa-base (Row = True Label, Column = Predicted Label)

Source: Own

Lower performance was observed for anger, which achieved a recall of 0.486 and an F1-score of 0.531, as well as sadness (F1 = 0.578) and neutral (F1 = 0.588). These results suggest that certain emotional states may be expressed through more subtle linguistic cues, making them harder to capture using textual signals alone.

In contrast, the model achieved strong performance for surprise (F1 = 0.835) and joy (F1 = 0.767), both of which appear frequently in the dataset and are often associated with explicit lexical indicators.

Table 4: Per-Class Metrics Table

Emotion	Precision	Recall	F1-score
Surprise	0.848	0.823	0.835
Joy	0.723	0.817	0.767
Anger	0.586	0.486	0.531
Neutral	0.556	0.625	0.588
Sadness	0.574	0.582	0.578
Fear	0.745	0.603	0.667
Disgust	0.600	0.667	0.632

4.2.1 High-Arousal Confusion Patterns

A recurring confusion pattern was observed between several high-arousal emotional categories. While anger maintained moderate recall (0.486), 18 of the 35 anger instances were misclassified, most frequently as joy or surprise.

Inspection of misclassified examples indicates that the model often responds to intensity-related lexical signals without consistently distinguishing emotional valence. In the absence of acoustic prosody or facial cues, strongly expressive language may be interpreted as high-arousal positivity rather than high-arousal negativity. This suggests that arousal and valence dimensions are not fully disentangled within the textual representation learned by the model.

4.3 Explainable AI Audit

To investigate model behavior beyond aggregate performance metrics, a Gradient × Input analysis was conducted to analyze token-level attribution patterns.

4.3.1 Gradient × Input Attribution

Gradient × Input (Shrikumar et al., 2017) was used to compute the partial derivative of the output with respect to input embeddings, producing token-level relevance scores. Across multiple examples, the model frequently assigned disproportionate weight to isolated, emotionally salient tokens. In several misclassifications, these tokens dominated the prediction despite contradictory contextual information within the same sentence.

To visually represent these attributions, **bold text** is used below to denote tokens with the highest positive attribution weights (those most responsible for the model's chosen label). For example, the sentence "High stakes, you only win by taking **risks**" was labeled anger but predicted as disgust. The mathematical audit indicated that the intensity-related surface-level signal of the word **risks** outweighed the broader competitive context of the statement.

In another case, "I **love** to eat good things" was labeled joy but predicted neutral, illustrating how positive lexical cues like **love** do not always dominate classification when the surrounding contextual signals are structurally weak. A different failure pattern is observed in the sentence "we all **expected** something completely **different**," which was labeled as neutral but predicted as joy. Here, the model appears to overgeneralize and assign emotional meaning to the words **expected** and **different**, suggesting that emotionally suggestive words can trigger classification even in the absence of a clear affective signal. Overall, these token-level attributions suggest that the model relies on sparse lexical cues rather than distributed contextual reasoning, often misclassifying utterances when isolated words conflict with the broader conversational context.

5 Discussion

5.1 The Ground Truth Paradox and Epistemological Validity

The study identifies a critical epistemological risk defined as the "Ground Truth Paradox"**.** Evaluating the classifier against AI-generated labels creates a circular dependency where models are optimized to reproduce the patterns of other models rather than verified emotional states. Even with high accuracy, such systems may

reflect system-level consistency within a specific labeling framework rather than definitive emotional truth. This dependency suggests that future research must prioritize native human-annotated "gold standards" to move beyond automated bias.

5.2 Regulatory Implications: The EU AI Act and "Context Blindness"

The identified 9.7% translation error rate and the model's reliance on isolated lexical cues highlight the fragility of cross-lingual affective pipelines. For organizations deploying these tools under the EU AI Act, particularly in high-risk domains such as recruitment, mental health, or creditworthiness, this "Context Blindness" represents a significant technical and ethical liability.

Because text-only systems are prone to misinterpreting high-arousal states due to a lack of situational "situatedness," their deployment risks automating systemic bias. We conclude that responsible AI development must transition from unimodal text-mining to multimodal, context-aware architectures to bridge the gap between categorical convenience and the complex reality of human affective experience.

5.3 Limitations

Several factors bound the generalizability of this study. First, the evaluation focuses on a single-domain dataset (*MasterChef Polska*), where the emotional repertoire is skewed toward high-arousal interactions. Second, a distinct domain and language shift exists between the training and evaluation phases. While the inclusion of diverse unscripted international programming provided a broad emotional foundation for training, the model was evaluated on a specific Polish-to-English pipeline that was only partially represented during the learning stage. Finally, reliance on a single-sentence unit of analysis prevents the model from utilizing broader conversational flow.

Acknowledgment

The authors would like to express their sincere gratitude to Mekselina Doğanç for her excellent mentorship and valuable guidance throughout the project.

This research article was developed with the assistance of Gemini (Google, Gemini 3 Flash) and ChatGPT (OpenAI, GPT-5). These tools were utilized as a writing aid to assist with adjusting and rephrasing original sentences and providing feedback on the logical structure of the text.

References

Azari, B., Westlin, C., Satpute, A. B., Hutchinson, J. B., Kragel, P. A., Hoemann, K., Khan, Z., Wormwood, J. B., Quigley, K. S., Erdogmus, D., Dy, J., Brooks, D. H., & Barrett, L. F. (2020). Comparing supervised and unsupervised approaches to emotion categorization in the human brain, body, and subjective experience. *Scientific Reports*, *10*(1), 20284. https://doi.org/10.1038/s41598-020-77117-8

Barrett, L. F. (2006). Solving the Emotion Paradox: Categorization and the Experience of Emotion. *Personality and Social Psychology Review*, *10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F. (2011). Constructing emotion. *Psihologijske Teme*, *20*(3), 359–380.

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23. https://doi.org/10.1093/scan/nsw154

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21*, 610–623. https://doi.org/10.1145/3442188.3445922

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550–553. https://doi.org/10.1037/0033-295x.99.3.550

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Youcanprint.

Khare, Y., Peyash, T., Vanzo, A., & Yoshioka, T. (2025). *Universal-2-TF: Robust All-Neural Text Formatting for ASR* (arXiv:2501.05948). arXiv. https://doi.org/10.48550/arXiv.2501.05948

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, The Biopsychology of Emotion: Current Theoretical and Empirical Perspectives*, *84*(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach* (arXiv:1907.11692). arXiv. https://doi.org/10.48550/arXiv.1907.11692

Loshchilov, I., & Hutter, F. (2019). *Decoupled Weight Decay Regularization* (arXiv:1711.05101). arXiv. https://doi.org/10.48550/arXiv.1711.05101

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209–237. https://doi.org/10.1080/02699930802204677

Picard, R. W. (2000). *Affective Computing*. MIT Press.

Radford, A., Kim, J. W., Xu, T., Brockman, G., Mcleavey, C., & Sutskever, I. (2023). Robust Speech Recognition via Large-Scale Weak Supervision. *Proceedings of the 40th International Conference on Machine Learning*, 28492–28518. https://proceedings.mlr.press/v202/radford23a.html

Regulation (EU) 2024/1689 (2024). http://data.europa.eu/eli/reg/2024/1689/oj

Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J.-M. (2003). Facial and Vocal Expressions of Emotion. *Annual Review of Psychology*, *54*(Volume 54, 2003), 329–349. https://doi.org/10.1146/annurev.psych.54.101601.145102

Schuller, B., Mallol-Ragolta, A., Almansa, A. P., Tsangko, I., Amin, M. M., Semertzidou, A., Christ, L., & Amiriparian, S. (2026). Affective computing has changed: The foundation model disruption. *Npj Artificial Intelligence*, *2*(1), 16. https://doi.org/10.1038/s44387-025-00061-3

Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning Important Features Through Propagating Activation Differences. *Proceedings of the 34th International Conference on Machine Learning*, 3145–3153. https://proceedings.mlr.press/v70/shrikumar17a.html

Sugiyama, A., & Yoshinaga, N. (2019). Data augmentation using back-translation for context-aware neural machine translation. In A. Popescu-Belis, S. Loáiciga, C. Hardmeier, & D. Xiong (Eds.), *Proceedings of the Fourth Workshop on Discourse in Machine Translation (DiscoMT 2019)* (pp. 35–44). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-6504

Tiedemann, J., & Thottingal, S. (2020). OPUS-MT – Building open translation services for the World. In A. Martins, H. Moniz, S. Fumega, B. Martins, F. Batista, L. Coheur, C. Parra, I. Trancoso, M. Turchi, A. Bisazza, J. Moorkens, A. Guerberof, M. Nurminen, L. Marg, & M. L. Forcada (Eds.), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (pp. 479–480). European Association for Machine Translation. https://aclanthology.org/2020.eamt-1.61/

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., Platen, P. von, Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T. L., Gugger, S., … Rush, A. M. (2020). *HuggingFace's Transformers: State-of-the-art Natural Language Processing* (arXiv:1910.03771). arXiv. https://doi.org/10.48550/arXiv.1910.03771

Wu, Y., Mi, Q., & Gao, T. (2025). A Comprehensive Review of Multimodal Emotion Recognition: Techniques, Challenges, and Future Directions. *Biomimetics*, *10*(7). https://doi.org/10.3390/biomimetics10070418

Future of Work in the Gastronomy Sector: Challenges of Adopting Artificial Intelligence in a Service Industry

Abdullah Sardogan,[1] Jan Wirsam,[1] Rainer Alt[2]

[1] HTW Berlin, Faculty 3 Economics and Law, Berlin, Germany
sardogan@htw-berlin.de, jan.wirsam@htw-berlin.de
[2] Uni Leipzig, Faculty of Economic Sciences, Leipzig, Germany
rainer.alt@uni-leipzig.de

The digital transformation and Artificial Intelligence (AI) integration promise to restructure complex service operations from back-stage management to front-stage customer interactions. However, while theoretical benefits of automation and augmentation are recognized, practical implementation in gastronomy frequently stalls, overlooking specific socio-technical realities of physical environments. To address this gap, we conducted a qualitative, exploratory study based on expert interviews (N=6) using a dyadic sample of AI providers and gastronomy operators. Findings indicate that AI integration is rarely hindered by algorithmic flaws but collides with a multi-layered socio-technical integration gap: (1) unstandardized analog processes preventing deterministic data extraction, (2) infrastructural debt from legacy IT, and (3) monopolistic vendor politics enforcing lock-ins and stifling scalability. This paper contributes to the Future of Work discourse by demonstrating that algorithms cannot optimize fragmented physical workflows, highlighting process standardization, workflow virtualization, and open data standards as critical prerequisites.

DOI
https://doi.org/10.18690/um.fov.4.2026.45

ISBN
978-961-299-152-4

Keywords:
artificial intelligence, future of work, socio-technical systems, task automation, task augmentation, legacy IT, vendor lock-in, complex service systems

University of Maribor Press

1 Introduction

Rapid advancements in artificial intelligence (AI) are fundamentally transforming the world of work by automating routine tasks, improving productivity, and shifting skill requirements (Mossavar-Rahmani & Zohuri, 2024). While AI adoption accelerates globally, the traditional service industry and particularly the gastronomy and hospitality sectors faces unique hurdles. Historically slow to adopt technological innovations (Gupta & Godiyal, 2025), these sectors operate through complex, intertwined front-stage and back-stage processes. Today, severe labor shortages and growing customer demand exert immense pressure on this industry to urgently integrate AI and evaluate robotic process innovations for both task automation and workforce augmentation (Ferdous et al., 2023; Blöcher & Alt, 2021).

However, the existing "Future of Work" literature often fails to adequately address the specific dynamics of the service sector, particularly the acute labor shortages, escalating human resource costs, and the rapid push towards self-service innovations (Dăniloaia & Turturean, 2024; Lorenz et al., 2023; Pennathur et al., 2024). Consequently, there is a distinct lack of empirical understanding regarding the practical barriers to AI implementation at the operational base. While broader studies indicate that digital initiatives rarely fail due to technological immaturity but rather structural obstacles (Ivchyk, 2024; Kar & Kushwaha, 2023; Rjab et al., 2023), the specific manifestations of these barriers in gastronomy where modern algorithms collide with fragmented physical workflows and monopolistic legacy IT remain underexplored.

To address this gap, this study poses the following research question: What socio-technical integration gaps emerge as primary barriers during the adoption of Artificial Intelligence in the service industry?

To answer this question, we conducted an exploratory, qualitative study based on expert interviews (N=6) using a dyadic sample of leading AI founders and CEOs (supply-side) alongside senior corporate and institutional gastronomy operators (demand-side)[1]. Our findings reveal that successful AI adoption is primarily hindered not by algorithmic flaws or employee resistance, but by a multi-layered

[1] The sample intentionally comprises C-level executives, founders, and senior directors who are directly responsible for AI scaling and strategic IT procurement, thereby ensuring profound domain expertise.

socio-technical gap. Integration frequently stalls at unstandardized analog processes, rigid legacy IT systems, and, critically, deliberate *vendor lock-in* strategies by established software monopolies. These barriers disrupt both autonomous task automation and workforce augmentation, forcing manual workarounds and paralyzing the realization of "New Work" concepts.

By uncovering these structural bottlenecks, this paper contributes to the Future of Work discourse by illustrating that standardizing and virtualizing physical workflows and breaking historical software monopolies are prerequisites for AI integration. The paper proceeds as follows: Section 2 details the theoretical background, followed by the qualitative methodology (Section 3) and the empirical findings (Section 4). Section 5 discusses the socio-technical implications, before Section 6 concludes the paper.

2 Theoretical Background

2.1 AI in Complex Service Systems: Front-Stage and Back-Stage Operations

The gastronomy and broader hospitality sectors are fundamentally structured as complex service systems. Unlike purely digital or manufacturing environments, value creation in these sectors relies on a highly interdependent network of physical workflows, human coordination, and informational exchanges. Service management literature traditionally divides these operations into two distinct but deeply interconnected spheres: "front-stage" processes directly visible to the customer (e.g., ordering, checkout) and hidden "back-stage" processes (e.g., kitchen production, inventory management) (Glushko & Tabas, 2009).

While the restaurant industry has recently begun navigating a broader digital transformation across both of these operational layers to drive efficiency (Alt, 2021), historically, these physical workflows have been characterized by a low degree of digitalization and highly heterogeneous, unstandardized manual routines. However, facing acute labor shortages and efficiency pressures, the industry is currently attempting to integrate Artificial Intelligence (AI) to optimize these operations. While the theoretical deployment of AI promises significant benefits, integrating highly deterministic, data-centric technologies into a historically unstandardized,

physical environment introduces immense socio-technical complexity. AI algorithms require clean, standardized data points to function; thus, the underlying analog processes must be strictly standardized before any digital optimization can occur.

2.2 Restructuring Service Work: Task Automation and Augmentation

The integration of AI into these complex systems is intended to profoundly restructure the nature of service work. Within this context, the literature conceptualizes the impact of AI through two primary mechanisms: task automation and task augmentation (Agrawal et al., 2023; Lei & Kim, 2024).

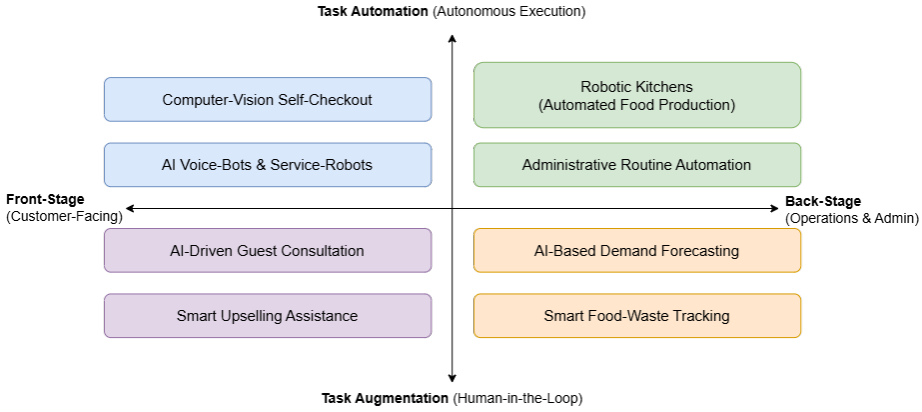

Figure 1: An AI Restructuring Matrix in Complex Service Systems. Categorizing exemplary AI use cases based on their operational domain (front-stage vs. back-stage) and their impact on human labor (augmentation vs. automation)

Source: Own

Task automation refers to the deployment of AI and robotics to autonomously execute specific, routine processes, effectively substituting human labor for those tasks. This is increasingly recognized as a crucial process innovation in high-contact service environments (Blöcher & Alt, 2021). Examples include autonomous computer-vision checkout systems and robotic kitchen production. Conversely, task augmentation deliberately keeps the human worker "in the loop." AI systems, such as demand forecasting algorithms, are utilized to support and enhance human decision-making in the back-stage, elevating the employee's capabilities without

replacing them. The ultimate goal of this restructured "New Work" environment in gastronomy is to free the workforce from repetitive administrative burdens, allowing them to focus on high-value, specialized tasks (Lavoie-Gagne et al., 2025). However, achieving this relies heavily on seamless data flows and interoperability across the entire service system.

To systematically categorize how AI reshapes these distinct operational domains, Figure 1 synthesizes the theoretical dimensions of service work. By crossing the operational domain (front-stage vs. back-stage) with the systemic impact on human labor (task automation vs. task augmentation), exemplary AI use cases emerge within the gastronomy sector.

2.3 Socio-Technical Integration Barriers: Legacy IT and Politics

To understand why the restructuring of service work frequently stalls in practice, gastronomy operations must be viewed through the lens of Socio-Technical Systems (STS) theory. A core principle is "joint optimization": implementing AI will only yield benefits if the structural and political foundation is adapted simultaneously (Sony & Naik, 2020). When modern AI collides with rigid or outdated structures, "socio-technical integration gaps" emerge.

Current literature identifies infrastructural debt commonly referred to as legacy IT as a primary technical barrier (Arora, 2025). Many traditional gastronomy businesses rely on fragmented, proprietary legacy systems (e.g., old Point-of-Sale or Enterprise Resource Planning systems) that lack the open Application Programming Interfaces (APIs) required for modern AI integration, creating massive data silos.

Crucially, recent critical perspectives expand on this by highlighting that these technological silos are often maintained by *vendor politics* and *institutional macro-politics* (Archer et al., 2025; Thylstrup et al., 2024). Rather than mere technical inability, established legacy system providers frequently utilize their market monopolies to enforce vendor lock-in effects. By deliberately restricting API access or charging exorbitant integration fees, these providers strategically shield their closed ecosystems from innovative third-party AI solutions. Furthermore, in highly regulated sectors such as public catering, this integration gap is exacerbated by macro-political fragmentation, where differing bureaucratic requirements across

municipalities stifle scalable digital rollouts. While these monopolistic and institutional barriers are theoretically acknowledged, empirical evidence exploring how they actively disrupt the restructuring of frontline service work remains scarce.

3 Methodology

3.1 Research Design and Case Selection

To explore the socio-technical integration gaps of AI in complex service systems, this study employs a qualitative, exploratory research design. Given the underexplored intersection of modern AI, legacy IT, and physical workflows in gastronomy, an interview-based approach provides the depth required to uncover underlying socio-technical dynamics.

3.2 Data Collection and Sampling

Data was gathered through semi-structured expert interviews using purposive sampling to ensure high domain expertise. We utilized a dyadic sample of six (N=6) senior experts from both the supply side (AI founders/CEOs) and the demand side (corporate and institutional leaders) to ensure a holistic perspective. This seniority and dyadic approach provide robust "information power" regarding structural barriers despite the compact sample size. Interviews were conducted in German, audio-recorded, and transcribed verbatim. Table 1 summarizes the expert profiles.

Table 1: Expert Interview Sample Profiles

Interview ID	Role	Organization Type	Focus
A	Founder & CEO	AI Tech Startup	Computer Vision & Autonomous Self-Checkout
B	Founder & CEO	AI Tech Startup	Machine Learning Demand Forecasting
C	Founder & CEO	AI Tech Startup	Digital Guest Consultation & Storytelling
D	Founder & CEO	AI Tech Startup	Automating and digitizing invoice verification
E	Corporate Gastronomy Expert	Multinational Corporate Catering & Facility Management	AI Scaling in Corporate Catering, Cultural Transformation, Piloting food service innovations
F	Foundation Expert	Non-Profit Foundation for Food & Nutrition Education	Providing dietary consultation, nutritional education, and Institutional kitchen management

3.3 Data Analysis

Transcripts were analyzed using thematic analysis following Braun and Clarke (2006). Given the exploratory nature and compact sample (N=6), coding was conducted manually. An iterative process combined deductive coding based on our theoretical framework (e.g., task automation, infrastructural debt) with inductive open coding to capture emergent, industry-specific phenomena (e.g., vendor politics, lack of process standardization).

These initial codes were continuously refined and grouped into the four higher-order socio-technical integration gaps conceptualized in Figure 2. To ensure analytical rigor, emerging themes were constantly compared across the dyadic sample contrasting AI providers with gastronomy operators to validate the structural nature of the identified barriers. Finally, representative quotes were translated into English, preserving contextual nuances.

4 Findings

The analysis of the expert interviews suggests that the primary hurdles to successfully restructuring service work through AI are rarely algorithmic flaws or a fundamental unwillingness of the service workforce. By analyzing the dyadic perspectives of both AI providers and gastronomy operators, it becomes evident that the implementation of both autonomous systems and augmenting tools consistently stalls at two major socio-technical integration gaps: severe infrastructural debt combined with unstandardized processes, and complex market and institutional politics.

4.1 The Precondition of Process Standardization and Infrastructural Debt

Before technical interfaces can even be addressed, the demand-side experts emphasized a severe lack of operational process standardization as the foundational barrier. The integration of AI requires deterministic structures; however, the physical preparation and administration in complex service systems remain highly heterogeneous. As one corporate gastronomy expert highlighted, AI cannot optimize a broken or unstandardized workflow:

> "In many cases, AI startups and solutions do not fail because of a flawed technology or problem-solving approach, but rather because of the rigid, fragmented organizational and IT structures they are confronted with." (Corporate Gastronomy Expert 1)

Without a unified analog process, AI algorithms lack the necessary data points to function effectively:

> "If the underlying operational processes are not strictly standardized, there is simply no point of attack for the AI. This severe lack of process standardization makes it incredibly difficult for AI systems to extract valid data and generate any actual added value." (Corporate Gastronomy Expert 1)

The sheer scale of this operational fragmentation often outpaces technological advancements, rendering "plug-and-play" AI deployments impossible for large operators:

> "The fundamental hurdle is operating 400 canteens with 400 completely different workflows. Ultimately, AI technology will advance much faster than the operational ability of the gastronomy sector to align and standardize its internal processes." (Corporate Gastronomy Expert 1)

4.2 Infrastructural Debt and the Cost of Legacy IT

When standardized processes do exist, the next major barrier is the reliance on rigid, historically grown IT infrastructures. Traditional gastronomy operations are built on legacy Point-of-Sale (POS) and Enterprise Resource Planning (ERP) systems that are fundamentally incompatible with modern, cloud-based AI.

> "Many legacy point-of-sale systems have grown historically and are simply not cloud-native solutions. They are based on architectures that are 20 or 30 years old. Instead of investing in new, open infrastructure, providers continue to just 'milk the cow,' adding superficial web plugins to heavily closed systems." (AI Provider 3)

This infrastructural debt severely disrupts front-stage task automation. AI providers report massive resource drains just trying to establish basic connectivity:

> "It is the very old-school point-of-sale systems that are the real pain. The fundamental pain point of the industry right now is the sheer lack of standardized protocols and interfaces that would allow systems to communicate with each other." (AI Provider 1)

> "I would estimate that we invest over a third of our entire development resources solely into system integration with the existing IT landscape." (AI Provider 1)

For back-stage augmentation, such as AI-driven demand forecasting, missing automated interfaces actively create new administrative burdens. The resulting data delays drastically reduce the algorithm's predictive value:

> "Imagine you are a kitchen manager updating the menu on Monday. Because automated interfaces are missing, our AI only receives the manual data export on Friday. For the entire week, the algorithm operates on outdated information, severely limiting its predictive value." (AI Provider 2)

Consequently, employees are forced into manual workarounds, contradicting the core premise of AI-driven relief:

> "Currently, the kitchen manager has to keep their legacy inventory system open alongside our AI tool. They are forced to work across two different formats and manually transfer the data between them, which prevents any seamless task augmentation." (AI Provider 2)

4.3 Vendor Politics, Monopolies, and the Illusion of Open APIs

Beyond the technical challenges, the interviews exposed massive barriers within the political subsystem. Notably, the integration gap is deliberately widened by external vendor politics. AI providers unanimously reported that legacy IT providers maintain closed systems not out of technical inability, but as a strategic mechanism to protect their market share.

> "The biggest absolute showstopper for AI integration are the political blockades imposed by legacy POS providers. The main difficulty is not the technology itself, but rather the deliberate lack of integration-willingness from established vendors protecting their market share." (AI Provider 3)

> "Most legacy systems have interfaces in some form, and the actual technical effort to develop them is completely manageable. Instead, the biggest—the absolute biggest—issue is politics." (AI Provider 1)

This gatekeeping is largely driven by a systemic fear of replacement. As AI startups introduce highly advanced augmentation and automation tools, legacy providers act defensively to secure their ecosystems:

> "A new AI system docks onto an existing system, and the immediate reflex of the established system—and the market—is fear and concern that the new technology will take something away from them. The initial reaction is always: 'Oh, something new is coming, they want to replace us, we have to be careful.'" (AI Provider 1)

Established providers intentionally utilize this monopolistic power to artificially inflate integration costs, effectively pricing innovative AI out of the market:

> "The established legacy system providers know exactly what kind of leverage and structural power they hold over the market. Because there are few alternatives, they use this monopoly to charge exorbitant fees for basic data interfaces, artificially creating massive political and financial barriers for AI integration." (AI Provider 2)

To mask these blockades, legacy providers often rely on deceptive marketing regarding their technical openness:

> "For marketing reasons, every legacy system provider claims to have an 'open API.' But the reality is highly restrictive. You might only be able to retrieve aggregated daily revenue, but you cannot process real-time orders. The lack of open standards is often a deliberate political blockade rather than a technical limitation." (AI Provider 3)

This culminates in the aggressive pushing of "all-in-one" solutions, which creates severe vendor lock-in and prevents gastronomy operators from utilizing specialized AI tools:

> "Large legacy providers often push 'one-solution-fits-all' systems to intentionally create a vendor lock-in effect. They play on the management's fear of digital complexity. However, this political strategy severely harms the gastronomy business, because a single closed system can never seamlessly cover every specific operational need or integrate the best AI tools." (AI Provider 3)

4.4 The Cost of Uncertainty and Cultural Resistance

The culmination of rigid infrastructure, monopolistic gatekeeping, and unstandardized processes creates an environment of immense uncertainty. For AI providers, this socio-technical friction is the primary reason for project failure:

> "When we evaluate the deals we lose, the number one 'Reason for Loss' in our CRM is missing integration. The sheer uncertainty—not knowing if an interface will be built, when it will be ready, or what it will cost—is the factor that ultimately destroys projects." (AI Provider 1)

Finally, when gastronomy management attempts to navigate these external blockades, they frequently underestimate the internal, human dimension of the restructuring process. The successful implementation of task automation and augmentation requires a workforce prepared to embrace new workflows:

> "Digital transformation is, at its core, a cultural transformation. Any technological innovation is only as effective as the foundational cultural will within the organization to actually implement it and adapt to it." (Corporate Gastronomy Expert 1)

Ultimately, the collision of modern AI algorithms with fragmented analog workflows, monopolistic vendor politics, and cultural inertia acts as the definitive socio-technical integration gap preventing the Future of Work in the service industry.

5 Discussion

5.1 Synthesizing the Socio-Technical Integration Gaps

This study examined how legacy IT and organizational politics emerge as socio-technical integration gaps during AI adoption. Our findings confirm that AI restructuring rarely fails due to algorithmic flaws. Rather, it stalls at a multi-layered collision between advanced algorithms and rigid socio-technical structures. This cascading failure is conceptualized in **Figure 2**, illustrating how the theoretical ambition of AI is systematically blocked across four layers.

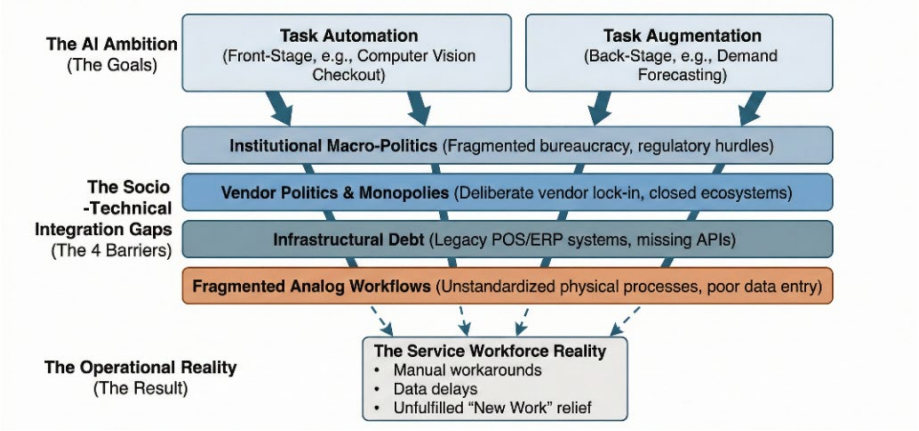

Figure 2: Multi-Layered Socio-Technical Integration Gap in Complex Service Systems. This conceptual model illustrates how the intended restructuring of service work (automation and augmentation) is systematically disrupted by institutional, political, technical, and operational barriers, resulting in manual workarounds for the workforce

Source: Own

While current 'Future of Work' literature predominantly focuses on individual AI adoption factors like skill shifts or algorithm aversion (Mossavar-Rahmani & Zohuri, 2024; Dăniloaia & Turturean, 2024), the model in Figure 2 reveals that integration rarely reaches the human-algorithm interface. Viewed through the STS lens of "joint optimization", AI stalls at structural hard-stops that preemptively override individual-level acceptance. Specifically, the restructuring is paralyzed by unstandardized analog workflows and "infrastructural debt" a debt deliberately maintained by monopolistic legacy providers enforcing vendor lock-ins.

Compounded by institutional macro-politics, these systemic barriers fundamentally block the restructuring of service work.

5.2 Theoretical Contributions

This paper offers two distinct theoretical contributions. First, it expands Future of Work literature beyond knowledge workers to physical service systems, identifying analog process standardization as a foundational prerequisite for AI integration. As our findings strongly suggest, algorithms struggle to optimize fragmented workflows.

Second, our dyadic approach broadens the perspective on the political dimension of STS theory. While existing literature focuses on internal resistance, like "algorithm aversion," our exploratory findings indicate that external forces, market monopolies (vendor politics), and institutional fragmentation (macro-politics), act as major political barriers that prevent scalable task automation.

5.3 Practical Implications

For practitioners, these insights demand a paradigm shift and the adoption of holistic AI integration frameworks to successfully navigate complex service environments. For AI startups and technology providers, the "plug-and-play" illusion has to end. Instead of exhausting resources on costly, customized 1-to-1 workarounds, providers must actively participate in industry-wide alliances to develop vendor-agnostic middleware architectures and open data protocols (comparable to the E-Class standard). For gastronomy management and corporate operators, the mandate is twofold. First, they must prioritize the rigorous standardization or complete virtualization (e.g., shifting on-site ordering to digital pre-order platforms) of their physical, analog processes before deploying AI. Second, management must leverage its purchasing power to break vendor lock-ins. During IT procurement, decision-makers must strictly mandate genuinely open APIs, thereby forcing the market towards interoperability.

6 Conclusion, Limitations, and Future Research

6.1 Conclusion

The digital transformation of the gastronomy sector promises unprecedented operational efficiency, yet AI's theoretical potential currently collides with the harsh realities of the service industry. This study highlights that successful AI integration is a profound socio-technical challenge, not merely an algorithmic one. Our exploratory study highlights that AI adoption stalls across interconnected layers: unstandardized analog workflows at the foundation , infrastructural debt from legacy IT , and monopolistic vendor politics that actively block innovative tools. To unlock the "Future of Work," the industry must undergo a paradigm shift. Operators must rigorously standardize and virtualize physical processes, while leveraging their collective purchasing power to enforce open, vendor-agnostic data standards. Only by modernizing this socio-technical foundation can AI successfully elevate the nature of service work.

6.2 Limitations and Future Research

While providing deep empirical insights, this qualitative study is limited to six highly specialized experts in the DACH region, restricting statistical generalizability. Future research should employ large-scale quantitative surveys among gastronomy operators and IT procurement officers to measure the exact financial and temporal impacts of these integration gaps. Additionally, longitudinal studies could observe how political dynamics between legacy providers and AI startups evolve amid increasing regulatory pressure for interoperable middleware architectures.

References

Agrawal, A., Gans, J. S., & Goldfarb, A. (2023). Do we want less automation?. *Science*, *381*(6654), 155-158.

Alt, R. (2021). Digital transformation in the restaurant industry: Current developments and implications. *Journal of smart tourism*, *1*(1), 69-74.

Archer, M., Ravn, L., & Thylstrup, N. B. (2025). The political economy of platformed silos: Theorizing data storage reconfigurations in the age of interoperability capitalism. *Big Data & Society*, *12*(2), 20539517241303144.

Arora, A. (2025). Challenges of Integrating Artificial Intelligence in Legacy Systems and Potential Solutions for Seamless Integration. *Available at SSRN 5268176*.

Blöcher, K., & Alt, R. (2021). AI and robotics in the European restaurant sector: Assessing potentials for process innovation in a high-contact service industry: K. Blöcher, R. Alt. *Electronic Markets*, *31*(3), 529-551.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Dăniloaia, D. F., & Turturean, E. (2024). Knowledge Workers and the Rise of Artificial Intelligence: Navigating New Challenges. *SEA: Practical Application of Science*, *12*(35).

Ferdous, S. F., Badar, M. A., & Lin, Y. (2023). Effect of Rising Cost and Worker Shortage on Industry. *Computer*, *12*(1).

Glushko, R. J., & Tabas, L. (2009). Designing service systems by bridging the "front stage" and "back stage". *Information Systems and E-Business Management*, *7*(4), 407-427.

Gupta, P., & Godiyal, M. AI Adoption in the Hospitality Industry: Opportunities and Challenges. *INDIAN JOURNAL OF HOSPITALITY MANAGEMENT*, 80.

Ivchyk, V. (2024). Overcoming barriers to artificial intelligence adoption. *Three Seas Economic Journal*, *5*(4), 14-20.

Kar, A. K., & Kushwaha, A. K. (2023). Facilitators and barriers of artificial intelligence adoption in business–insights from opinions using big data analytics. *Information Systems Frontiers*, *25*(4), 1351-1374.

Lavoie-Gagne, O., Woo, J. J., Williams III, R. J., Nwachukwu, B. U., Kunze, K. N., & Ramkumar, P. N. (2025). Artificial intelligence as a tool to mitigate administrative burden, optimize billing, reduce insurance-and credentialing-related expenses, and improve quality assurance within health care systems. *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, *41*(8), 3270-3275.

Lei, Y. W., & Kim, R. (2024). Automation and augmentation: artificial intelligence, robots, and work. *Annual Review of Sociology*, *50*(1), 251-272.

Lorenz, H., Stephany, F., & Kluge, J. (2023). The future of employment revisited: How model selection affects digitization risks. *Empirica*, *50*(2), 323-350.

Mossavar-Rahmani, F., & Zohuri, B. (2024). Artificial intelligence at work: Transforming industries and redefining the workforce landscape. *Journal of Economics & Management Research*, *5*(2), 2-4.

Pennathur, P. R., Boksa, V., Pennathur, A., Kusiak, A., & Livingston, B. A. (2024). The future of office and administrative support occupations in the era of artificial intelligence: A state of the art review and future research directions. *International Journal of Industrial Ergonomics*, *104*, 103665.

Rjab, A. B., Mellouli, S., & Corbett, J. (2023). Barriers to artificial intelligence adoption in smart cities: A systematic literature review and research agenda. *Government information quarterly*, *40*(3), 101814.

Sony, M., & Naik, S. (2020). Industry 4.0 integration with socio-technical systems theory: A systematic review and proposed theoretical model. *Technology in society*, *61*, 101248.

Thylstrup, N. B., Archer, M., & Steiner, H. (2024). Desiloization and its discontents: The politics of data storage in the age of platformization. *Information, Communication & Society*, *27*(13), 2419-2437.

User Frictions in Process Intelligence Dashboards: Implications for AI Assistants

Candan Cetin,[1,2] Teodora Lata[2]

[1] Technical University of Munich, Munich, Germany
candan.cetin@tum.de
[2] Celonis SE, Munich, Germany
candan.cetin@tum.de, t.lata@celonis.com

Process intelligence platforms increasingly support operational decisions, but their realized value is also related to the degree that users can interpret dashboards and act on insights effectively. We derived user-friction themes from the process intelligence and dashboard literature, operationalized them into a survey, and analyzed responses from 58 active and experienced dashboard users with hypothesis-based OLS regressions using robust standard errors and false discovery rate correction. The results showed that decision-related outcomes were most strongly associated with visual interpretability, question-to-query translation, trust in data credibility, and metric actionability. The most robust supported relationships indicated that clear layouts and labels, effective translation of analytical intent into filters and views, complete and up-to-date data, and decision-relevant metrics aligned with higher ease of use, usefulness, confidence, and decision effectiveness. Based on these findings, we discuss how AI assistants should complement dashboards by reducing friction in understanding, querying, verifying, and acting on process intelligence data.

DOI
https://doi.org/10.18690/um.fov.4.2026.46

ISBN
978-961-299-152-4

Keywords:
process intelligence, dashboards, AI assistants, user experience, information systems

University of Maribor Press

1 Introduction

Process Intelligence (PI) solutions have evolved from niche analytical tools into a core layer of digital operations. By extracting event logs from any source system, PI provides organizations with fact-based visibility into their actual workflows (van der Aalst, 2016; van der Aalst et al., 2012). As the PI software market experiences sustained double-digit growth (Everest Global, Inc., 2025), PI platforms become a part of the companies` digital core.

However, software investments do not automatically yield operational returns. In practice, value realization hinges on end-user adoption; humans must interact with dashboards, interpret results, and act. In this context, end-users encounter challenges related to complex layouts, opaque data origins, and the technical expertise required to extract answers.

At the same time, artificial intelligence (AI) actively reshapes the PI user landscape. Vendors are embedding conversational copilots designed to lower skill barriers and shorten the time-to-value. While current PI literature focuses heavily on backend algorithmic discovery and data integration, specifically the end-user experience aspect of dashboard consumption remains largely underexplored (Ammann et al., 2025; Ankomah, 2025, and Meduri et al., 2021). Consequently, empirical evidence is scarce regarding concrete challenges end-users face and how vendors should design AI assistants` roadmaps to alleviate those frictions.

Accordingly, this study asks: Which dashboard consumption frictions should AI assistants target in process intelligence platforms? Although this study did not test a specific AI assistant intervention, it instead identified empirically supported dashboard frictions that AI assistants may target in the future.

2 Background and Related Work

To understand the landscape of user challenges, we conducted a systematic literature review and the Gioia methodology to synthesize findings into Second-order Themes (SoTs) (Gioia, 2013).

The review revealed 24 SoTs and highlighted 10 of them, which we used as the baseline of our research. To summarize, the top 4 most frequently mentioned themes were:

- Visual Comprehension,
- Trust Issues,
- Expertise Requirement,
- Lack of Actionability.

The most consistently reported problem is the comprehension effort required to interpret highly dense, complex visual representations. Misaligned charts and poor labelling reduce the user`s ability to recognize trends or inefficiencies (Atighehchian et al., 2023; Rehse et al., 2022; Zuidema-Tempel et al., 2022). Secondly, users frequently doubt the accuracy and completeness of dashboard data. A lack of transparency regarding data processing and timestamps leads users to manually verify raw data, undermining the dashboard`s value (Alhamadi et al., 2022; Chinnathai & Alkan, 2023; Simic et al., 2023). Thirdly, dashboards often require users to possess specialized technical skills or familiarity with query languages in order to translate analytical business questions into the appropriate interface filters (Barbieri et al., 2023; Ullrich et al.; Zimmermann, 2023). Lastly, dashboards frequently display retrospective insights without supporting timely, decision-relevant follow-up, forcing users to rely on manual workflows to execute tasks (Morelli et al., 2024; Rehse et al., 2022; Simic et al., 2023; Zimmermann, 2023).

Most of the PI vendors have introduced AI and Large Language Models (LLMs) into their ecosystems (Carpenter, 2025; Everest Global, Inc., 2025; Wilson, 2024). Current AI assistant capabilities aim to democratize process data (Boareto et al., 2025). LLMs provide natural-language analytics and situation-aware explainability (Fahland et al., 2025).

3 Methodology

3.1 Conceptualization and Operationalization

To systematically translate qualitative literature into testable quantitative metrics, we used a traceable approach. To illustrate this approach, we observed the development of Hypothesis 1 (H01) (Eisenhardt, 1989). We started by extracting raw literature excerpts from articles in our sample space. H01 was traced back to the excerpts of works of Alhamadi et al. (2022)[1] and Zuidema-Tempel et al. (2022)[2]. Then, we translated these raw excerpts into First-Order Concepts (FoCs) which were problem statements such as "Some users verify data manually, indicating distrust or lack of confidence in the dashboard." and "Inaccurate or incomplete data undermines user confidence in dashboard outputs." We grouped together related FoCs under the theoretical Second-Order Themes (SoTs). In our example above, these two FoCs fall under the "Visual Comprehension" SoT. To create a falsifiable claim, we refined this theme into a specific, design-actionable sub-theme (Visual Layout & Label Clarity) and linked it to an Information Systems (IS) outcome. This generated "H01: Clear layouts and labels increase perceived ease of use." After the hypothesis formation, we measured the variables using a 6-point Likert scale in the survey by user responses. The predictor was measured using the survey item: "The visual layout, labels, and axes make data easy to understand" and the outcome was measured using the Technology Acceptance Model (TAM) item (Davis, 1989): "Learning to use this dashboard is easy for me." (PEOU).

3.2 Survey Design and Sample

We operationalized the literature-based hypotheses into a survey measuring user frictions alongside established IS outcomes, specifically the TAM constructs (Davis, 1989; Ives et al., 1983; Wixom & Todd, 2005). We tried to operationalize each

[1] When users are not sure of the accuracy of the information presented, they usually question the origin of the data. If the dashboard does not capture the depth of the data that satisfies users' needs, users tend to check the raw data sets frequently. Developers try to prevent this workaround as they believe it undermines the value of a data overview dashboard Alhamadi et al. (2022).

[2] Zuidema-Tempel et al. (2022) reported that all interviewed experts identified data quality as a major bottleneck, especially because of incorrect data formats and incomplete data entry by employees. One expert explained that "it was difficult to get the data out of the systems," which reduced trust in the availability of good-quality data. The same study further noted that responses such as "we have tried that before but it did not work" were common, and that some process mining projects were terminated because the required data were unavailable.

construct with more than one survey item without excessively extending the survey's length. All attitudinal items were measured on a 6-point Likert scale ranging from 1 = strongly disagree to 6 = strongly agree. Closed-ended survey questions captured the prevalence and perceived severity of the literature-based dashboard-related issues. We also added open-ended questions to each of the three predictor-measuring survey blocks. This gave respondents space to report issues not covered by the predefined survey items and added qualitative richness to the analysis. Specifically, 13 respondents (22.4%) gave input to optional open-ended VSC_FREE, 5 (8.6%) to TRUST_FREE, and 10 (17.2%) to ACT_MER_FREE questions shown after the predictor items blocks of the survey. These rates seem to be consistent with academic benchmarks for optional open-ended items in digital surveys, which typically range between 18%-24% for optional qualitative feedback (Asare-Marfo, 2021; Miller & Lambert, 2014).

Our sample consisted of experienced and active dashboard users with 52% of them having 3+ years experience in the PI area and 79% of them using dashboards on a weekly or daily basis. Regarding AI exposure, 34.5% reported that AI assistant functionality was available and that they had used it at least once, while 41.4% indicated that AI was available but not yet used. Respondents indicated that they need the most help with finding the right dashboard/page/view, digging into anomalies, and understanding metrics consecutively. When they were asked to rank their selected needs based on their AI assistance preference, the order that arose was first around finding the right dashboard, followed by turning insights into actions, and then digging into anomalies. Appendix Section 7.1 documents the recruitment procedure and Appendix section 7.2 summarizes the survey structure and measurement blocks.

3.3 Statistical Analysis

In the data preparation phase, before hypothesis testing, we assessed the internal consistency of multi-item survey constructs using reliability analysis (Cronbach`s alpha) (Cronbach, 1951). Constructs demonstrating acceptable internal consistency were aggregated into composite measures by averaging their items. For instance, Decision Effectiveness (outcome construct) showed satisfactory internal consistency, and, therefore, we used it as a composite of DE ("This dashboard helps me make decisions quickly when needed.") and DC ("I feel confident in the

decisions I make based on this dashboard.") items for hypothesis testing. Several constructs did not demonstrate adequate internal consistency. In particular, items within the Visual Clarity & Findability survey block (VSC1, VSC2, VSC3, VSC4) did not exhibit sufficient internal consistency and were therefore treated as conceptually singleton predictors representing their subthemes (Visual Layout & Label Clarity, Visual Time Context & Trend Readability, Visual Emphasis & Colour Encoding, Visual Glanceability & Information Hierarchy). Appendix Section 7.3 reports the reliability diagnostics and the resulting operationalization decisions for all multi-item constructs. To evaluate which dashboard features were reliably associated with user outcomes, we estimated separate OLS regression models for each of the 60 hypotheses using all completed responses ($N = 58$). We controlled for role, experience, usage frequency, and AI availability. Since the sample size is moderate, we applied the HC3 robust standard errors to obtain conservative and more reliable p-values (MacKinnon & White, 1985; White, 1980). We only considered a hypothesis supported if the three following criteria were satisfied:

- The estimated coefficient had the theoretically expected direction,
- The significance threshold of $p < .05$ was met,
- The Benjamini-Hochberg False Discovery Rate (FDR) adjustment produced $q < .10$ (Benjamini & Hochberg, 1995).

Using the conventional p-value threshold indicated that the observed effects were unlikely under the null model and were therefore less likely to reflect random sampling variation alone (Greenland et al., 2016; Wasserstein & Lazar, 2016). However, FDR adjusted values controlled for false positives arising from testing many hypotheses simultaneously (Bender & Lange, 2001, Benjamini & Hochberg, 1995).

4 Results

Before hypothesis testing, we examined mean predictor ratings to identify the lowest- and highest- rated dashboard components in our sample. Descriptively, items related to data lineage transparency, timestamp correctness, and manual effort exhibited comparatively weaker agreement. Conversely, items capturing visual layout, label clarity, and the ability to work without technical help received comparatively stronger agreement. Across all outcomes, respondents reported the

highest agreement for decision confidence and the lowest agreement for perceived usefulness for job performance. The mean differences across outcome constructs remained small in this sample. Overall, the spread of mean ratings across predictors and outcomes was modest, indicating that respondents did not report severe dissatisfaction but rather relative differences in friction levels within an otherwise positively evaluated dashboard experience.

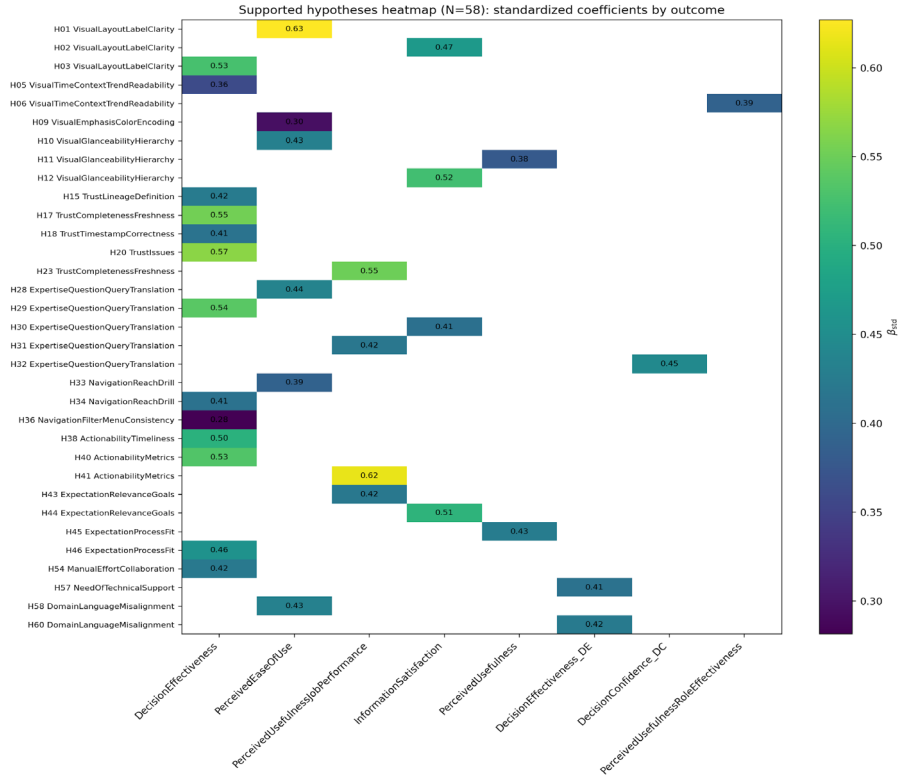

Figure 1: Accepted Hypotheses
Source: Own

Under our strict criteria (positive expected direction, $p < .05$, and FDR-adjusted $q < .10$), 33 of the 60 hypotheses were statistically supported. Figure 1 visualizes all 33 accepted hypotheses (H1, H2, H…) by mapping predictors (VisualLayoutLabelClarity,...) to outcomes (PerceivedEaseofUse, DecisionEffectiveness, PerceivedUsefulness, ...) and colouring the links according

to standardized effect size β_{std}. The results indicated that specific usability, trust, and actionability factors aligned strongly with decision-related outcomes.

Decision Effectiveness, operationalized as the average of items "This dashboard helps me make decisions quickly when needed." (DE) and "I feel confident in the decisions I make based on this dashboard." (DC), emerged as the most central outcome in our sample, with 13 supported predictors.

Table 1: Top 8 Accepted Hypotheses

ID	Hypothesis	Predictor	Outcome
H01	Clear layouts and labels increase perceived ease of use.	Visual Layout & Label Clarity	Perceived Ease of Use
H29	Translating questions into the right filters/views improves decision effectiveness.	Question-to-Query Translation	Decision Effectiveness-composite
H02	Clear layouts and labels increase information satisfaction.	Visual Layout & Label Clarity	Information Satisfaction
H23	Complete, up-to-date data increase perceived usefulness for job performance.	Data Completeness & Freshness	Perceived Usefulness (Job Performance)
H03	Clear layouts and labels improve decision effectiveness.	Visual Layout & Label Clarity	Decision Effectiveness-composite
H41	Specific and appropriately filtered metrics increase perceived usefulness for job performance.	Metric Actionability	Perceived Usefulness (Job Performance)
H20	Higher overall trust in dashboard data improves decision effectiveness.	Trust-composite (Lineage, freshness, timestamps)	Decision Effectiveness-composite
H32	Translating questions into the right filters/views increases decision confidence.	Question-to-Query Translation	Decision Confidence

Source: Own

Although we reported all 33 supported hypotheses in Figure 1, we emphasized the relationships with the smallest FDR-adjusted q-values because they remained the most robust against multiple-testing corrections and were the least likely to reflect chance findings. The eight strongest hypotheses (H01, H29, H02, H23, H03, H41, H20, H32) have the smallest FDR-adjusted q-values and therefore provided the

empirical basis for our main conclusions. Table 1 summarizes these hypotheses. Beyond statistical support (direction, $p < .05$, and $q < .10$), the top eight hypotheses also showed moderate-to-large, standardized effects and tight uncertainty bounds under HC3 robust standard errors.

The strongest result was H01, which linked Visual Layout & Label Clarity to Perceived Ease of Use ($\beta_std = 0.627$, 95% CI [0.442, 0.813], $q = 2.09 \times 10^{-9}$). Other highly robust effects included H29, which linked Question-to-Query Translation to Decision Effectiveness ($\beta_std = 0.539$, 95% CI [0.318, 0.759], $q = 5.3 \times 10^{-5}$), H02, which linked Visual Layout & Label Clarity to Information Satisfaction ($\beta_std = 0.469$, 95% CI [0.270, 0.668], $q = 5.9 \times 10^{-5}$), and H23, which linked Data Completeness & Freshness to Perceived Usefulness for Job Performance ($\beta_std = 0.551$, 95% CI [0.317, 0.785], $q = 5.9 \times 10^{-5}$). H03 further showed that Visual Layout & Label Clarity aligned with Decision Effectiveness ($\beta_std = 0.526$, 95% CI [0.294, 0.758], $q = 1.06 \times 10^{-4}$). H41 linked Metric Actionability to Perceived Usefulness for Job Performance ($\beta_std = 0.616$, 95% CI [0.312, 0.920], $q = 7.28 \times 10^{-4}$), H20 linked the Trust composite to Decision Effectiveness ($\beta_std = 0.567$, 95% CI [0.283, 0.851], $q = 7.67 \times 10^{-4}$), and H32 linked Question-to-Query Translation to Decision Confidence ($\beta_std = 0.445$, 95% CI [0.212, 0.679], $q = 1.42 \times 10^{-3}$). All eight effects had strictly positive 95% confidence intervals, which reinforced that these relationships represented robust positive associations rather than borderline estimates.

5 Discussion

5.1 Limitations

This study relies on self-reported perceptions rather than objective usage or performance measures. Respondents evaluated both predictors and outcomes within the same cross-sectional survey, which may have inflated associations because the method and response tendencies can influence all measures in a similar way. The analytical sample was moderate in size with $N = 58$ complete responses, and it skewed toward experienced users with combined builder-and-viewer roles, which may limit generalizability. Similarly, the employed recruitment approach was network-based rather than a defined sampling framework which might not produce

a holistic representative sample. Besides, several constructs were operationalized as single- item measures (singletons) such as DLM1, PRV1, and IS1 or split into single-item sub-themes after reliability checks as in the case of VSC1, VSC2, VSC3, and VSC4. These choices may have introduced measurement error or reduced coefficient stability. Finally, the regression models identified associations under controls and robust standard errors, but they did not establish causality or test the effects of a specific AI assistant intervention. Future research should validate these findings with larger, more balanced samples and with behavioural outcomes such as task accuracy, time, and interaction traces in controlled studies or field deployments.

5.2 Implications for AI Assistant Design

Under these limitations, the results still identified four friction areas that future AI assistants may target: interpretability, expertise translation, credibility, and actionability. This interpretation also aligned with the broader survey context. The sample consisted of experienced and active dashboard users, and respondents most frequently reported needing help with finding the right dashboard or view, digging into anomalies, and understanding metrics. When they ranked their preferred areas for AI support, they placed finding the right dashboard first, followed by turning insights into actions. These priorities fit well within the supported hypothesis pattern, which converged on four mechanisms: interpretability, expertise translation, credibility, and actionability.

First, the results pointed to interpretability as a foundational mechanism. H01, H02, and H03 showed that clearer layouts and labels aligned with higher perceived ease of use, higher information satisfaction, and higher decision effectiveness. These relationships suggested that respondents evaluated dashboards more positively when the interface reduced visual ambiguity at the point of consumption. This interpretation also matched the open-ended responses. Users described filters as "not seen that easily" and noted that the information hierarchy was sometimes "off," which supports the idea that visual frictions affected how efficiently they could orient themselves in a dashboard. To illustrate, we may consider an order-management dashboard as an example. In such a dashboard, users may struggle to understand whether top KPI tiles, lateness buckets, drilldown tables, and opportunity lists all reflect the same active filters, or how labels such as "Avg. Days Late" and "Open Deliveries" should be interpreted together. In this situation, an AI

assistant could alleviate the friction by explicitly summarizing the active filter context, explaining how the panels relate to one another, clarifying what a KPI or bucket represents, and answering view-specific questions such as whether the displayed average delay refers to all open deliveries or only the currently filtered subset. However, H01, H02, and H03 all originated from the dashboard's visual layout and label clarity rather than from assistant behaviour. We therefore interpreted AI support as complementary rather than substitutive. Assistants could reduce residual interpretive friction within a given view, but our findings did not support treating them as a replacement for clear layout, labelling, and information hierarchy in the dashboard itself (Ullrich et al., 2024).

Second, the results highlighted that expertise translation was a major source of friction. H29 and H32 showed that the ability to translate questions into the right filters and views aligned with Decision Effectiveness and higher Decision Confidence. These results suggested that users often knew what they wanted to ask but struggled to express that intent through dashboard interaction primitives. This interpretation also aligned with the screening questions, in which finding the right dashboard or page emerged as the most frequently selected and highest-ranked need for AI support. For instance, this friction could arise when a user wanted to identify which customers drove the highest overdue value in the accounts receivable context. It could also arise when the user wanted to determine whether the issue came mainly from disputes, credit, or collections. In other cases, the user may have wanted to understand which filtered subset of data explained the top KPIs shown in the current view. Without guidance, the user may need to infer which panel contained the relevant evidence, manually open several filter dropdowns, and test multiple slices before reaching the correct view. In this context, an AI assistant could reduce that effort by identifying the most relevant panel, opening or pre-populating the relevant filters, highlighting the chart or table that contained any necessary evidence. It could also have shared a direct link to the appropriate drilldown state and explained why a view was relevant. For example, it could have indicated that disputes and overdue days should be examined together because both indicators were already visible in the current dashboard context. Therefore, we interpreted AI assistance in this area as a translation layer between user intent and dashboard operations. Assistants could guide users toward the correct interaction state, but they still depended on the dashboard as the environment in which the evidence, filters, and drill paths remained visible and auditable.

H20 and H23 indicated that decision outcomes are stronger when respondents perceive dashboard data as credible, complete, and up to date. In our sample, higher overall trust aligned with higher decision effectiveness. Additionally, perceived completeness and freshness of dashboards aligned with higher reported usefulness for job performance. These results suggested that users needed more than an answer, they also needed fast access to credibility cues before they acted on that answer. Open-ended responses supported this interpretation by respondent expressions saying that unclear lineage or calculation logic forced them to "guess" meanings and to "cross compare with [external tools] just to make sure the numbers are accurate." This behaviour showed that verification remained part of the decision workflow when credibility remained unclear. Accordingly, we thought extending headless assistants to enable users to query KPIs without opening the dashboard on external tools should be implemented with caution to preserve data credibility. Our results suggested that this pattern risked increasing validation friction due to assistants rarely exposing structured metadata such as definitions, filter context, and refresh status in a way that supports quick auditing. We therefore positioned the dashboard as the primary representation layer for credibility and auditability, and we framed assistant design as complementary. Assistants should surface KPI definitions, active filters, freshness timestamps, and lineage on demand while staying anchored to the dashboard evidence.

Fourth, the results emphasized actionability. H41 showed that specific and appropriately filtered metrics aligned with higher perceived usefulness for job performance. This finding suggested that dashboards became more valuable when they helped users move from observation to next steps. The ranked assistance needs reinforced this point, placing insights into action immediately after finding the right dashboard. Open-ended responses pointed in the same direction, as some users reported that they needed to export results to spreadsheets or perform manual calculations before acting. For example, this friction could arise when a user recognized that overdue deliveries, disputed invoices, or high days-sales-outstanding values required follow-up queries in the current dashboard view, but still had to determine which customer segment, document type, or operational owner should be prioritized first. In this area, an AI assistant could strengthen user experience by summarizing the most decision-relevant slices, highlighting which filtered subset contributed most to the KPI, comparing the current value to target or best-in-class benchmarks, and proposing concrete next steps such as opening the affected

customer list, exporting the relevant case subset, or generating a shareable follow-up summary for operations teams. It could also reduce manual effort by performing simple comparative calculations within the current filter context and by recommending the next drilldown state needed to turn a dashboard observation into an operational action.

Taken together, these mechanisms showed that the findings suggested that AI assistants may create the greatest value when they reduce friction in understanding, querying, verifying, and acting on dashboard information. The broader survey context supported this interpretation. Respondents did not primarily ask for generic automation, but for help in locating evidence, understanding metrics, and operationalizing results. In our sample, AI assistants did not appear most useful as standalone answer engines; they appeared most useful when they remained grounded in the dashboard context and supported users at the points where decision-related friction emerged most clearly.

6 Conclusion

Even in experienced and relatively satisfied user groups, specific dashboard frictions were systematically aligned with decision-related outcomes. The results showed that decision effectiveness depended on a users` ability to interpret visual information, translate analytical intent into system interactions, verify data credibility, and act on decision-relevant metrics. These findings identify four friction areas that future AI assistants may productively target when designed to remain grounded in dashboard context.

Acknowledgment

The authors used AI-based tools to support language editing and to improve the structure and clarity of the paper. The authors reviewed and revised all AI-assisted suggestions and took full responsibility for the final content.

References

Alhamadi, M., Alghamdi, O., Clinch, S., & Vigo, M. (2022). Data quality, mismatched expectations, and moving requirements: The challenges of user-centered dashboard design. *NordiCHI '22: Nordic Human-Computer Interaction Conference*. doi:10.1145/3546155.3546708

Ammann, J., Lohoff, L., Wurm, B., & Hess, T. (2025). How do process mining users act, think, and feel? an explorative study of process mining use patterns. *Business & Information Systems Engineering*. doi:10.1007/s12599-025-00931-9

Ankomah, D. (2025). Evaluating the impact of AI-powered chatbots on user experience [Preprint]. *TechRxiv*. doi:10.36227/techrxiv.173750043.33655505/v1

Asare-Marfo, D. (2021). Why do some open-ended survey questions result in higher item nonresponse rates than others? Pew Research Center. https://www.pewresearch.org/decoded/2021/10/14/why-do-some-open-ended-survey-questions-result-in-higher-item-nonresponse-rates-than-others/

Atighehchian, A., Alidadi, T., Mohammadi, R. R., Lotfi, F., & Ajami, S. (2023). Identifying the application of process mining technique to visualize and manage in the healthcare systems. In Z. Hu, I. Dychka, & M. He (Eds.), *Advances in computer science for engineering and education vi* (Vol. 181, pp. 299-308). Springer, Cham. doi:10.1007/978-3-031-36118-0_26

Barbieri, N., Madeira, E., Stroeh, K., & van der Alst, W. (2023). A natural language querying interface for process mining. *Journal of Intelligent Information Systems*, *61*, 113-142. doi:10.1007/s10844-022-00759-9

Bender, R., & Lange, S. (2001). Adjusting for multiple testing-when and how? *Journal of Clinical Epidemiology*, *54*, 343-349. doi:10.1016/S0895-4356(00)00314-0

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

Bernerth, J. B., & Aguinis, H. (2016). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, *69*(1), 229-283. doi:10.1111/peps.12103

Boareto, P. A., Loures, E. d. F. R., Santos, E. A. P., & Deschamps, F. (2025). Generative assistant for digital twin simulations [12th CIRP Global Web Conference (CIRPE 2024)]. *Procedia CIRP*, *132*, 129-134. doi:10.1016/j.procir.2025.01.022

Boer, L. d. (2024, August). New process AI capabilities for SAP Signavio offer instant, tailored recommendations [Press Release]. https://news.sap.com/2024/08/new-sap-signavio-process-ai-capabilities-instant-tailored-recommendations/

Calvia, P. (2024, November). New AI and copilot capabilities open up new ways businesses interact with SAP Signavio solutions for business transformation [Press Release]. https://news.sap.com/2024/11/joule-ai-copilot-sap-signavio-business-transformation/

Carpenter, B. (2025, October). Unlocking AI success: The critical role of process intelligence in SAP transformations. Mygo Consulting Inc. Retrieved December 24, 2025, from https://mygoconsulting.com/blog/sap-garage/unlocking-ai-success-the-critical-role-of-process-intelligence-in-sap-transformations/

Chinnathai, M. K., & Alkan, B. (2023). A digital life-cycle management framework for sustainable smart manufacturing in energy intensive industries. *Journal of Cleaner Production*, *419*, 138259. doi:10.1016/j.jclepro.2023.138259

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334. doi:10.1007/BF02310555

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340. doi:10.2307/249008

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532-550. doi:10.5465/amr.1989.4308385

Everest Global, Inc. (2025). Process mining state of the market 2025 (State of the Market Report). Everest Group. https://www.everestgrp.com/report/egr-2025-38-r-7549/?utm

Fahland, D., Fournier, F., Limonad, L., Skarbovsky, I., & Swevels, A. J. E. (2025). How well can a large language model explain business processes as perceived by users? *Data & Knowledge Engineering*, *157*, 102416. doi:10.1016/j.datak.2025.102416

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, *16*(1), 15-31. doi:10.1177/1094428112452151

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, p values, confidence intervals, and power: A guide to

misinterpretations. *European Journal of Epidemiology, 31*(4), 337-350. doi:10.1007/s10654-016-0149-3

González-García, J., Estupiñán-Romero, F., Tellería-Orriols, C., González-Galindo, J., Palmieri, L., Faragalli, A., Pristás, I., Vuković, J., Misinš, J., Zile, I., & Bernal-Delgado, E. (2021). Coping with interoperability in the development of a federated research infrastructure: Achievements, challenges and recommendations from the JA-InfAct. *Archives of Public Health*, *79*, 221. doi:10.1186/s13690-021-00731-z

Harati Nik, M., van der Aalst, W. M. P., & Fani Sani, M. (2019). BIpm: Combining BI and process mining. *Proceedings of the 8th International Conference on Data Science, Technology and Applications*, 123-128. doi:10.5220/0007741901230128

Ibanez-Sanchez, G., Castillo, P. A., Juarez, J. M., Fernandez-Llatas, C., Valencia, P., Ruiz-Fernandez, D., & Rojas, I. (2024). Pmapp: An interactive process mining toolkit for building healthcare dashboards. *Explainable AI in Healthcare and Process Mining for Healthcare (XAI-Healthcare/PM4H 2023)*, 2020, 75-86. doi:10.1007/978-3-031-54303-6_8

Ives, B., Olson, M. H., & Baroudi, J. J. (1983). The measurement of user information satisfaction. *Communications of the ACM*, *26*(10), 785-793. doi:10.1145/358413.358430

Kecht, C., Egger, A., Kratsch, W., & Röglinger, M. (2023). Quantifying chatbots' ability to learn business processes. *Information Systems*, *113*, 102176. doi:10.1016/j.is.2023.102176

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, *29*(3), 305-325. doi:10.1016/0304-4076(85)90158-7

Martin, N., Fischer, D. A., Kerpedzhiev, G. D., Goel, K., Leemans, S. J. J., Röglin, M., van der Aalst, W. M. P., Dumas, M., La Rosa, M., & Wynn, M. T. (2021). Opportunities and challenges for process mining in organizations: Results of a delphi study. *Business & Information Systems Engineering*, *63*(5), 511-527. doi:10.1007/s12599-021-00720-0

Martinez-Millana, A., Lizondo, A., Gatta, R., Vera, S., Traver Salcedo, V., & Fernandez-Llatas, C. (2019). Process mining dashboard in operating rooms: Analysis of staff expectations with analytic hierarchy process. *International Journal of Environmental Research and Public Health*, *16*(2), 199. doi:10.3390/ijerph16020199

Meduri, V. V., Quamar, A., Lei, C., Efthymiou, V., & Ozcan, F. (2021). BI-REC: Guided data analysis for conversational business intelligence. *arXiv preprint*. doi:10.48550/arXiv.2105.00467

Miller, A. L., & Lambert, A. D. (2014). Open-ended survey questions: Item nonresponse nightmare or qualitative data dream? *Survey Practice*, *7*(5). doi:10.29115/SP-2014-0024

Molléri, J. S., Petersen, K., & Mendes, E. (2019). An Empirically Evaluated Checklist for Survey Research in Software Engineering. *Information and Software Technology*, *106*, 20-37. doi:10.1016/j.infsof.2019.106240

Morelli, F., Sekulovska, A., & Schätter, F. (2024). Dashboard use case for supply chain resilience management and future research direction. *Communications of the ECMS, Proceedings*, *38*(1). https://www.scseurope.net/dlib/2024/ecms2024acceptedpapers/0416_simo_ecms2024_0039.pdf

Nai, R., Sulis, E., Audrito, D., Trifiletti, V. M. S., Meo, R., & Genga, L. (2025). Leveraging process mining and event log enrichment in european public procurement analysis: A case study. *Computer Law & Security Review*, *57*, 106144. doi:10.1016/j.clsr.2025.106144

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673-690. doi:10.1007/s11135-006-9018-6

Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *arXiv preprint arXiv:1704.04579*. doi:10.48550/arXiv.1704.04579

Rehse, J.-R., Pufahl, L., Grohs, M., & Klein, L.-M. (2022). Process mining meets visual analytics: The case of conformance checking. *arXiv preprint arXiv:2209.09712*. doi:10.48550/arXiv.2209.09712

Sanchez-Obando, J. W., Duque-Méndez, N. D., & Herrera, O. M. B. (2024). Construction of the invoicing process through process mining and business intelligence in the Colombian pharmaceutical sector. *Computers*, *13*(10), 245. doi:10.3390/computers13100245

SAP Signavio. (2023, November). Harnessing the power of generative AI for business process transformation. https://www.signavio.com/events/harnessing-the-power-of-generative-ai-for-business-process-transformation-apj-emea/

SAP Signavio. (2024a). AI interplays with the process world [Ebook]. https://www.signavio.com/downloads/white-papers/e-book-ai-interplays-with-the-process-world/?utm_source=chatgpt.com

SAP Signavio. (2024b). Streamline and optimize processes with sap business AI. https://www.signavio.com/business-ai/

Simic, D., Leible, S., Schmitz, D., Gücük, G., & Kučević, E. (2023). Enhancing project-based learning through data-driven analysis and visualization: A case study. *ACIS 2023 Proceedings*, 139. https://aisel.aisnet.org/acis2023/139

Ullrich, C., Teodora, L., & Benton, G. (2024). Building process mining dashboards manually or AI-assisted with Celonis views. *Proceedings of the International Conference on Process Mining (ICPM) 2024, Tool Demonstration Track*. https://ceur-ws.org/Vol-3783/paper_332.pdf

van der Aalst, W. M. P. (2016). *Process mining: Data science in action* (2nd ed.). Springer. doi:10.1007/978-3-662-49851-4

van der Aalst, W. M. P., Adriansyah, A., van Dongen, B. F., Günther, C. W., Rozinat, A., Rubin, V., Verbeek, H. M. W., Weijters, T., van der Werf, J., & Wynn, M. T. (2012). Process mining manifesto. In *Business process management workshops* (Vol. 99, pp. 169-194). Springer. doi:10.1007/978-3-642-28108-2_19

Velásquez, I., & Sepúlveda, M. (2024). Analyzing the devil's quadrangle of process instances through process mining. In J. De Weerdt & L. Pufahl (Eds.), *Business process management workshops: Bpm 2023* (Vol. 492, pp. 272-284). Springer. doi:10.1007/978-3-031-50974-2_21

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Quarterly*, *37*(1), 21-54. https://www.jstor.org/stable/43825936

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician, 70*(2), 129-133. doi:10.1080/00031305.2016.1154108

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817-838. doi:10.2307/1912934

Wilson, C. (2024, October). Digital transformation unleashed: The cube explores AI-powered process intelligence innovations. Retrieved December 24, 2025, from https://siliconangle.com/2024/10/25/process-intelligence-celosphere/

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, 29-39. https://www.cs.unibo.it/~danilo.montesi/CBD/Beatriz/10.1.1.198.5133.pdf

Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, *16*(1), 85-102. doi:10.1287/isre.1050.0042

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (7th ed.). Cengage Learning.

Zimmermann, L. (2023). Enhancing support for individual analysts in process mining. *Proceedings of the ICPM Doctoral Consortium and Demo Track 2023*, 3648. https://ceur-ws.org/Vol-3648/paper_7878.pdf

Zuidema-Tempel, E., Effing, R., & van Hillegersberg, J. (2022). Bridging the gap between process mining methodologies and process mining practices: Comparing existing process mining methodologies with process mining practices at local governments and consultancy firms in the Netherlands. In C. D. Ciccio et al. (Eds.), *Business process management: Bpm 2022 industry track* (Vol. 458, pp. 70-86). Springer. doi:10.1007/978-3-031-16171-1_5

7 Appendix

7.1 Recruitment Procedure

The target population involved users of dashboards on process intelligence platforms. We recruited participants through a vendor-agnostic non-probability sampling approach that combined convenience and snowball sampling. First, we approached individuals in the authors` networks who were known to use dashboards on process intelligence platforms. After confirming their eligibility and willingness to participate, we asked them to refer other eligible users in their own networks. Because this network-based recruitment did not rely on a defined sampling frame, the sample should not be interpreted as statistically representative of all PI dashboard users. We screened all respondents to ensure that they currently use a dashboard on a process intelligence platform before including them in the study. We recorded 63 eligible responses via Qualtrics, of which 58 complete responses entered the regression analyses after excluding five partial completions.

7.2 Survey Structure and Measurement Blocks

Appendix Table 2 summarizes the survey structure, measured constructs, variable codes, and survey blocks.

Table 2: Survey Structure and Measurement Blocks

Construct/ Measure	Variable codes	Short Description	Section/Block
Informed Consent	CONSENT1-CONSENT3	Eligibility (18+) and consent to participate (incl. data handling).	Informed Consent
Screeners & demographics	SCR1; PURPOSE; ROLE; EXP_YEARS; FREQ; NPS; AI_AVAIL	Dashboard use inclusion criterion followed by usage context, role/experience, frequency, advocacy, and AI availability/use.	Screeners & Demographics
Dashboard Anchor	DASH_PICK; DASH_DESC	Select one dashboard actually used and briefly describe its use case to contextualize subsequent ratings.	Dashboard Anchor
Visual Comprehension Issues	VSC1, VSC2, VSC3, VSC4	Clarity of layout, labels, time context and salience via color/emphasis.	Visual Clarity & Findability

Construct/ Measure	Variable codes	Short Description	Section/Block
Navigation difficulty	NAV1, NAV2	Speed to the right page/detail; consistency of filters/menus.	Visual Clarity & Findability
Domain language misalignment	DLM1	Labels/metrics aligned to team language; clarity of codes/IDs.	Trust & Data Credibility
Expectation misalignment	EXM1, EXM2, EXM3	Alignment of KPIs/detail/process view and responsiveness to feedback.	Trust & Data Credibility
Trust issues	TRI1, TRI2, TRI3	Lineage/definitions; completeness/freshness; correct timestamp handling.	Trust & Data Credibility
Fear of surveillance	PRV1	Context and fairness of comparisons.	Trust & Data Credibility
Expertise requirement	EXP1, EXP2	Ability to use the dashboard without specialized technical skills and to translate questions into filters/views without help.	Skills, Guidance & Getting to Action
Lack of user understanding	LKU1	Knowledge of steps/terms/KPIs needed to interpret the dashboard correctly.	Skills, Guidance & Getting to Action
Need of technical support	NTS1	In-tool help/onboarding sufficient to proceed unaided.	Skills, Guidance & Getting to Action
Lack of actionability	ACT1, ACT2	Update cadence supports timely decisions; metrics are specific/filtered enough to guide next steps.	Skills, Guidance & Getting to Action
Manual effort requirement	MER1, MER2, MER3	Thresholds/filters can be set quickly without trial-and-error; routine questions can be answered in-tool without exporting.	Skills, Guidance & Getting to Action
Decision effectiveness	DE1, DE2	Timeliness and confidence of decisions made with the dashboard.	Outcomes
Perceived usefulness	PU1, PU2	Contribution to job performance and role effectiveness.	Outcomes
Perceived ease of use	PEOU1	Ease to learn/operate the dashboard.	Outcomes
Information satisfaction	IS1	Sufficiency/fit of dashboard information for the job.	Outcomes
AI task needs	AIA_TASKS	Which user tasks most need AI help (multi-select).	AI Assistant Feasibility
AI task priorities	AIA_RANK	Rank selected tasks from most to least helpful.	AI Assistant Feasibility

Source: Own

7.3 Reliability Diagnostics

Table 3: Internal consistency (Cronbach's α) and operationalization decisions for multi-item constructs

Construct	Items	k	α	Operationalization in analysis
Visual Comprehension	VSC1, VSC2, VSC3, VSC4	4	0.675	Used as sub-themes: VSC1 → VisualLayoutLabelClarity; VSC2 → VisualTimeContextTrendReadability; VSC3 → VisualEmphasisColorEncoding; VSC4 → VisualGlanceabilityHierarchy
Navigation Difficulty	NAV1, NAV2	2	0.684	Used as sub-themes: NAV1 → NavigationReachDrill; NAV2 → NavigationFilterMenuConsistency
Expectation Misalignment	EXM1, EXM2, EXM3	3	0.652	Used as sub-themes: EXM1 → ExpectationRelevanceGoals; EXM2 → ExpectationProcessFit; EXM3 → ExpectationFeedbackIteration
Trust Issues	TRI1, TRI2, TRI3	3	0.710	Kept as composite: TrustIssues = mean(TRI1, TRI2, TRI3)
Expertise Requirement	EXP1, EXP2	2	0.589	Used as sub-themes: EXP1 → ExpertiseTechnicalIndependence; EXP2 → ExpertiseQuestionQueryTranslation
Lack of actionability	ACT1, ACT2	2	0.626	Used as sub-themes: ACT1 → ActionabilityTimeliness; ACT2 → ActionabilityMetrics
Manual effort requirement	MER1, MER2, MER3	3	0.501	Used as sub-themes: MER1 → ManualEffortConfiguration; MER2 → ManualEffortNoExport; MER3 → ManualEffortCollaboration
Decision effectiveness	DE1, DE2	2	0.811	Kept as composite: DecisionEffectiveness = mean(DE1, DE2)
Perceived usefulness	PU1, PU2	2	0.874	Kept as composite: PerceivedUsefulness = mean(PU1, PU2)

Source: Own

In general, Cronbach's Alpha values are commonly used as a minimum threshold for acceptable internal consistency (Nunnally & Bernstein, 1994). In this study, we used the common rule of thumb for interpreting alpha:

- ↑ 0.90 – excellent
- 0.80–0.89 – good
- 0.70–0.79 – acceptable

- < 0.70 – questionable or poor (items treated separately) (Nunnally & Bernstein, 1994).

Below shown all constructs measured with multiple items in the survey and their Cronbach's alpha values in the collected sample. Single item constructs like Fear of Surveillance, Need of Technical Support, PEOU, and IS; screeners and demographics; and free-text questions are not shown here because construct reliability check was not applicable to these constructs. Based on the operationalization decisions made in this step, constructs kept and referred as either singletons or composites.

Being There: Using Combined Experience Data Sources to Capture Tourist Interactions

Ondrej Mitas,[1] Loïs Camp,[1] Ruben Raidt,[2] Kyle Woosnam,[3] Joël Pardaan,[1] Simone Moretti,[1] Jeroen Klijs,[1] Sait Durgun[1]

[1] Breda University of Applied Sciences, Breda, the Netherlands
mitas.o@buas.nl, loiscamp53@gmail.com, pardaan.j@buas.nl, moretti.s@buas.nl, klijs.j@buas.nl, durgun.s@buas.nl
[2] Academiehuis, Zwolle, the Netherlands
rubenraidt@academiehuis.nl
[3] University of Geogria, Warnell School of Forestry and Natural Resources, Athens, Georgia, United States of America
woosnam@uga.edu

DOI
https://doi.org/10.18690/um.fov.4.2026.47

ISBN
978-961-299-152-4

Tourist-resident interactions are central to tourism experience yet notoriously difficult to capture in naturalistic settings. Drawing on Collins' interaction ritual theory, this study refines its core construct of "emotional energy" by redefining it as emotional arousal-a momentary psychophysiological response that accrues, over repeated interactions, into the more enduring attitude of emotional solidarity. We tested this two-level theoretical model using an innovative combination of continuous skin conductance recording, 360-degree camera monitoring, and questionnaires with 115 participants (49 tourists, 66 residents) across two Dutch tourism experiences designed to facilitate tourist-resident encounters. Results confirmed that interactions triggered elevated emotional arousal at the within-person level, consistent with interaction ritual theory. However, between-person effects on emotional solidarity, destination recommendation, and support for tourism were not observed, likely because interactions were remarkably few and brief. We discuss implications for tourism experience design and the methodological potential of combined psychophysiological and video-based field measurement.

Keywords:
social interaction, 360 degree video, skin conductance, psychophysiology, emotion, tourism

University of Maribor Press

1 Introduction

Interactions between tourists and local residents are inherent to tourism. When people visit a place for leisure purposes, they usually encounter other people going about daily life in the same spaces. Tourism scholars have thoroughly documented that these encounters can give tourist experiences tremendous value, ranging from memorable experiences (Jiang et al., 2022) to lifelong friendships (Cai et al., 2025). Scholars have also extensively explored the negative aspects of interactions between tourists and residents. Negative resident-tourist interactions are a manifestation of 'overtourism,' a situation of locally excessive tourist flows that leads to resident resistance (Gössling et al., 2020). Thus, explaining how these interactions and their consequences unfold is an urgent issue for tourism research. Joo et al. (2020) proposed an adaptation of the Collins (2014) interaction ritual theory to explain why tourists and residents interact, and how these interactions lead to downstream behavior.

Unfortunately, it has proved challenging for researchers to systematically record interactions between individuals in natural settings, especially such dynamic settings as tourism experiences. In the present study, we bring an innovative technological development to bear on this issue, namely the combination of continuous skin conductance recording with 360-degree camera monitoring. While these technologies are not new, their novel combination allows us to precisely isolate psychophysiological responses to social interactions.

Applying this powerful technological combination to the issue of tourist-resident interactions, we undertake a multi-method, multilevel analysis to measure and explain interaction outcomes. First, we review how interaction ritual theory has been adapted to tourism and attempt to sharpen its conceptual precision, making a distinction between *emotional arousal* and *emotional solidarity*, arguing that *arousal* is a momentary physiological response which, in sufficient levels and frequency, might lead to *solidarity*, an enduring affective attitude, over time. Then, we test the effects of interactions on a between-person level, focusing on experience evaluation and support for tourism as desirable outcomes of interactions in the context of tourism. Finally, we examine the extent to which interactions indeed contribute to emotional arousal on a within-person, moment-to-moment level.

2 Literature review

2.1 Interaction rituals

In his influential book about interaction rituals, Randall Collins (1981, 2014) presented a remarkably clear set of psychological hypotheses about how social interactions unfold between people. According to Collins, interactions are ritualized at the societal level and thus follow a predictable sequence of mechanisms. First, multiple people find themselves in the same physical space. Second, they focus their attention on one and the same object. Third, they share a common or comparable emotional reaction to the object. Fourth, as the interaction progresses, both attention and emotion amplify. Finally, the members of the interaction begin to identify with one another, as members of a small and temporary in-group (Collins, 1989).

Unfortunately, Collins did not interface with the discipline of emotion psychology (Barrett, 2017; Ekman et al., 1983; Frijda, 2017; LeDoux, 1998; Russell, 1978), which at that time already had much to contribute to individual-level analyses of social interactions. Collins also faced obvious measurement challenges with fleeting social, affective, and cognitive phenomena as they occurred in naturalistic settings. As interaction rituals are crucial to understanding tourism (Joo et al., 2020), tourism scholars have tried to address these methodological challenges, at first tangentially as explain in the following section.

2.2 Interaction rituals in tourism

Joo et al. (2020) suggested that the construct of emotional solidarity may replace emotional energy in Collins' theory for methodological purposes. They claimed that "although doing so may not be an ultimate solution, it can be an effective means of circumnavigating the conceptual issue associated with emotional energy and the interpretative challenges associated with understanding physiological data" (p. 10). In the present paper, we address these issues and challenges to advance tourism knowledge.

The "conceptual issue associated with emotional energy" (p. 10) that Joo and colleagues touch upon is that emotional energy is only vaguely defined in Collins (2014), even by his own admission. To sharpen this definition requires looking across disciplines to emotion psychology. In Collins' book, emotional energy is apparently a reaction to social stimuli that hold personal meaning, a definition rather similar to emotion in general (Rosenberg, 1998), though specifically in a social context. Collins' emotional energy is something that emerges in interactions between people, rather than within an individual, however. Thus, besides emotion in general, Collins' conceptualization touches on the construct of emotional contagion. Emotional contagion is the process whereby an emotional expression in one person triggers a similar reaction in another (Mitas et al., 2022). Further adding to the ambiguity, Collins' model posits emotional energy as a necessary condition for an interaction to continue or be repeated, a reaction associated with positive emotions in particular, whereas Joo et al. (2020) make a distinction between positive and negative emotional energy. We choose to approach emotional energy at the individual rather than dyadic level, that is, conceptualizing emotional energy as something *one person* feels as a result of interacting with others, rather than something that connects one person's emotions to that of another's. Furthermore, we follow Joo et al. (2020) in understanding emotional energy as something with no inherent valence. That is, emotional energy may have a positive or negative valence.

Applying the well-established valence-arousal dimensional model of emotion (Barrett, 2017; Posner et al., 2005; Russell, 1978) to these conceptual clues, we posit that *emotional energy* as described by Collins (2014) and Joo et al. (2020) should be redefined as the psychophysiological construct of *emotional arousal*. The word 'energy' is often defined as vigor or force, whereas arousal represents the intensity with which an emotion is experienced, and thus the extent to which it invigorates an individual to react with behavior. In both interaction ritual theory and in the dimensional model of emotion, energy and arousal respectively are akin to intensity. Thus, Collins' choice to use the term 'emotional energy' evokes the function of emotional arousal in the dimensonal model of emotion. Furthermore, just as Joo et al. (2020) describe emotional energy as being possibly either positive or negative, the dimensional model of emotion positions arousal as independent from valence. The dimension of valance refers to how positive or negative an emotion is, whereas arousal refers to intensity, independent of valence (Barrett, 2017; Posner et al., 2005; Russell, 1978). Finally, in both Collins' description of emotional energy and in empirical research

on how the brain constructs emotion (Barrett, 2017), emotional arousal or energy *precede* any conscious thought about feeling as well as any behavioral reaction. In conclusion, we adopt an operational definition of *emotional energy* in Collins' model as *emotional arousal*—the intensity of emotional experience that can be either positive or negative, and largely determines whether an emotional experience will be encoded in memory and lead to downstream behavior (Bastiaansen et al., 2019).

Joo et al. (2020) then suggest the concept of emotional solidarity as a useful proxy for emotional energy, as it is sometimes described by Collins "with little distinction from emotional solidarity" (p. 7). If emotional energy is delineated as arousal, however, a substantial distinction from emotional solidarity emerges. Emotional solidarity as conceptualized and operationalized by Woosnam and Norman (2010) is not a momentary experience, but rather an enduring attitude. Items in their Emotional Solidarity Scale measure two types of psychological phenomena: enduring patterns of past behavior, such as "I have made friends with some visitors in [destination]," and enduring emotionally-driven attitudes that can change over longer periods of time, but not moment to moment, such as "I feel affection toward visitors in [destination]" (p. 375). Accordingly, studies using the scale have been exclusively between-individual in design. Rather than using a relatively more temporally stable construct (emotional solidarity) as a between-individual proxy of a temporally unstable construct (emotional energy), we hypothesize that transient experiences of emotional arousal could potentially accrue to emotional solidarity. Thus, we divide the theoretical model posited by Joo et al. (2020) into a within-person level, wherein interactions might affect emotional arousal, and a between-person-level, wherein interactions are related to emotional solidarity. Furthermore, we expect interactions that contribute to emotional solidarity on a between-person level to feature higher emotional energy on a within-person level.

2.3 Hypotheses

We posit several between-person hypotheses (Figure 1 in blue):

- H1. The number and total duration of interactions between residents and tourists is associated with intent to recommend the destination.
- H2. The number and total duration of interactions between residents and tourists is associated with support for tourism among residents.

And we also posit several within-person (in interaction with between-person effect) hypotheses (Figure 1 in green):

- H3. Interactions between tourists and local inhabitants trigger emotional elevated emotional arousal.
- H4. The effect of interactions between tourists and local inhabitants on emotional arousal is accentuated among individuals with higher levels of tourist-inhabitant emotional solidarity.

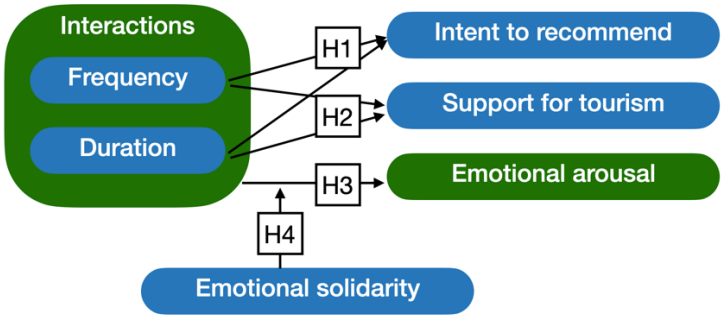

Figure 1. Study hypotheses
Source: Own

These hypotheses have been technically difficult to test until recently. Thus, besides extending theory on interaction rituals in a tourism context, our study represents a methodological contribution in experience measurement technology. Interactions between people, especially strangers, have been difficult to capture because they tend to be rather fleeting. As such, there have been few serious attempts to isolate continuous measures of experiences based on discrete epochs of social interaction, at least outside of laboratory settings. However, the combination of 360-degree cameras and wearables which measure physiological arousal make isolating individual reactions possible. The accessibility of these technologies allowed us to not only address hypotheses raised by interaction ritual theory, but to demonstrate how it is technically feasible to isolate emotional responses to interactions between people.

3 Methods

The present study used a multi-method longitudinal approach to measure tourist-resident interactions at two locations in the early summer of 2022. Two tourist experiences designed to facilitate tourist-resident interactions were selected on basis of theoretical relevance and availability to measure participants' experiences in detail. Then, participants—who comprised 49 tourists and 66 residents—filled in questionnaires before and after the experience, and were continuously monitored for interaction and emotional arousal using 360-degree camera and wearable recordings. Thus, three methods are brought together in our study: self-response questionnaires, continuous physiological recording, and camera-based continuous behavior monitoring. While each method has limitations, the three methods are brought together to create an unusally complete and temporally detailed assessment of social interactions and their consequences during these experiences.

3.1 Study contexts

We selected two experiences which were designed for and consumed by roughly equal proportions of residents and tourists. Each of these experiences highlights a specific attraction from a mid-size city in the Netherlands, and each has a stated, secondary goal of facilitating resident-tourist interactions. St. Michael's Church in Zwolle is operated as the Academiehuis (Academy House), offering a self-guided visit of varying cultural products, including Gothic church interior architecture, an art gallery, a secondhand bookshop, a café, and recorded as well as live music. The Blind Walls Gallery tours in Breda offer an urban walk experience focused on a selection of the city's over 100 (in 2022; now over 200) contemporary murals. The murals are commissioned from internationally renowned painters and graphic artists and express site-specific stories from the city's history. They are guided by knowledgeable tour guides, often students or recent graduates of the local art school.

3.2 Sample

Due to the different structures of the two experiences studied (self- vs. professional-guided) we also approached sampling differently. At the Academiehuis, we used intercept sampling, choosing peak hours on weekend days which were likely to be busy, to maximize both the sampling frame of potential participants and to capture

maximum possible interactions within each participant. At the Blind Walls Gallery tours, we arranged tour dates and times at moments when guides were available, then used a combination of convenience and snowball sampling to recruit participants from researchers' personal and professional social networks. This sampling approach resulted in 74 participants recruited over three weekends at the Academiehuis, and 45 participants recruited over the course of five guided tours.

The distinction between tourists and residents was important to study hypotheses. Unfortunately, this distinction is challenging in both theory and practice. To simplify matters, we included both overnight and day visitors in the tourist category, while residents were defined as individuals self-reporting long-term (e.g., permanent or higher education) residence in the same city as the attraction.

3.3 Measures

We combined multiple measurement streams by participant ID and time to create a multilevel dataset suitable for testing our hypotheses. Key between-person variables of emotional solidarity, support for tourism, and intent to recommend were measured on a questionnaire administered immediately post-visit using established and validated items from (Woosnam & Norman, 2010) and Reichheld (2003), respectively. Emotional arousal and social interaction were measured within each individual, continuously throughout the visit. To record social interaction, we recorded video using a 360 degree camera. Placed centrally on a tripod within the Academiehuis, this allowed us to record all visible behavior throughout the data collection period simultaneously. During Blind Walls Gallery tours, the camera was instead mounted on a monopod which was inserted into a backpack worn by a research assistant who followed all the tours. This approach allowed us to record all visible behavior simultaneously. Social interactions were coded with second-accurate time stamps by trained research assistants.

We recorded skin conductance as an index of emotional arousal from each participant using Empatica E4 wearable wristbands and external pre-gelled single-use electrodes attached to the inner medial phalanges of the first and second digits. The inner (palm) side of the digits contains sufficient eccrine sweat glands to give a relatively high-quality signal in skin conductance changes over the course of an experience (Braithwaite et al., 2015). Any potential motion artifacts in the raw data

were highlighted using the ArtifactZ algorithm (Bastiaansen et al., 2022), checked by a trained member of the research team (2nd author), and replaced by linear interpolation. Then, the skin conductance signal was separated into its tonic component, which represents gradual change due to ambient temperature and wearing of the device, and its phasic component, which contains rapid changes due to emotional arousal. Finally, we z-standardized the phasic component within each participant to eliminate between-participants differences in skin conductance reactivity to physiological arousal, and log-transformed it to reduce skew.

3.4 Data analysis

Our analysis approach included three steps. First, we computed descriptive statistics on all variables to get a general impression of the experiences and opinions of our sample. Second, we used linear regressions to test between-individual hypotheses. Finally, to test within-individual hypotheses, we used mixed-effects linear models with random intercepts per participant, testing coefficients by Satterthwaite approximation.

4 Findings

4.1 Descriptive findings

Most participants (61 of 82 who provided complete data), including all residents (n=28), were not seen in tourist-resident interactions. Among tourist participants, 13 engaged in one interaction with residents, while eight engaged in more than one (max = 8 interactions). For those who did interact, they spent 7.25 to 239 seconds interacting (mean = 20.15 seconds, sd = 49.45 seconds, skew = 2.78). The low rate and duration of interactions was remarkable, and must be kept in mind while interpreting further results.

Participants felt generally positive about the city destinations where the research took place, with mean intent to recommend of 8.13 (sd = 1.18) on a 0-to-10 scale. Residents were moderately positive in their support for tourism, above the scale midpoint of 3 (mean = 3.79, sd = 0.46). Relatively strong support for tourism was consistent with both Breda and Zwolle being midsize cities where any crowding situations include substantial proportions of local residents. Moderate emotional

solidarity with the other was reported by both tourists (mean = 3.21, sd = 0.42) and local residents (mean = 3.38, sd = 0.58).

4.2 Hypothesis testing

Our first hypothesis was that interactions between residents and tourists are associated with intent to recommend the destination (city). We modelled this using a linear regression model, with intent to recommend the city as the dependent variable, and number of continuous runs of data points in interaction between tourists and locals as well as total duration of those points in seconds as predictors. For the 82 participants with complete data for these variables, there was no effect of interacting on intent to recommend, with an R-square of less than 1% and non-significant F test for the model overall (F = 0.2435 (2,79), p = 0.7844). This hypothesis was therefore rejected.

Table 1: Linear regression of intent to recommend on frequency and duration of interaction between tourists and resident

	Coefficient estimate	SE	t
(intercept)	8.1032	0.1431	56.640***
frequency	0.0944	0.1833	0.515
duration	-0.0012	0.0055	-0.219

R-square: 0.006; p values: *** <0.001< ** <0.01< * 0.05

Our second hypothesis was that interactions with tourists are associated with support for tourism among the residents in the sample. Because no residents interacted with tourists, this hypothesis could not be tested and was therefore not supported.

Our third hypothesis was that emotional arousal would be triggered by moments of social interaction between tourists and residents. We modelled this hypothesis using the full (four times per second) skin conductance data as our dependent variable. For the predictor, we used a dummy coded 1 if the given participant was in a tourist-resident interaction at that moment, else 0. A random intercept per participant was estimated. This model (Table 2) revealed a very weak, significant effect of tourist-resident interactions on skin conductance (coefficient = 0.0329, se = 0.0014, p < 0.001; marginal-R-square = 0.001). Adding a second dummy variable, representing tourist-tourist and resident-resident interactions revealed an almost identical effect

(coefficient = 0.0333, se = 0.0010, p < 0.001; marginal-R-square with both predictors = 0.002). Therefore, H3 was partly supported, in that interacting was more emotionally arousing than not interacting, but interactions within only tourists or only residents were essentially as arousing as those between tourists and residents.

Table 2: Linear mixed effects model of emotional arousal on interactions between tourists and residents

	Coefficient estimate	SE	T estimate
(intercept)	0.14205	0.0083	17.08***
Tourist-resident interaction	0.0329	0.0014	22.95***
Tourist-tourist or resident-resident interaction	0.0333	0.0010	32.18***

Conditional R-square estimate: 0.308; Marginal R-square estimate: 0.002; p values: *** <0.001< ** <0.01< * 0.05

Our fourth and final hypothesis was that the effect of social interactions on skin conductance would be reinforced by emotional solidarity, leading to an interaction effect. We entered emotional solidarity as an interaction term to the mixed effects model of skin conductance as a function of tourist-resident interactions. As two participants did not provide emotional solidarity data, this model was estimated based on 80 participants (Table 3). While this addition substantially improved the marginal-R-square estimate (from 0.001 to 0.009), complexity-penalized model fit was marginally worse (BIC from -919763.3 to -907056.8) and neither the interaction term nor the zero-order effect of emotion solidarity was significant. Therefore, our hypothesis of emotional solidarity as an accelerant of emotional arousal in social interaction was not supported.

Table 3: Linear mixed effects model of emotional arousal on interactions between tourists and residents moderated by emotional solidarity

	Coefficient estimate	SE	T estimate
(intercept)	0.2178	0.0574	3.794***
Tourist-resident interaction	0.0493	0.0110	4.468***
Emotional solidarity	-0.0231	0.0173	-1.329
Tourist-resident interaction X Emotional solidarity	-0.0051	0.0034	-1.488

Conditional R-square estimate: 0.308; Marginal R-square estimate: 0.009; p values: *** <0.001< ** <0.01< * 0.05

5 Discussion and conclusions

We studied two tourist experiences aimed at fostering interactions between tourists and local residents through the lens of Collins' interaction ritual theory, which posits that interactions trigger emotional 'energy.' We extended this theory to a greater degree of conceptual precision using contemporary psychological understanding of emotion. We specified that social interactions trigger momentary peaks in emotional arousal, which might build solidarity over repeated exposure and eventually accrue to lasting social bonds. We found support for the theoretical proposition that tourists and local residents find social interactions emotionally arousing, a novel contribution to the academic discussion around interaction rituals.

No between-person differences in emotional solidarity, support for tourism, or intent to recommend the destination were linked to interactions between tourists and residents, however. This is likely because the studied experiences triggered remarkably few, very brief interactions between tourists and residents. None of the local residents studied and merely 39% of tourists interacted with the other at all, and not a single participant spent as much as 5 minutes interacting, despite one of the studied experiences (Blind Walls Gallery tour) necessarily lasting over an hour. Therefore, rather than rejecting possibility of tourist-resident interactions contributing to destination evaluation, quality, and support for tourism, we rather posit that the experiences we studied were not *yet* sufficiently centred on tourist-resident interaction to fully test Collins' theory.

Conditions have changed since the data were collected in 2022. Perhaps most importantly, in 2022, the last social distancing restrictions of the Covid-19 pandemic were still being phased out in the Netherlands. Thus, people's habitual behaviors, including their physical closeness to one another, were still somewhat affected. That is no longer the case today. Over the same period, the Blind Walls Gallery has doubled in the number of murals. The gallery has also increased the number of guided tours, and tourism to Breda has increased generally. These developments could facilitate interactions between tourists and residents. At the Academiehuis, the visitor experience has evolved even more substantially, with additions such as a more extensive shop, guided tours, and children's routes. Reconstruction of the church, which was ongoing in 2022, has also been completed. These elements more actively structure visitor presence and potentially increase opportunities for interaction.

Thus, future research should re-examine both the Blind Walls Gallery and Academiehuis to determine to what extent these strategic developments have been effective. Also, now that the Academiehuis offers guided tours, a 2x2 design including both guided and self-guided experience in each location could yield additional experience design insights. It would also be possible to broaden the research further still to contexts where longer-lasting interactions might be structurally more prevalent (i.e., festivals, guided outdoor recreation and adventure, and homestays). Finally, a different sampling approach could target populations, such as backpackers or solo tourists, which are presumed to be more interested in interacting with strangers.

Besides contributing to the academic discussion around interaction rituals in tourism, our study showcases technical solutions to the dilemma of studying social interactions in the field. The valid, reliable, and informationally dense data show that the collected data streams-skin conductance, 360-degree video, and questionnaires-can be readily and accessibly combined to capture social interactions between tourists. There is a 'weak link' in the system for future technological development to address, however, which is the identification and coding of social interaction from video.

At the time the data were collected and coded (2022), broadly available algorithms were not yet sufficient to identify specific participants from 360 degree video and code their interactions. Recent developments in artificial intelligence are making video processing more efficient on an almost-daily basis. The video data we struggled with in 2022 could now be coded with partial automation based on supervised machine learning. These technical advances would support research on social interactions by allowing for larger samples and potentially richer data.

In sum, our findings give a practical as well as methodological snapshot in time. Collected in spring of 2022, at the tail end of the Covid-19 pandemic, we showed that in this context, two novel tourism experiences triggered relatively few, rather brief interactions between tourists and residents. Consistent with Randall Collins' theory of ritual interactions, these interactions triggered emotional arousal, but were not protracted enough to build emotional solidarity or positive evaluations of destinations. Furthermore, we demonstrated that interactions are now technically accessible using an innovative combination of skin conductance wearable recording, 360 degree video, location tracking, and mobile questionnaires, albeit with manual

coding of video. Given current technological developments in these areas, we hope our snapshot from 2022 inspires other tourism, psychology, and sociology scholars to pursue the previously-difficult domain of social interaction research.

References

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

Bastiaansen, M., Lub, X., Mitas, O., Jung, T. H., Passos Acenção, M., Han, D., Moilanen, T., Smit, B., & Strijbosch, W. (2019). Emotions as core building blocks of an experience. *International Journal of Contemporary Hospitality Management*, *31*, 651-668.

Bastiaansen, M., Oosterholt, M., Mitas, O., Han, D., & Lub, X. (2022). An emotional roller coaster: Electrophysiological evidence of emotional engagement during a roller-coaster ride with virtual reality add-on. *Journal of hospitality & tourism research*, *46*(1), 29-54.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2015). *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments*. U. o. B. Behavioural Brain Sciences Centre, UK.

Cai, J., Wu, M.-Y., Wang, L., Ye, S., Chan, J. H., & Li, Q. (2025). 'My host is my buddy': revisiting guests' emotional solidarity with hosts in Peer-to-Peer(P2P) accommodation. *Current Issues in Tourism*, 1-18. https://doi.org/10.1080/13683500.2025.2512003

Collins, R. (1981). On the microfoundations of macrosociology. *American journal of sociology*, *86*(5), 984-1014.

Collins, R. (1989). Toward a neo-Meadian sociology of mind. *Symbolic Interaction*, *12*(1), 1-32.

Collins, R. (2014). Interaction ritual chains. In *Interaction Ritual Chains*. Princeton university press.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions.

Frijda, N. H. (2017). *The laws of emotion*. Psychology Press.

Gössling, S., McCabe, S., & Chen, N. C. (2020). A socio-psychological conceptualisation of overtourism. *Annals of tourism research*, *84*, 102976. https://doi.org/https://doi.org/10.1016/j.annals.2020.102976

Jiang, L., Eck, T., & An, S. (2022). A Study on the Effect of Emotional Solidarity on Memorable Tourism Experience and Destination Loyalty in Volunteer Tourism. *Sage Open*, *12*(1), 1-19. https://doi.org/10.1177/21582440221087263

Joo, D., Cho, H., Woosnam, K. M., & Suess, C. (2020). Re-theorizing social emotions in tourism: Applying the theory of interaction ritual in tourism research. *Journal of Sustainable Tourism*, 1-16.

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster.

Mitas, O., Bastiaansen, M., & Boode, W. (2022). If you're happy, I'm happy: Emotion contagion at a tourist information center. In *Planning and managing the experience economy in tourism* (pp. 122-140). IGI Global.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, *17*(3), 715-734.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, *81*(12), 46-55.

Rosenberg, E. L. (1998). Levels of analysis and the organization of affect. *Review of general psychology*, *2*(3), 247-270. https://doi.org/10.1037/1089-2680.2.3.247

Russell, J. A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of personality and social psychology*, *36*(10), 1152.

Woosnam, K. M., & Norman, W. C. (2010). Measuring residents' emotional solidarity with tourists: Scale development of Durkheim's theoretical constructs. *Journal of Travel Research*, *49*(3), 365-380.

Process Model-guided Post-filtering for Wearable Activity Recognition in Ergonomic Work Processes

Andreas Emrich,[1] Alessa Wein,[1]
Janaki Viswanathan,[2] Michael Frey,[3] Peter Loos[1, 4]

[1] German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany
andreas.emrich@dfki.de, alessa.wein@dfki.de, peter.loos@dfki.de
[2] Department of Language Science & Technology, Saarland University, Germany
jviswanathan@lsv.uni-saarland.de
[3] TRON – Translationale Onkologie an der Universitätsmedizin der Johannes Gutenberg-Universität Mainz gemeinnützige GmbH, Mainz, Germany
michael.frey@tron-mainz.de
[4] Saarland University, Saarbrücken, Germany
peter.loos@dfki.de

Ergonomic assessment of manual work processes requires continuous recognition of workers' physical activities – a task complicated by privacy constraints, sensor complexity, and noisy movement data. This paper proposes a process model-guided post-filtering approach that leverages existing process knowledge to refine the output of wearable activity classifiers, regardless of the underlying recognition method. Evaluated across two industry-derived use cases using body area networks with 3, 5, and 7 sensors, the approach yields dramatic accuracy improvements – from below 50 % unfiltered to over 82 % with as few as three sensors. Results demonstrate that explicit process model knowledge can substantially compensate for reduced sensor setups, lowering hardware costs and privacy risks without sacrificing recognition quality. This proof-of-concept establishes process model filtering as a promising component in ergonomic monitoring pipelines for structured manual work.

DOI
https://doi.org/10.18690/um.fov.4.2026.48

ISBN
978-961-299-152-4

Keywords:
process model filtering, human activity recognition, wearable sensing, body area network, ergonomic assessment, work process monitoring

University of Maribor Press

1 Introduction

Ergonomic assessment in the workplace aims to improve worker health and sustainably enhance productivity and performance. In this context, activity detection for personalized recommendations can help play a key role in preventing injuries and reducing physical strain. While many industrial processes are already automated and optimized using machine data for process mining, a considerable share of work is still performed manually by human operators.

Process mining has emerged as a powerful technique for modeling, analyzing, and improving business processes. In automated environments, processes are typically IT-enabled and equipped with sensors, allowing for the generation of structured system logs. These logs provide a direct source of event data, which serves as a foundation for process mining and performance optimization. For manual human tasks it is not as straightforward.

1.1 Problem Statement

Wearable activity recognition supports health self-monitoring and prevention, providing early warnings or accident alerts (Assenza et al., 2020). However, applying process mining to human-centered work presents significant challenges. Unlike machines, human activities are complex and do not naturally generate the structured, semantic event logs central to classical process mining. Instead, human-generated data is often noisy and heterogeneous, requiring extensive preprocessing to approximate event structures.

Furthermore, designing reliable detection algorithms and selecting sensors are non-trivial tasks. While expressive sensors increase accuracy, they also raise costs, complexity, and training effort. Privacy is an additional concern, as movement data can identify workers and assess performance, which is a particular sensitive issue in healthcare (cf. Munoz-Gama et al., 2022). Consequently, following the EU-GDPR principle of data minimization is essential (cf. GDPR (Regulation (EU) 2016/679)). Ultimately, human-centered process mining deviates from "happy path" scenarios, necessitating novel approaches to data acquisition, preparation, and interpretation.

1.2 Goal

This paper proposes a rule-based post-processing approach that leverages existing process models to refine activity recognition results. The underlying Human Activity Recognition (HAR) classifier is treated as an exchangeable component – the contribution lies in the filtering stage and its integration with process model knowledge. To validate this, two industrial use cases are evaluated across three sensor configurations within a Body Area Network (BAN), demonstrating that filtering substantially compensates for reduced sensor setups.

1.3 Methodology & Structure

Following the Design Science Research (DSR) methodology, the goal is to develop an IT artifact that solves a real-world problem. Different sensor setups are compared using the "Design as a search process" guideline by Hevner et al. (2004). In accordance with Peffer's DSR process (cf. Peffers et al., 2007), the problem at hand is analyzed in this introduction and in Section 2. A motivational use case subsequently serves to refine the solution's objectives (cf. Section 3). Sections 4 and 5 explicate the artifact's design, detailing the system architecture and the specific activity detection methods. Finally, Section 6 demonstrates and evaluates the approach via a proof-of-concept implementation (cf. Nunamaker et al., 2015) comparing the detection accuracy across unfiltered and filtered results for the three sensor setups.

2 Related Work

This section identifies research gaps in (human) activity recognition, sensor-based process mining, and model-based approaches.

Human Activity Recognition (HAR) automatically classifies activities using data from video, accelerometer, or wearables (cf. Lara & Labrador, 2013). Applications range from healthcare to worker safety. For example, Faramondi et al. (2019) used waist-mounted IoT sensors to detect anomalous, potentially hazardous situations for workers. Recent advancements are driven by large datasets and deep learning (cf. Chen et al., 2021). While some focus on simple smartphone sensors (cf. Wan et al., 2020), others utilize graph convolutional networks (GCNs) on skeleton data (cf.

Peng et al., 2019). Context-aware HAR further incorporates human-object interactions via directed semantic graphs, often using graph attention networks or edge convolutions (cf. Corona et al., 2020). Recent ergonomics research also explores low-cost setups without strict process notions (cf. Demrozi et al., 2023; Raso et al., 2019). Crucially, none of these approaches utilize a dedicated process model to interlink activities in a temporal-logical order.

Sensor-based Process Mining: Predictive Process Monitoring aims to predict the current/next steps or outcomes of a process instance (cf. Francescomarino et al., 2018). While many methods rely on atomic process logs, IoT-based logs must be derived from complex, continuous sensor data. Although sensor fusion (cf. Rebmann et al., 2019) and specific aggregation steps (cf. De Leoni & Pellattiero, 2022) improve mining capabilities, uncertainty remains high. Knoch et al. (2018) demonstrated sensor-based monitoring for manual work, yet process model knowledge is rarely used to mitigate this uncertainty or enhance data quality.

Hybrid and Model-based Process Mining approaches: Some predictive monitoring approaches utilize a-priori instance knowledge (cf. Francescomarino et al., 2018), but the specific benefit of using process models for this purpose has not been systematically compared or explicitly stated. Notably, existing literature lacks approaches that leverage the process model to filter incorrect machine learning results from detection algorithms.

3 Motivating Use Case

The presented use case involves an ergonomics assessment at a large European OEM with highly standardized procedures and implicit process models. To address the physical strain of daily work, this paper identifies bottlenecks and problematic activities. This foundation enables future work to provide notifications and process recommendations to prevent ergonomically unfavorable postures.

3.1 Background: Upper-Extremity Workplace Ergonomics

Daily loads and physical stress on the upper extremities significantly impact the prevalence of work-related musculoskeletal disorders (WMSD) (cf. Schaefer et al., 2007). The Ergonomic Assessment Worksheet (EAWS) was developed as an expert

screening procedure to assess physical stress by tracking static postures, movements, action forces, manual loads, and repetitive tasks (cf. Schaub et al., 2012).

Based on stress degree and duration, a score-based system determines if movements remain within recommended boundaries. Project partners developed a BAN-based automatic ergonomic assessment using EAWS. While not detailed here, validating these process mining methods helps locate ergonomic bottlenecks. This enables operative guidance to prevent harmful situations, long-term advice to mitigate unavoidable awkward postures (cf. Sundstrup et al., 2020), and the identification of workplace redesign opportunities.

3.2 Use Case 1 – Pre-control of Injection Rails

The first use case (cf. Figure 1) describes a process in injection rail production for the automotive OEMs. As Figure 1 indicates, the rails are being delivered to the left and right side of the user. In a time-fixed cycle the user performs his regular actions which include picking up the rail from either the left or right container, turning to the inspection table, and placing the rail into the inspection holder. There, the rail is being checked for defects. If the inspection is successful, the user lifts the rail above some obstacle into a basin of water, where the rail is being washed and transported to its next destination.

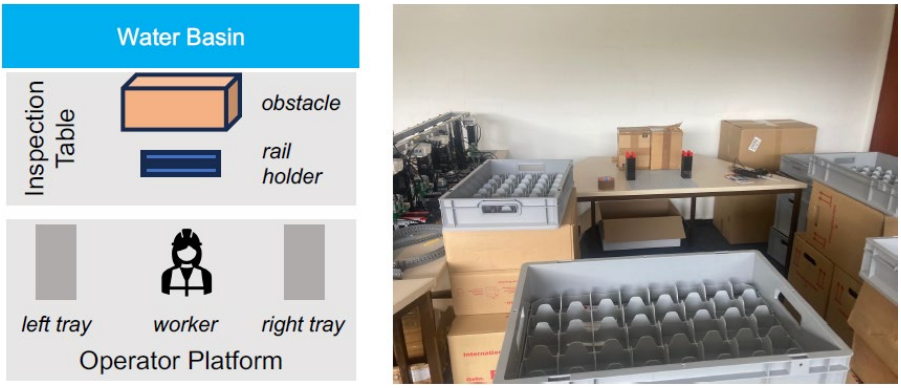

Figure 1: Use Case 1
Source: Own

3.3 Use Case 2 – Final Control and Packing of Rails

In Use Case 2, the rail is undergoing a final quality check before it is packed into boxes for delivery (cf. Figure 2). Once again, the delivery boxes are aligned left and right of the user. From a central conveyor belt, the user receives a rail that is being taken for inspection. To do so, the rail is put onto a holder. Now, the injection holes are closed with plastic caps, and the rail is being inspected with a magnifying glass. These two steps can be carried out in either order. Then the rail is taken again and deposited into one of the boxes.

Figure 2: Use Case 2
Source: Own

3.4 Relevance of the Use Cases for the Investigation

Although the use case processes are mostly linear, they expose some degrees of freedom which make classical HAR approaches hard to apply: In contrast to UC1, UC2 includes situations where both movement-based machine learning and process models may prove insufficient for accurate activity detection.

4 System Design

To capture the manual work processes described in Section 3, the system must balance detection accuracy against practical constraints such as sensor complexity, privacy, and worker comfort. The design therefore rests on two interdependent pillars: a body area network (BAN) that acquires upper-body movement data, and a process-aware software architecture that transforms raw sensor streams into activity predictions. The following sections detail both the hardware sensor setup and the

software components, with particular attention to how process model knowledge is embedded into the detection pipeline.

4.1 Hardware & Sensor Setup

The sensor suit consists of seven sensors of the type IMU (inertial measurement units) that are attached to the upper body and extremities via a sensor suit. The sensors as enumerated in the figure are: (0) for the lower back (root sensor), (1) left shoulder, (2) right shoulder, (3) left upper arm, (4) left lower arm, (5) right upper arm, (6) right lower arm. Each sensor is a digital signal processor (DSP) that returns three-dimensional angle data. The sensors are connected via cables to a multiplexing unit which is controlled by an Arduino board, that compresses the single sensor data into timed data frames. The Arduino board is connected to a Raspberry Pi 4 which contains all required local software capabilities of the body area network (BAN) (cf. section 4.2). The Raspberry Pi is connected to a power pack and a WiFi antenna to maintain connectivity to the main server.

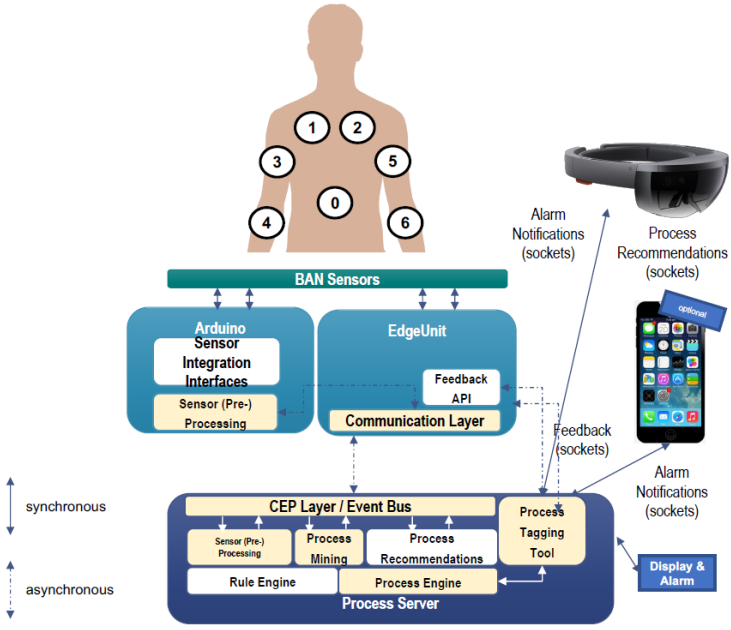

Figure 3: Hardware & Software Architecture (only yellow-highlighted components are relevant)

Source: Own

4.2 Software Architecture

As Figure 3 indicates, the overall system interacts with the following **components** in its infrastructure: the *Arduino* controller mainly has the functionality to multiplex and pre-process the sensor signals and collect them in common data frames and forwards them to the *Edge Unit*, which in our case is the Raspberry Pi. In this case, its main purpose is the data transport to a backend server via its communication layer, which is implemented with a socket connection.

The **Process Server** is the main backend server which is responsible for detecting, monitoring, analyzing and recommending the process and associated actions. In the presented task, the incoming data is received through the *Event Bus* via socket connections. The *Sensor Preprocessing* on this side prepares the data for the respective machine learning algorithms (cf. section 5), e.g. by aggregating individual data frames into windowed data frames, building stacker functions, etc. The *Process Mining component* performs the actual algorithms for determining the current and predicted process activities. This is the main focus of the evaluations in this work. The *Process Tagging Tool* is a Universal Web App that can be used both on PCs and tablets to tag the current activity. It is used to train the system in a supervised learning manner. Whenever a process model is chosen in this app, it shows the list of available activity types, and by selecting one the given activity type is started, while the previously running activity instance is stopped. The approach is sufficient for this purpose, as the use cases involve no parallel activities. Identifying these time segments allows filtering of relevant sensor data frames for training detection models. The *Process Engine* – an open-source jBPM implementation (cf. Cumberlidge, 2007) – is a standard business process management (BPM) engine used to enact and monitor instances. Triggered either by automated detection models or manual tagging via the Tagging App, it provides the ground truth for process data. The resulting process logs serve as output labels for detection models.

5 Activity Recognition and Model-based Post-filtering

The following describes the overall detection pipeline. The HAR component serves as an interchangeable baseline classifier; its specific architecture is not the focus of this work. The primary contribution is the rule-based filtering stage described in Section 5.4, which leverages process model knowledge to refine classifier output.

5.1 Input & Output Models

A DSP sensor captures 3-dimensional angles in space, with a set of sensors oriented in relation to a root sensor (sensor 0) located at the lower back. Consequently, a sensor data frame is defined as follows:

$$s_i = \langle x_i, y_i, z_i \rangle \; with \; x_i, y_i, z_i \in \mathbb{R}[-180, +180] \tag{1}$$

The input and output vectors for the learning models are defined as:

$$\langle s_1, \dots, s_i \rangle \rightarrow a \; with \; i \in \mathbb{N} = 3{,}5{,}7 \tag{2}$$

This setup utilizes either the full 7-sensor suit (as depicted in Figure 3) or specific subsets. The model predicts the label of the current activity.

5.2 Data Preparation & Stacking

To build a supervised predictive model, it is essential to build the data with proper labelling. The input vectors s_i and the output vectors a_i are obtained from two different components as described in Figure 3. The a_i vectors obtained from the process tagging tool are cleaned by identifying incomplete processes and wrong process tags due to human errors while tagging.

Once the data is cleaned, the input and output vectors are mapped to each other using the time stamp on which it was collected. Let a_j and a_{j+1} be two subsequent activities, any s_i collected between the corresponding timestamps of a_j and a_{j+1} will be labelled as a_j. Any s_i with no corresponding a_i is removed.

The cleaned and labelled data is then stacked with a windowing length of four, i.e. a 4x21 matrix is formed for every four sequential data points when we consider all the seven sensors with x_i, y_i and z_i for each. This enables the model to learn how an activity changes over a short period of time, making it more realistic and reliable.

The stacked data is then split into train, validation and test sets for model training. First, an 80:20 split is done on the stacked data to create a train set and a test set. A validation set is created by doing an 80:20 split on the train set created in the previous step. This results in an overall data split of 64 %|16 %|20 % for train, validation and test set respectively.

5.3 Network Architecture

A deep learning model is built to identify activities from sensor input vectors. To extract temporal features from sequential sensor data, a convolutional neural network (CNN) is integrated with a bidirectional long short-term memory (BiLSTM) layer. The CNN layers employ a 2x2 kernel and Rectified Linear Unit (ReLU) activation functions. For improved convergence speed, batch normalization is applied after each CNN layer, while dropout provides regularization to prevent overfitting. A BiLSTM layer with 64 units, followed by a linear layer and a softmax function, classifies the activities. Figure 4 illustrates this network architecture.

The Adam optimizer is employed with a learning rate of 10^{-4}. Its adaptive learning rate enables faster optimization compared to standard stochastic gradient descent. Cross-entropy serves as the loss function, as the softmax layer outputs a probability distribution across classes. Accuracy is chosen as model performance metric.

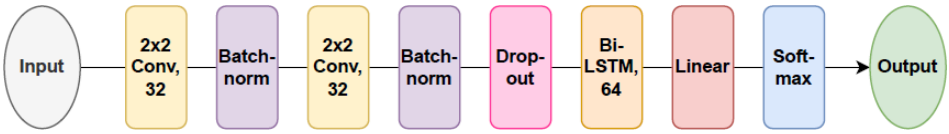

Figure 4: The network architecture used to do activity detection.
Source: Own

5.4 Rule-based Post-filtering of the Steps

Let M be the process model with activity types $\langle a_1, \ldots, a_n \rangle$ with n as the number of activity types contained in M. Let furthermore a_i' be instantiations of the respective activity type a_i. A successor candidate $sc(a_i)$ contains the set of all successor candidates of activity type a_i. The objective is to ensure that the current running activity instance a_i' matches the type detected by the machine learning algorithm.

The algorithm outputs probabilities for all available activity types $a_1, \ldots, a_n$. Through continuous evaluation of the active instance, all predictions are filtered out that neither satisfy the successor function $sc(a_i)$ nor match a_i itself. Given the prediction probability p for a candidate activity a_k, the activity type with the highest probability score hs among these candidates is selected (cf. Figure 5).

```
Data:   n ∈ ℕ
Data:   a'i ← current activity instance
Result: a* ← activity type with highest score in C

C ← sc(a'i);
if C ≠ ∅ then
    k ← 0;
    m ← |C|;                    # elements in C
    hs ← 0;
    a* = NIL;
    while k < m do
        s ← p(ak);
        if s > hs then
            hs ← s;
            a* ← ak;
        end
    end
end
```

Figure 5: Algorithm for post-filtering
Source: Own

6 Evaluation

To provide a proof-of-concept implementation, the approach was tested in a laboratory environment covering the use cases from Section 3. The evaluation is designed to isolate the contribution of the filtering stage. Accordingly, the comparison between unfiltered and filtered predictions – across varying sensor configurations – constitutes the primary experiment. The absolute performance of the underlying classifier is deliberately secondary; what matters is the delta introduced by filtering. The two use cases were tested with 27 participants. However, experiments were conducted under explicit instructions and supervised to ensure correct process execution. Models were trained during the test runs using the Tagging App described in Subsection 4.2. As stated in Section 3, users were free to choose between left/right variants, or in which order to perform certain tasks (UC2).

6.1 Experimental Setup

Three distinct sensor setups are evaluated for each use case: (i) root and shoulder sensors (0, 1, 2); (ii) root, shoulders, and upper arm sensors (0, 1, 2, 3, 5); and (iii) the full 7-sensor setup. To analyze the impact of post-filtering on predicted process activities, unfiltered and filtered approaches are compared across all sensor setups.

6.2 Results

Table 1 allows the comparison of different sensor setups and pre- and post-filtering of predicted process activities. The sensor setup with all seven sensors consistently performs best for both use cases and for the non-filtered and filtered approaches. The difference in test accuracies for the different sensor setups without any filtering is as high as approximately 10 %, however, when the filtering approach is used, it reduces significantly to 2.6 % – 5.36 %. It is evident that post-filtering predictions consistently and significantly improves test accuracy.

Table 1: Accuracy comparison of Filtering approaches

Case	Sensor Setup	Model Filtering	
		No	Yes
UC1	3	47.00 %	82.09 %
	5	50.40 %	83.35 %
	7	56.94 %	84.69 %
UC2	3	38.70 %	81.72 %
	5	42.04 %	85.12 %
	7	46.42 %	87.08 %

Source: Own

Figure 6 helps to understand class-wise performance. When no filtering is applied, the activities "check rail", "close openings" and "take rail" are often confused with "mover rail to the holder". Similarly, "bend down to box" is sometimes confused with "deposit rail in box". Applying the filtering approach, Figure 6 shows that almost all classes perform very well.

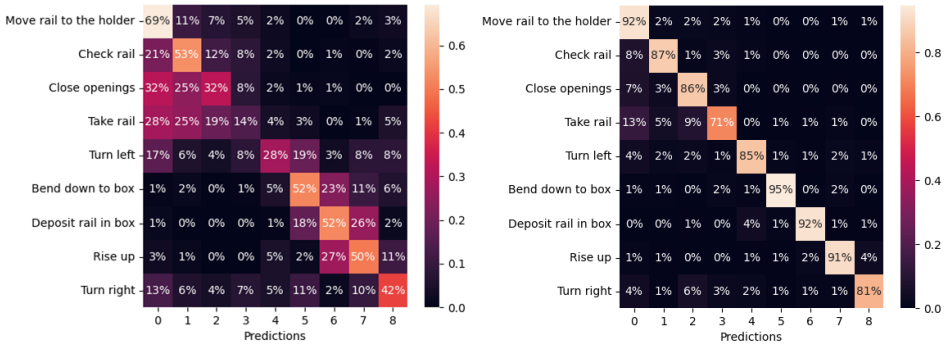

Figure 6: Confusion matrices for use case 2 without (left) and with (right) post-filtering of predicted activities. Rows represent actuals, columns represent predictions.
(Source: own)

6.3 Discussion

Regarding the **sensor setups**, Table 1 demonstrates that increasing the number of sensors improves detection accuracy. However, overall accuracy remains relatively low, with the 7-sensor configuration providing the greatest advantage. An analysis of the confusion matrices for UC2 reveals that the 3-sensor setup leads to high misclassification rates for manual activities such as "close openings" and "take rail". While these rates remain elevated across all configurations, they consistently decrease as sensors are added, indicating that the chosen sensor design does not fully capture hand movements. This is acceptable, however, as the primary goal described in Section 3 is to analyze upper-back ergonomics.

Model filtering consistently outperforms non-filtered alternatives. Notably, filtering considerably reduces the accuracy gap between the different sensor setups in both use cases. While results reach a "good" range, the peak accuracy (87.08 % for the UC2 7-sensor setup) remains modest. For both use cases, the frequency of confused activities (discussed in Section 6.2 and seen in Figure 6) improves significantly after filtering. **Limitations** that negatively impact performance are:

1. *Activity transitions:* Despite the largely linear process model which might suggest near 100 % accuracy the system must recognize activity boundaries. These transitions are not fully captured and should be addressed in future work.

2. *Inexperienced test subjects:* The use of laboratory participants rather than trained workers occasionally caused delays and execution errors, impacting the consistency of the sensor data.
3. *Simultaneous non-determinism of sensor setup and process model:* In cases with process decision splits, the process model alone cannot determine the next activity. If the sensor setup also fails to distinguish between these candidates (e.g., "check rail" and "close openings" in UC2), filtering approach lacks the necessary data to resolve the ambiguity reliably.
4. *Low raw accuracy of the baseline method:* Due to the very restricted sensor setup that intentionally prioritizes upper-body ergonomics over fine-grained finger or hand tracking. Consequently, these baseline results do not represent a failure of the machine learning model, but rather highlight the 'information gap' created by data minimization – a gap that our rule-based post-filtering approach is specifically designed to bridge.

7 Conclusion

This paper presents a novel approach for predictive monitoring of human work processes, leveraging model data to minimize sensor setups and machine learning requirements for activity recognition. Results from configurations with 3, 5, and 7 sensors across two laboratory-modeled industrial workplaces demonstrate that explicit process models enable a rules-based post-processing approach to recognize activities more efficiently and with fewer sensors than purely sensor-based methods. This proof-of-concept suggests the approach should be considered when benchmarking new capturing environments for human work processes.

Limitations: The models follow a sequential pattern with few splits or forks. While more advanced sensor setups or baseline algorithms reduce the accuracy gap between unfiltered and filtered approaches, filtered accuracy is still expected to remain significantly higher.

Outlook: Furthermore, the demonstrated ability to achieve over 80 % accuracy with only a 3-sensor setup has significant practical and economic implications. By reducing hardware requirements, the system lowers initial procurement costs and minimizes the time and training effort required for workers. These factors are critical for the adoption of ergonomic monitoring in real-world settings, particularly within

Small and Medium Enterprises (SMEs) where resource constraints often hinder the implementation of high-fidelity capturing environments.

Acknowledgements

This Paper has been written in the context of the project "ErgoBest" funded under the ID "16SV8672" by the German Federal Ministry of Education and Research (BMBF). We want to thank our project partners and especially our student assistants Jonybul Islam, Mejbah Uddin Shameem and Anushka Kothari for their assistance in the laboratory experiments.

References

Assenza, G., Fioravanti, C., Guarino, S., & Petrassi, V. (2020). New Perspectives on Wearable Devices and Electronic Health Record Systems. *2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT*, 740–745. 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT). https://doi.org/10.1109/MetroInd4.0IoT48571.2020.9138170

Chen, K., Zhang, D., Yao, L., Guo, B., Yu, Z., & Liu, Y. (2021). Deep Learning for Sensor-based Human Activity Recognition: Overview, Challenges, and Opportunities. *ACM Computing Surveys*, *54*, 1–40. https://doi.org/10.1145/3447744

Corona, E., Pumarola, A., Alenya, G., & Moreno-Noguer, F. (2020). *Context-Aware Human Motion Prediction*. 6992–7001. https://openaccess.thecvf.com/content_CVPR_2020/html/Corona_Context-Aware_Human_Motion_Prediction_CVPR_2020_paper.html

Cumberlidge, M. (2007). *Business Process Management with JBoss jBPM*. Packt Publishing. https://univ-senegal.scholarvox.com/catalog/book/88852238?_locale=en

De Leoni, M., & Pellattiero, L. (2022). The Benefits of Sensor-Measurement Aggregation in Discovering IoT Process Models: A Smart-House Case Study. In A. Marrella & B. Weber (Hrsg.), *Business Process Management Workshops* (Bd. 436, S. 403–415). Springer International Publishing. https://doi.org/10.1007/978-3-030-94343-1_31

Demrozi, F., Turetta, C., Kindt, P., Chiarani, F., Bacchin, R., Valè, N., Pascucci, F., Cesari, P., Smania, N., Tamburin, S., & Pravadelli, G. (2023). A Low-Cost Wireless Body Area Network for Human Activity Recognition in Healthy Life and Medical Applications. *IEEE Transactions on Emerging Topics in Computing*, *PP*, 1–12. https://doi.org/10.1109/TETC.2023.3274189

Faramondi, L., Bragatto, P., Fioravanti, C., Gnoni, M., Guarino, S., & Setola, R. (2019). *A Wearable Platform to Identify Workers Unsafety Situations* (S. 343). https://doi.org/10.1109/METROI4.2019.8792857

Francescomarino, C. D., Ghidini, C., Maggi, F. M., & Milani, F. (2018). *Predictive Process Monitoring Methods: Which One Suits Me Best?* (Bd. 11080, S. 462–479). https://doi.org/10.1007/978-3-319-98648-7_27

Hevner, A., R, A., March, S., T, S., Park, Park, J., Ram, & Sudha. (2004). Design Science in Information Systems Research. *Management Information Systems Quarterly*, *28*, 75.

Knoch, S., Ponpathirkoottam, S., Fettke, P., & Loos, P. (2018). Technology-Enhanced Process Elicitation of Worker Activities in Manufacturing. In E. Teniente & M. Weidlich (Hrsg.), *Business Process Management Workshops* (S. 273–284). Springer International Publishing. https://doi.org/10.1007/978-3-319-74030-0_20

Lara, O., & Labrador, M. (2013). A Survey on Human Activity Recognition Using Wearable Sensors. *Communications Surveys & Tutorials, IEEE*, *15*, 1192–1209. https://doi.org/10.1109/SURV.2012.110112.00192

Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O. A., Sepúlveda, M., Helm, E., Galvez-Yanjari, V., Rojas, E., Martinez-Millana, A., Aloini, D., Amantea, I. A., Andrews, R., Arias,

M., Beerepoot, I., Benevento, E., Burattin, A., Capurro, D., Carmona, J., Comuzzi, M., … Zerbato, F. (2022). Process mining for healthcare: Characteristics and challenges. *Journal of Biomedical Informatics*, *127*, 103994. https://doi.org/10.1016/j.jbi.2022.103994

Nunamaker, J., Briggs, R., Derrick, D., & Schwabe, G. (2015). The Last Research Mile: Achieving Both Rigor and Relevance in Information Systems Research. *Journal of Management Information Systems*, *32*, 10–47. https://doi.org/10.1080/07421222.2015.1094961

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Peng, W., Hong, X., Chen, H., & Zhao, G. (2019). *Learning Graph Convolutional Network for Skeleton-based Human Action Recognition by Neural Searching* (arXiv:1911.04131). arXiv. https://doi.org/10.48550/arXiv.1911.04131

Raso, R., Emrich, A., Burghardt, T., Sträter, O., Fettke, P., & Loos, P. (2019). *Einsatz eines textilbasierten Assistenzsystems zur Analyse von körperlich anstrengenden Arbeitsprozessen*. Springer. https://doi.org/10.22028/D291-33497

Rebmann, A., Emrich, A., & Fettke, P. (2019). Enabling the Discovery of Manual Processes Using a Multi-modal Activity Recognition Approach. In C. Di Francescomarino, R. Dijkman, & U. Zdun (Hrsg.), *Business Process Management Workshops* (S. 130–141). Springer International Publishing. https://doi.org/10.1007/978-3-030-37453-2_12

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance), 119 OJ L (2016). http://data.europa.eu/eli/reg/2016/679/oj

Schaefer, P., Boocock, M., Rosenberg, S., Jäger, M., & Schaub, Kh. (2007). A target-based population approach for determining the risk of injury associated with manual pushing and pulling. *International Journal of Industrial Ergonomics, Musculoskeletal Load of Push–Pull Tasks*, *37*(11), 893–904. https://doi.org/10.1016/j.ergon.2007.07.008

Schaub, K., Mühlstedt, J., Illmann, B., Bauer, S., Fritzsche, L., Wagner, T., Bullinger-Hoffmann, A., & Bruder, R. (2012). Ergonomic assessment of automotive assembly tasks with digital human modelling and the ' Ergonomics Assessment Worksheet ' (EAWS). *International Journal of Human Factors Modelling and Simulation*, *3*, 398. https://doi.org/10.1504/IJHFMS.2012.051581

Sundstrup, E., Seeberg, K., Bengtsen, E., & Andersen, L. (2020). A Systematic Review of Workplace Interventions to Rehabilitate Musculoskeletal Disorders Among Employees with Physical Demanding Work. *Journal of Occupational Rehabilitation*, *30*. https://doi.org/10.1007/s10926-020-09879-x

Wan, S., Qi, L., Xu, X., Tong, C., & Gu, Z. (2020). Deep Learning Models for Real-time Human Activity Recognition with Smartphones. *Mobile Networks and Applications*, *25*. https://doi.org/10.1007/s11036-019-01445-x

The Role of Supplier Trust in Enabling the Twin Transition

Alinda Kokkinou, Jan van Elderen
Breda University of Applied Sciences, Breda, the Netherlands
kokkinou.a@buas.nl, elderen2.j@buas.nl

The twin transition, the concurrent pursuit of digital and green transformation, requires organizations to develop specific dynamic capabilities. The present study examines how a Dutch textile company successfully transitioned from a linear to a circular business model. It identifies supplier trust as a critical yet underexplored enabler. Drawing on Dynamic Capability Theory and a longitudinal single-case study spanning 2021 to 2026, the study finds that trust-based supplier relationships underpin two key organizational dynamic capabilities: data analytic capability and sustainable supplier management capability. Supplier trust facilitates the cross-organizational data flows needed for evidence-based sustainability decision-making, while simultaneously enabling the collaborative R&D investments required for circular supply chain development. The findings challenge the assumption that the twin transition is primarily a technological challenge. Instead, the relational capabilities are foundational preconditions for effective digitalization.

DOI
https://doi.org/10.18690/um.fov.4.2026.49

ISBN
978-961-299-152-4

Keywords:
twin transition,
dynamic capabilities,
supplier trust,
circular business model,
data analytic capability

University of Maribor Press

1 Introduction

Climate change is forcing businesses and governments to take urgent action in order to mitigate future natural disasters (Ghadge et al., 2020). The manufacturing and production sectors are estimated to account for one-fifth of the world's carbon emissions (World Economic Forum, 2022). The textile industry, in particular, was estimated to be the fifth highest user of raw materials and greenhouse gas emissions in 2020, in addition to causing the third highest pressure on water and land use (European Environment Agency, 2024). In Europe, this has prompted the European Commission to formulate the European Green Deal, a comprehensive strategy to achieve climate neutrality by 2050. Recognizing digitalization as an enabler of sustainability, the European Union also prioritized digital transformation as part of the European Green Deal. The concurrent pursuit of green and digital transformations has been coined the "twin transition." This concept refers to the "*intertwined and simultaneous green and digital transition to offset companies' carbon footprint*" (Rehman et al., 2023, p. 1) and centers on the notion that the green and digital transformation are synergistic (Accenture Research, 2021). While extensive research has been conducted on the topic from a policy perspective, relatively few studies have examined it from a business perspective, on an organizational level (Tabares et al., 2025). Furthermore, research on digital transformation and research on sustainable and circular transformation have often been conducted from within discipline silos (Christmann et al., 2024). So, despite digitalization being widely acknowledged as a potential enabler of sustainability, the complex relationship between the two remains only marginally understood (Patrucco et al., 2025).

For organizations striving to develop their sustainable and circular business models, a better understanding of how the twin transition unfolds at the organizational level is therefore needed (Christmann et al., 2024). This includes better insight into the organizational capabilities needed for such a transformation (Belhadi et al., 2022). Central amongst these capabilities are digitalization capabilities, which enable the data exchange and transparency required to support circular business models (Kokkinou & Mitas, 2026). However, such capabilities are contingent on suppliers' willingness to share data and knowledge across organizational boundaries. This suggests that buyer-supplier trust is a critical yet underexplored antecedent (Gualandris & Kalchschmidt, 2016), leading to two research questions: (RQ1) *how does trust enable the development of digitalization capabilities in support of a circular business model*

transition? and (RQ2) *how does trust enable the development of sustainable supplier management capabilities in the context of the transition to a circular business model?* The purpose of the present study was therefore to explore the capabilities needed to successfully transition to a circular business model. Given the complex and emerging nature of twin transition in combination with a relative lack of empirical studies on the topic, a longitudinal single-case study design was used to examine how a focal organization successfully transitioned from a linear to a circular business model, identifying supplier trust as an enabler of both digitalization and sustainable supplier management capabilities. The study thereby contributes to the body of knowledge on technology-driven sustainability transitions, advancing our understanding of the mechanisms by which digitalization shapes sustainability- and circularity-oriented practices (Patrucco et al., 2025).

2 Review of the Literature

The Dynamic Capability Theory (DCT), a powerful theory explaining how organizations develop and maintain their competitive advantage in changing market and environmental conditions (Teece et al., 1997) forms the theoretical background for this study. Dynamic capabilities are combinations of assets, resources, and organization-specific processes that allow organizations to sense their environment, seize opportunities, and reconfigure accordingly (Teece, 2007). According to DCT, organizations develop dynamic capabilities to maintain their competitive advantage as market conditions change or as disruptions occur.

While sustainability and circularity are distinct, sustainability practices are essential for circularity (Romagnoli et al., 2023). The adoption of sustainable practices has been shown to contribute to circular supply chain functionality and thereby improve business performance. Gualandris & Kalchschmidt (2016) distinguish between sustainable process management and sustainable supplier management. Sustainable process management practices primarily focus on reducing waste, improving health and safety within the organization, and developing environmentally friendly products. They are often developed over time, through trial and error, by combining assets and resources in a unique way. Sustainable supplier management practices typically span organizational boundaries and include environmental and social assessments of suppliers and, more generally, collaboration within the supply chain.

Gualandris & Kalchschmidt (2016) find sustainable process management to be an antecedent of sustainable supplier management.

Even though circularity is supported by sustainability practices, the two are distinct. Whereas sustainability practices principally focus on reducing waste across the supply chain, the circularity paradigm is more comprehensive, extending to preserving (Dwivedi & Paul, 2022) and even revitalizing materials and products (Bhawna et al., 2024). Aligning organizational strategy with circular economy principles entails '*transitioning from the linear "take, make, dispose" approach to designing products and processes that prioritize extended usage cycles, waste reduction, and material recovery*' (Bhawna et al., 2024, p. 1053). However, this requires the acquisition, processing, and exchange of information within and across organizations within the supply chain, leading to the hypothesis that digitalization capabilities could contribute to innovative and circular business models (Patrucco et al., 2025).

In the context of sustainability and circularity, digitalization capabilities enable organizations to enhance their competitive advantage by adopting and implementing digital technologies that improve their sustainable performance. Several such digital technologies have been shown to contribute to improved sustainable performance and sustainable reporting (Kokkinou, 2025; Mehedintu & Soava, 2023). For example, machine learning algorithms can improve forecast accuracy, thereby reducing overproduction, surplus inventory, and associated waste. Internet of Things (IoT) sensors and Blockchain technology can be used to track products across the supply chain, increasing transparency and visibility, and improving sustainability reporting. Big data analytics can improve sustainability by enabling organizations to make more complex decisions that involve a broader set of variables and criteria. Digitalization capabilities can thus serve as a foundation for exploring alternative sustainability scenarios and informing sustainable decision-making. Nevertheless, the mechanisms through which organizations should develop their digitalization capabilities in support of circular business models remain poorly understood (Patrucco et al., 2025), motivating RQ1.

A specific digitalization capability, DAC, has been identified as an enabler of improved supply chain performance, supply chain resilience, and competitive advantage (Kokkinou et al., 2023). DAC is defined as an organization's capability to deploy data, technology, and people to quickly access and analyze information in

support of complex decision-making (W. Yu et al., 2021). DAC is conceptualized as a dynamic capability as it allows organizations to sense their environment (e.g., by identifying new demand patterns), seize opportunities (e.g., by using data-driven decision-making to evaluate alternatives), and make progress towards transforming and reconfiguring their activities (e.g., through improved knowledge management) (Kokkinou, 2024). In the context of sustainability and circularity, a well-developed DAC is linked to an organization's ability to facilitate stakeholders, both within and outside the organization, in making transparent sustainability-related choices by drawing, among other sources, on product and service lifecycle data (Christmann et al., 2024). Critically, however, DAC depends on the availability and quality of data from supply chain partners. If suppliers are unable or unwilling to share data, DAC will have limited effectiveness, underscoring the importance of trust as a precondition for DAC development in the context of the twin transition.

In the context of sustainable supply chain performance, an important yet frequently overlooked antecedent is supplier trust (Gualandris & Kalchschmidt, 2016). Buyer-supplier trust is described as an organization's confidence in the reliability and integrity of its supply chain partner, within an exchange relationship based on the expectation of cooperation (Gualandris & Kalchschmidt, 2016; Moorman et al., 1992). In the context of sustainable supply chain practices, Anisul Huq et al. (2014, p. 630) showed that "*a combination of monitoring and trust building, actively developing and training suppliers, and taking the cultural context into account may be the most efficient way forward.*" In the context of supplier collaboration, Braziotis and Tannock (2011) distinguished between contractual, enabling, and enhancing factors. Once contracts have covered the explicit aspects of the relationship (contractual factors), the practical aspects of the relationship (enabling factors, such as the levels and methods of information sharing) kick in, influencing the implicit aspects of the relationship (enhancing factors, such as trust and commitment). Trust is a form of non-economic influence among supply chain partners (Parmigiani et al., 2011) and is closely linked to supply chain collaboration. It is a relational capability that organizations can develop to promote flexibility and innovation. While digitalization capabilities have received extensive attention as facilitators of information exchange across organizational boundaries, they are contingent on supply chain partners' willingness to share information and, by extension, trust. First and foremost, information sharing is needed for supplier audits and assessments. More broadly, trust between supply chain stakeholders facilitates knowledge transfer (Parmigiani et al., 2011).

Open dialogue and trust between supply chain partners also facilitate joint problem-solving. Trust between supply chain partners can also support investments in relationship-based assets (Parmigiani et al., 2011). In addition to intangible assets such as recycling practices and patents, these can include tangible assets such as remanufacturing equipment or the development and implementation of new production lines. To summarize, buyer-supplier trust emerges as a relational dynamic capability that enables both the information flows required for DAC (RQ1) and the collaboration required for sustainable supplier management (RQ2).

3 Methods

A longitudinal in-depth case study was used to investigate the transition of a mid-sized focal company from a linear to a circular business model. In-depth case study research is particularly appropriate for the study of emerging and complex phenomena, such as the transition to a circular business model, where prior empirical research is scarce, but an opportunity to examine it in the field arises (Heldt & Pikuleva, 2025). The unit of analysis for the study was the focal company and its supply chain.

The selection of the focal case for the present study was serendipitous. Based on purposive sampling, the focal case company was selected in 2021 for its strong sustainability orientation and geographical proximity to the researchers. After two waves of data collection in 2022 and 2024, it came to the attention of the researchers that the focal company had successfully "closed the loop" and data collection was therefore resumed. During each data collection wave, unstructured interviews were conducted with key informants from the focal company. Additionally, internal company documents, publicly available sustainability reports for the period 2021-2024, and press releases linked to third-party sustainability and/or entrepreneurship award nominations were collected and analyzed to complement, validate, and triangulate the interview findings. The preliminary findings highlighted the substantial role of suppliers. Therefore, additional data collection with suppliers, IT suppliers, and other members of the company's ecosystem has been scheduled but not yet completed, as shown in Figure 1. The preliminary findings are therefore subject to bias, as the data captured only opinions and perspectives from the focal company.

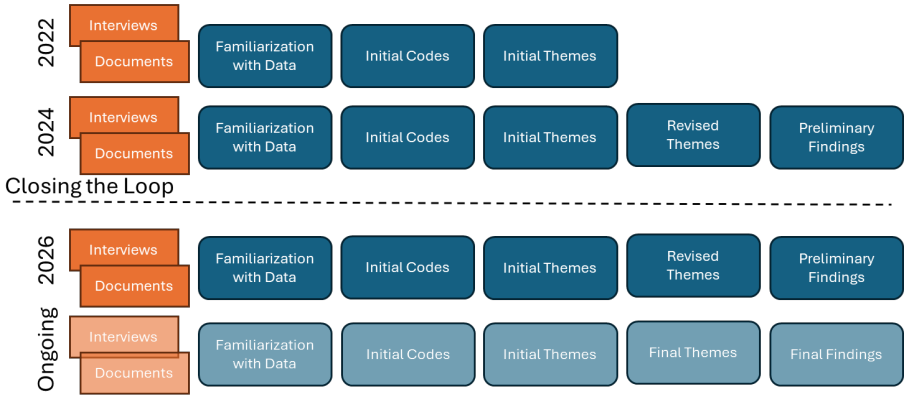

Figure 1: Data Collection & Analysis Timeline
Source: Authors' Own Work

The focal company is a Dutch family-owned mid-sized textile company. This B2B company aims to produce sustainable clothing following a made-to-order business model. The focal company embraces inclusivity. This implies that sustainability objectives need to be balanced against affordability, and that clothes produced need to fit a variety of body types. The company has been in existence for approximately 20 years. The design of clothing and the commercial activities take place in the Netherlands, while clothing manufacturing is outsourced to companies in China. The organizational and cultural context of the study is thus quite distinctive, as it takes place at the intersection of Dutch and Chinese cultural contexts, where the former is characterized by strong individualism and the latter by guanxi, while the long-lasting relationships at its heart are made possible by the Dutch company's family ownership. While the focal company has embraced sustainability since its inception, it was constrained in its early years by available technology and production processes. It iteratively developed and/or adopted new production techniques in cooperation with its suppliers, enabling it to reduce the volumes of water, dye, and chemicals it used and to lower CO2 emissions. As these changes required adjustments to production and other supplier processes, it further developed close collaborations with its Chinese suppliers, involving them in research & development activities. Furthermore, it prioritized monitoring its Tier 1, Tier 2, and Tier 3 suppliers for social sustainability. The focal company also proactively embraced sustainability reporting to communicate transparently with customers and other stakeholders.

Table 1: Sample of Themes, Codes and Illustrative Quotes (not exhaustive)

Theme	Code	*Illustrative Quote* [Data Source and Date]
Emphasis on Measurable Impact	Preference for 'impact' over 'sustainability'	*"I don't want to be the most sustainable company in the world. I want to make products with impact, and with the right impact* [Interview 2026].
	Quantification through LCA and product testing	*"we do a lot based on data. Everything is tested based on functionality, physical properties, and sometimes chemical properties. LCAs are used for all major material decisions"* [Interview 2026].
	Third-party certifications and audits	*In addition, [focal company] invests in initiatives and certifications such as the Business Social Compliance Initiative (BSCI), the Global Recycled Standard (GRS) and Green Button to ensure that all processes are carried out under defensible conditions.* [Jury Statement of Award Nomination, 2022]
Experimentation and Data-Driven Decision-Making	Innovation initiated without pre-existing data	*'It's not like we first have a whole lot of data before we say "hey let's try this"* [Interview 2026]
	Structured experimentation and validation process	*We have a process where (1) we define what circular, affordable, functional means, (2) explore alternatives, (3) experiment with promising options (4) validate with data and LCA and (5) decide on launch* [Interview 2026]
	Data validates rather than initiates	*you don't wait until you have the data to experiment, the goal of the experiment is to get the data.* [Interview 2026]
Trust / Relationship with Suppliers	Long-term supplier relationships	*Together with its Chinese partners [names of suppliers], [focal company] has been pursuing the goal of making its supply chains sustainable and fair for over 20 years.* [Jury Statement of Award Nomination, 2022]
	Personalized rapport with each supplier	*At each team meeting [with the Chinese office], I would bring chocolate bonbons. With COVID I could not travel, so I had them shipped* [Interview 2022]
	Close collaboration enabling supplier investment in new capabilities	*Through significant investments of time,resources, and capital, we've built long-term partnerships with key players across our value chain. Together, we've developed strategies to promote sustainable resources, ethical labour practices, and responsible processes, ensuring that our growth is both shared and sustainable.* [Sustainability Report 2023-2024]

Source: Authors' Own Work

Thematic analysis was used to code the full set of interview transcripts, interview notes, and internal and external company documents, following the procedures of Braun and Clarke (2006). After each round of data collection, the researchers familiarized themselves with the data and coded it using a combination of inductive and deductive coding. Subsequently, initial themes were formulated, iteratively refined, and named. After the second round of data collection in 2024, preliminary findings were documented in a case description. After the company's successful "closing of the loop" and corresponding rebranding, another round of data

collection was conducted in early 2026. The procedures were repeated on the complete dataset, and the resulting themes were documented in the present manuscript. Table 1 illustrates this process using a small subset of themes, codes, and data.

4 Preliminary Findings

The preliminary analysis of the data collected so far highlighted several aspects linked to digitalization capabilities that enabled the case company to transition from a linear to a circular business model, as illustrated in Figure 2:

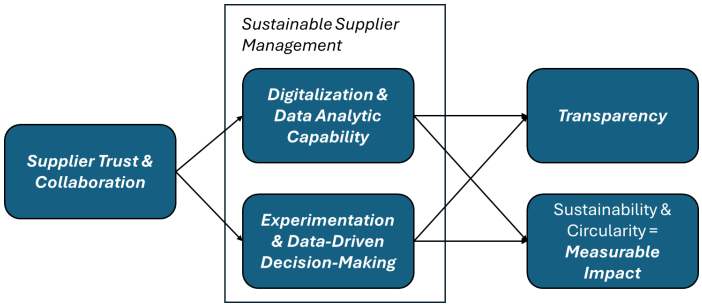

Figure 2: Results of Preliminary Thematic Analysis
Source: Authors' Own Work

Focus on Measurable Impact. The focal company consistently frames sustainability as something that must be demonstrably through measurement, and not merely as an abstract moral ideal. The CEO emphasizes the importance of making products with the *right* impact. This is quantified in liters of water saved and kilograms of CO2 reduction per item of clothing. Life Cycle Analyses (LCAs) are the primary tool for this and are also applied consistently to major product and material decisions. However, the *right* impact is also assessed in relation to product functionality, affordability, and commercial viability. This multi-stakeholder perspective is particularly relevant given the focal company's objective to offer affordable yet sustainable clothing. This focal company's data-driven orientation is clearly linked to its anti-greenwashing position: sustainability achievements must be earned through measurement, not merely claimed. As noted by the CEO during the interviews: "*for me, the most important is to show impact, and to validate that what we say we do, we really do achieve.*" The company's sustainability reports operationalize this

position by providing transparency and concrete key achievements, along with a roadmap to a specific date. A comparison of sustainability reports over time supports the growing importance of evidence-based sustainability reporting, not only in terms of KPIs, but also through certifications and audits.

Transparency in the supply chain. For the focal company, supply chain transparency is both a strategic priority and an ongoing operational challenge. The company has made significant investments in supply chain visibility. It uses a third-party supplier to track its orders across the supply chain using Blockchain technology. Retailers can use it to collect and verify product and supply chain data directly from the primary source. Even though data provision is required of suppliers, data collection from them remains fragmented. Not all suppliers see the added value in providing the requested information, and the quality of the data they provide also varies. Even though some data comes from Blockchain, much is still shared via Excel sheets of varying completeness and quality. Internally, the sustainability report is still compiled through an annual process of pulling data from multiple siloed sources. The focal company aspires to implement a dedicated supplier portal to standardize data entry and improve data quality.

Experimentation and the use of data insights to support decision-making. The focal company operates according to a deliberate logic of experimentation. New project opportunities are pursued based on clear strategic intent and reasoning, without pre-existing data, as such data is usually unavailable at the outset of the project. The project typically begins with defining an objective, such as a circular, functional, and affordable alternative, followed by an exploration of possible solutions. Once a plausible solution is identified, experimentation begins. The experimentation also generates the data required to validate or reject the project opportunity. As an interviewee describes: "*we do a lot based on data. Everything is tested based on functionality, physical properties, and sometimes chemical properties.*" The company's sustainability report confirms this, identifying company-specific techniques developed through internal innovation and subsequently validated through LCA and product testing. Furthermore, the focal company clearly recognizes the threats of data-driven decision-making for innovation. Waiting until data is available prior to initiating a new project may lead to delays and missed opportunities. As an interviewee explained: "*you don't wait until you have the data to experiment, the goal of the*

experiment is to get the data. […] *Don't go looking for all kinds of data before deciding to try something. You lose a lot of speed that way.*"

Role of Digitalization and DAC. The focal company recognizes digitalization as essential for better sustainability decisions. As an interviewee explains: "*my idea is that you really need digitalization to make better decisions regarding circularity*." However, the current state shows a significant gap between the company's internal data maturity and its ability to capture data across its supply chain. Internally, the company's "own" data is well captured and consistently used for data-driven decision-making. Important decisions based on measurable impact, such as actual reduction in CO2 emissions, water usage, etc., through e.g., LCAs. However, as manufacturing and other supply chain processes are outsourced, the focal company needs data from its suppliers. This data collection is still taking place ad hoc: some data is captured via blockchain technology, while other data is provided by suppliers using Excel sheets, requiring varying effort to safeguard data quality. The focal company would prefer to collect data in real time. However, this requires significant investments in digital tools (e.g., a supplier portal) and in relational work.

Relationships with Suppliers. For the focal company, building a sustainable supply chain is as much a people project as a technical one. Supplier relationships at the focal company are explicitly long-term, trust-based, and personal. The CEO visits suppliers regularly and invests significant time in building genuine rapport. This rapport is tailored to each individual, as some are convinced by data, others by commercial continuity, and yet others by social interaction. This depth of relationship also serves a strategic function: suppliers who trust that orders will not disappear overnight are far more likely to invest in new production lines, adapt to unfamiliar materials, and accept the added complexity occasionally required by sustainable production. Furthermore, suppliers engaged in such longstanding relationships are more likely to undertake the collaborative R&D that is necessary "to close the loop". The CEO rejects the transactional view that suppliers should be abandoned after a single mistake. Instead, he expressed strong support for a collaborative approach, saying in the context of sustainable practices: "*alone I can do a lot, but together we can get everything done.*" However, these relationships were not unconditional. Suppliers who refused to engage with the sustainability direction were let go, even after decades of collaboration. Supplier trust is also seen as a prerequisite for data sharing. Suppliers are willing to provide data and invest in data quality,

however, within the context of the buyer-supplier relationship. As an interviewee summarizes, "*I think you first need to have trust in the supply chain. Only then can you get the data […] transparency only comes after maintaining good relationships with your suppliers.*"

5 Discussion

The present study set out to examine two research questions. Regarding RQ1, the findings demonstrate that supplier trust is an enabler of the data flows that support the development of DAC. The company's ability to conduct LCAs, report on CO2 reductions, and build a real-time sustainability dashboard depends on the willingness of Tier 1, Tier 2, and Tier 3 suppliers to share accurate and timely data, supporting Parmigiani et al. 's (2011) view that trust is essential to information exchange. A well-developed DAC, in turn, enables the organization to make data-driven decisions about which sustainable practices to pursue. However, the study findings nuance the role of data in the innovation process, finding that a data-driven approach can act as an inhibitor. DAC plays a relatively limited role in the initiation phase of innovation, and becomes more relevant during the validation phase, when data from product testing and LCAs inform decision-making. The study thereby contributes to our understanding of the role that DAC plays in sustainable transition by demonstrating that a requirement of data-driven justification early in this process can impede innovation and progress towards more sustainable practices (Annarelli et al., 2021). It illustrates Patrucco et al.'s (2025) view that the effectiveness of digital tools in enhancing sustainable performance is contingent on how organizations integrate them into their sustainable practices. Regarding RQ2, the findings point towards trust-based supplier relationships acting as an enabler towards the development of sustainable supplier management capabilities over time. The transition to a circular business model was made possible by collaborative research & development activities undertaken by the focal company with suppliers. The case findings also provide further support for Anisul Huq et al.'s (2014) call for a more collaborative relationship between supply chain stakeholders. Yet, our findings somewhat contradict Braziotis and Tannock (2011) who find that information exchange precedes relations of trust, reversing the order in their model. Our findings show that trust and commitment (enhancing factors) are needed for information exchange (enabling factors). Together, the findings suggest that trust acts as dual enabler for the twin transition: it supports both the data flows needed for digitalization

capabilities (RQ1) and the collaborative investments needed for sustainable supplier management (RQ2).

The present study makes several contributions to the theoretical underpinnings of the twin transition by examining it from a business perspective and at the organization level (Belhadi et al., 2022; Tabares et al., 2025). First, the present study empirically substantiates the role of supplier trust (Gualandris & Kalchschmidt, 2016) as an antecedent of both DAC and sustainable supplier management capabilities, which have until now mostly been studied in isolation. By examining them jointly, the present study demonstrates that they are mutually reinforcing: trust-based relationships generate the data needed for impact measurement (Schilling & Seuring, 2024), while data credibility can, in turn, reinforce supplier confidence and commitment to the partnership. Second, the study supports and extends the DCT framing of the twin transition by identifying trust as a relational dynamic capability that supports the previously identified digitalization capabilities (Annarelli et al., 2021; Mehedintu & Soava, 2023). Third, the findings challenge the assumption that data sharing is primarily a technical challenge. Instead, the study findings support Galandris and Kalchschmidt's (2016) view that sustainable supplier management is a higher-order capability that depends on relational antecedents. From this perspective, these relational antecedents are at least as important as the technical infrastructure.

The present study is subject to several limitations. First, the findings are based on preliminary data from a single longitudinal case study, which limits their generalizability. While the case offers rich contextual insights, future research should examine the role of supplier trust in enabling the twin transition across multiple organizations and industries. Second, data collection so far has focused primarily on the focal company's perspective. Interviews with suppliers, partners, and other supply chain partners should provide a more comprehensive picture of the relational dynamics at play. Third, as noted by Galandris and Kalchschmidt's (2016), buyer-supplier trust may operate differently across cultures and contexts. In the particular context of this study, collectivism and guanxi may facilitate the development of relational capabilities in ways that differ from other contexts. Future research should therefore examine whether the trust mechanisms captured in this study hold across different cultural and contextual settings.

References

Accenture Research. (2021). *The European Double Up: A twin strategy that will strengthen competitiveness.* https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-144/accenture-the-european-double-up.pdf

Anisul Huq, F., Stevenson, M., & Zorzini, M. (2014). Social sustainability in developing country suppliers: An exploratory study in the ready made garments industry of Bangladesh. *International Journal of Operations & Production Management*, *34*(5), 610–638. https://doi.org/10.1108/IJOPM-10-2012-0467

Annarelli, A., Battistella, C., Nonino, F., Parida, V., & Pessot, E. (2021). Literature review on digitalization capabilities: Co-citation analysis of antecedents, conceptualization and consequences. *Technological Forecasting and Social Change*, *166*, 120635. https://doi.org/10.1016/j.techfore.2021.120635

Belhadi, A., Kamble, S., Gunasekaran, A., & Mani, V. (2022). Analyzing the mediating role of organizational ambidexterity and digital business transformation on industry 4.0 capabilities and sustainable supply chain performance. *Supply Chain Management: An International Journal*, *27*(6), 696–711. https://doi.org/10.1108/SCM-04-2021-0152

Bhawna, Kang, P. S., & Sharma, S. K. (2024). Bridging the gap: A systematic analysis of circular economy, supply chain management, and digitization for sustainability and resilience. *Operations Management Research*. https://doi.org/10.1007/s12063-024-00490-4

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braziotis, C., & Tannock, J. (2011). Building the extended enterprise: Key collaboration factors. *The International Journal of Logistics Management*, *22*(3), 349–372. https://doi.org/10.1108/09574091111181363

Christmann, A.-S., Crome, C., Graf-Drasch, V., Oberländer, A. M., & Schmidt, L. (2024). The Twin Transformation Butterfly: Capabilities for an Integrated Digital and Sustainability Transformation. *Business & Information Systems Engineering*, *66*(4), 489–505. https://doi.org/10.1007/s12599-023-00847-2

Dwivedi, A., & Paul, S. K. (2022). A framework for digital supply chains in the era of circular economy: Implications on environmental sustainability. *Business Strategy and the Environment*, *31*(4), 1249–1274. https://doi.org/10.1002/bse.2953

European Environment Agency. (2024). *What are the environmental impacts of the textile industry?* https://www.eea.europa.eu/en/about/contact-us/faqs/what-are-the-environmental-impacts-of-textiles

Ghadge, A., Wurtmann, H., & Seuring, S. (2020). Managing climate change risks in global supply chains: A review and research agenda. *International Journal of Production Research*, *58*(1), 44–64. https://doi.org/10.1080/00207543.2019.1629670

Gualandris, J., & Kalchschmidt, M. (2016). Developing environmental and social performance: The role of suppliers' sustainability and buyer–supplier trust. *International Journal of Production Research*, *54*(8), 2470–2486. https://doi.org/10.1080/00207543.2015.1106018

Hald, K. S., & Mouritsen, J. (2018). The evolution of performance measurement systems in a supply chain: A longitudinal case study on the role of interorganisational factors. *International Journal of Production Economics*, *205*, 256–271. https://doi.org/10.1016/j.ijpe.2018.09.021

Heldt, L., & Pikuleva, E. (2025). When upstream suppliers drive traceability: A process study on blockchain adoption for sustainability. *International Journal of Physical Distribution & Logistics Management*, *55*(3), 196–222. https://doi.org/10.1108/IJPDLM-01-2024-0022

Kokkinou, A. (2024). How organizations can harness continuous improvement practices to develop their data analytic capability: A conceptual paper. In T. van Kollenburg, A. Kokkinou, & O. McDermott (Eds.), *Challenging the Future with Lean: 9th IFIP WG 5.7 European Lean Educator Conference, ELEC 2023, 's-Hertogenbosch, the Netherlands, October 24-26, 2023, Proceedings.* Springer.

Kokkinou, A. (2025). Data analytic capability as an enabler of CSRD readiness: A moderated mediation analysis. *Journal of Global Responsibility*, 1–21. https://doi.org/10.1108/JGR-04-2025-0119

Kokkinou, A., Mandemakers, A., & Mitas, O. (2023). Developing Resilient and Robust Supply Chains through Data Analytic Capability. *Continuity & Resilience Review*, *5*(3), 320–342. https://doi.org/10.1108/CRR-07-2023-0013

Kokkinou, A., & Mitas, O. (2026). Driving the circular economy: How artificial intelligence could empower transportation and logistics companies to lead the transition. *Cleaner Logistics and Supply Chain*. https://doi.org/10.1016/j.clscn.2026.100300

Mehedintu, A., & Soava, G. (2023). Approach to the Impact of Digital Technologies on Sustainability Reporting through Structural Equation Modeling and Artificial Neural Networks. *Electronics*, *12*(9), 2048. https://doi.org/10.3390/electronics12092048

Mikalef, P., & Krogstie, J. (2020). Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities. *European Journal of Information Systems*, *29*(3), 260–287. https://doi.org/10.1080/0960085X.2020.1740618

Moorman, C., Zaltman, G., & Deshpande, R. (1992). Relationships between Providers and Users of Market Research: The Dynamics of Trust within and between Organizations. *Journal of Marketing Research*, *29*(3), 314–328. https://doi.org/10.1177/002224379202900303

Parmigiani, A., Klassen, R. D., & Russo, M. V. (2011). Efficiency meets accountability: Performance implications of supply chain configuration, control, and capabilities★. *Journal of Operations Management*, *29*(3), 212–223. https://doi.org/10.1016/j.jom.2011.01.001

Patrucco, A., Seuring, S., Fosso Wamba, S., Kaliyan, M., & Appolloni, A. (2025). Guest editorial: The missing link between supply chain technologies and sustainability issues: advancing theory and practice. *International Journal of Physical Distribution & Logistics Management*, *55*(3), 177–195. https://doi.org/10.1108/IJPDLM-04-2025-559

Rehman, S. U., Giordino, D., Zhang, Q., & Alam, G. M. (2023). Twin transitions & industry 4.0: Unpacking the relationship between digital and green factors to determine green competitive advantage. *Technology in Society*, *73*, 102227. https://doi.org/10.1016/j.techsoc.2023.102227

Romagnoli, S., Tarabu', C., Maleki Vishkaei, B., & De Giovanni, P. (2023). The Impact of Digital Technologies and Sustainable Practices on Circular Supply Chain Management. *Logistics*, *7*(1), 1. https://doi.org/10.3390/logistics7010001

Schilling, L., & Seuring, S. (2024). Linking the digital and sustainable transformation with supply chain practices. *International Journal of Production Research*, *62*(3), 949–973. https://doi.org/10.1080/00207543.2023.2173502

Tabares, S., Parida, V., & Chirumalla, K. (2025). Twin transition in industrial organizations: Conceptualization, implementation framework, and research agenda. *Technological Forecasting and Social Change*, *213*, 123995. https://doi.org/10.1016/j.techfore.2025.123995

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350. https://doi.org/10.1002/smj.640

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(7), 509–533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7%253C509::AID-SMJ882%253E3.0.CO;2-Z

World Economic Forum. (2022). *Reducing the carbon footprint of the manufacturing industry through data sharing*. https://www.weforum.org/impact/carbon-footprint-manufacturing-industry/

Yu, D., Tao, S., Hanan, A., Ong, T. S., Latif, B., & Ali, M. (2022). Fostering Green Innovation Adoption through Green Dynamic Capability: The Moderating Role of Environmental Dynamism and Big Data Analytic Capability. *International Journal of Environmental Research and Public Health*, *19*(16), 10336. https://doi.org/10.3390/ijerph191610336

Yu, W., Wong, C. Y., Chavez, R., & Jacobs, M. A. (2021). Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture. *International Journal of Production Economics*, *236*, 108135. https://doi.org/10.1016/j.ijpe.2021.108135

The Need for Deployment Norms in Value Sensitive Data-Sharing Infrastructures

Christel van de Wal, Vanessa Dirksen,
Roger Bons, Johan Versendaal

Open Universiteit, Faculty of Science, Heerlen, the Netherlands
christel.vandewal@ou.nl, vanessa.dirksen@ou.nl, roger.bons@ou.nl,
johan.versendaal@ou.nl

Governmental deployment of decentralized data-sharing infrastructures embeds governance functions within technical architectures. Although Value Sensitive Design (VSD) is widely used to integrate public values into socio-technical design, practical uncertainty remains regarding its capacity to guide binding ex ante deployment decisions under conditions of value pluralism and institutional accountability. A normative conceptual analysis, informed by a conceptual literature review, identifies three justificatory limits of VSD in governmental deployment contexts: (i) the absence of ethical commitment, (ii) the reliance on empirical convergence without normative authority, and (iii) the lack of a mechanism for resolving value conflicts. These limits risk embedding value priorities as durable governance patterns. Governance-level justificatory standards are therefore required for accountable infrastructural deployment. To accommodate this, the paper concludes that an additional normative framework is required to guide accountable decision-making.

DOI
https://doi.org/10.18690/um.fov.4.2026.50

ISBN
978-961-299-152-4

Keywords:
Value Sensitive Design, accountability, decentralized data sharing infrastructures, governance, technology deployment

University of Maribor Press

1 Introduction

Decentralized governmental data-sharing infrastructures (DDSIs) increasingly enable the combination of datasets across organizational and sectoral boundaries. Administrative records, mobility data and environmental datasets can be shared and integrated to support policy objectives such as climate mitigation, urban planning or fraud detection. For example, in urban mobility policy, combining transport and environmental data can support low-emission zones, while also raising questions about efficiency, privacy, and fairness, as certain groups may be disproportionately affected or surveilled. Such integrations both expand governmental capacity and introduce broader considerations about how value pluralism, power asymmetries, and institutional accountability become embedded in data practices. As such, they operate as socio-technical ecosystems of interdependent actors and institutional arrangements (Jacobides et al., 2018; Jacobs et al., 2021; Stahl, 2025). Therefore, socio-technical design- and deployment decisions need to address questions about proportionality, accountability, and acceptable risk beyond technical feasibility and legal permissibility.

Research in data protection governance shows that formal legal compliance, while necessary, is insufficient to secure accountable or legitimate outcomes in complex data-processing environments (Hendry et al., 2021; Stahl, 2025). Similarly, research on algorithm regulation highlights that legal compliance alone does not resolve broader concerns regarding legitimacy and accountability in complex data-processing environments (Yeung, 2018). The challenge concerns not the permissibility of data sharing as such, but the absence of structured ex ante procedures capable of aligning technological deployment with plural and possibly conflicting public values in a consistent and accountable manner.

What is at stake in deployment decisions is that technological systems in public governance do not merely implement policy objectives; they also institutionalize interpretations of accountability, proportionality, and acceptable risk (Bovens, 2007; De Reuver et al., 2020; Kaminski, 2019; Lüthi et al., 2021; Stahl, 2025). Decisions about the design and deployment of data-sharing infrastructures institutionalize normative commitments that shape how statutory objectives are pursued and how public authority is exercised. Technological design and organizational governance thus become inseparable once such commitments are formalized.

Against this background, procedural and deliberative theories of governance conceptualize legitimacy as dependent on the structuring of decision-making under conditions of values pluralism (Cenci & Cawthorne, 2020). Stahl (2025) notes that uncertainty about how values are realized in technological systems and how such systems are ultimately used complicates attempts to operationalize values in institutional practice. Within the Design for Values tradition, Value Sensitive Design (VSD) has emerged as a leading framework for integrating public values into the design of socio-technical systems (Friedman & Hendry, 2019; Friedman et al., 2006). VSD contributes to identifying and empirically exploring the public values at stake in technological systems.

This raises the question of whether VSD can guide justified institutional commitments where governance choices require prioritization and binding deployment decisions (Manders-Huits, 2011; Van de Poel, 2015). The issue becomes particularly salient in the deployment of governmental decentralized data-sharing infrastructures, where technical protocols predefine access conditions, computational logic, and accountability pathways. This article examines whether VSD can be applied ex ante to guide the design and deployment of decentralized data-sharing infrastructures in governmental ecosystems. The next section outlines the theoretical foundations of VSD in socio-technical governance.

2 Theoretical Foundations of VSD in Socio-Technical Governance

This section analyzes how VSD conceptualizes values, value tensions, and stakeholder concerns in relation to institutional decisions concerning technological deployment. VSD is a well-established theoretical and methodological approach for systematically addressing human values in the design of socio-technical systems (Friedman & Hendry, 2019; Friedman et al., 2006). A core premise underpinning this literature is that technologies are not value-neutral artefacts but actively mediate moral and social relations. Values such as privacy, autonomy, fairness, and trust are therefore treated as integral design concerns rather than as external constraints to be addressed post hoc.

Methodologically, the VSD approach integrates conceptual, empirical, and technical analyses to identify and articulate values that are implicated in technology design (Friedman & Hendry, 2019; Friedman et al., 2006). Rather than prescribing

substantive normative outcomes, VSD emphasizes sensitivity to context, stakeholder perspectives, and value pluralism. VSD primarily addresses ethical dynamics associated with anticipated engagement and interaction, rather than with the institutional decision-making that governs deployment of the system (De Reuver et al., 2020). The Design for Values literature explicitly acknowledges this analytical emphasis on anticipated use and interaction. Manders-Huits (2011) observes that VSD is particularly effective in uncovering value conflicts associated with technology use, while offering limited guidance for ethical questions that arise prior to institutional implementation. Recent work reinforces this limitation: Stahl (2025) notes that uncertainty about how values are realized in technological systems, and about how such systems will ultimately be used, complicates attempts to secure ethical outcomes through design alone.

Similarly, a systematic review of almost three decades of VSD research finds that while the framework integrates conceptual, empirical, and technical investigations, the practical implementation of technical investigations remains heterogeneous and requires further methodological development (Friedman et al., 2021; Gerdes & Frandsen, 2023). Decisions concerning the deployment of technological infrastructures may involve contextually competing normative demands or may reveal limited normative tension depending on the configuration of relevant variables. Ex ante deployment decisions establish the conditions under which use may proceed, including whether additional safeguards are required, whether specific configurations are permissible, or whether deployment should not occur at all. Hence, the use-oriented focus invites closer examination of whether VSD methods can inform accountable deployment decisions structured by explicit normative conditions, rather than remaining confined to analyzing anticipated use.

VSD methods are designed to render ethically salient aspects of socio-technical systems. Yet, as Van de Poel (2024) notes, "the emphasis has often mainly been on the early design phases, such as value elicitation, value specification and dealing with value conflicts, and less on the complexities and pitfalls of translating values into technological choices and realizing values in practice" (p. 308). This observation highlights the difficulty of moving from ethical articulation to binding institutional commitments. Although public values and accountability demands may become visible, VSD methods do not differentiate between varying degrees and configurations of normative tension. They therefore do not provide operational

criteria for structuring accountable deployment decisions, including whether processes should be automated, organized in hybrid arrangements, subjected to case-by-case assessment, or accompanied by additional protective measures aligned with the level of normative risk present.

This limitation has significant implications from a governance perspective. Cenci and Cawthorne (2020) argue that ethical legitimacy in complex socio-technical systems depends on procedural structures that enable values to be aggregated, justified, and translated into binding outcomes. In the absence of such procedures, ethical analysis remains reflective rather than action-guiding. This limitation does not reflect a defect in the VSD methods themselves. It reflects the boundary of its theoretical scope. Recognizing this boundary clarifies why a normative operational framework is required to structure accountable deployment conditions, particularly in decentralized data-sharing contexts.

The normative significance of deployment decisions becomes particularly pronounced in such infrastructures. Unlike stand-alone applications, decentralized data-sharing infrastructures stabilize patterns of coordination, automate or preconfigure decision logics, and embed assumptions about trust, control, and responsibility within their technical infrastructure. Once deployed, these infrastructures shape not only operational data flows but also the institutional environment within which legal norms, policy objectives, and accountability mechanisms are interpreted and enacted. As Kaminski (2019) observes, technological infrastructure can reconfigure the conditions under which legal and political norms operate, effectively relocating governance choices from explicit institutional decision-making to infrastructural design.

Deployment decisions concerning decentralized data-sharing infrastructures therefore constitute normative commitments that structure future action, constrain alternative governance arrangements, and affect the distribution of responsibility across actors. They do not merely implement prior design choices but institutionalize interpretations of public values, structuring the conditions under which later ethical reconsideration can occur. This dynamic aligns with the "problem of many hands" as formulated by Thompson (1980), in which responsibility becomes difficult to attribute within complex administrative systems. In governmental contexts, such

diffusion of responsibility poses challenges, as accountability relationships depend on clearly identifiable decision-makers and contestable justifications (Bovens, 2007).

Deployment choices therefore determine not only who can be held accountable, but also whether accountability mechanisms remain practically meaningful (Bovens, 2007). While VSD provides robust tools for identifying public value tensions, it does not by itself offer decision-oriented procedures for making and justifying binding deployment choices (Gerdes & Frandsen, 2023). Failure to differentiate ex ante between varying configurations of normative tension may entrench governance arrangements that are either unnecessarily restrictive or insufficiently safeguarded.

These limitations point to the need for a design-oriented methodological approach capable of translating value identification into structured ex ante deployment decisions aligned with the degree and configuration of normative tension within a specific context. The next section outlines the three steps to investigate what reference literature provides in addressing this need. First, we analyze how VSD can guide value trade-offs in institutional decision-making. Second, we examine how decentralized data-sharing infrastructures generate structural risks for accountability and governance in the absence of normative guidance for deployment decisions. Finally, we evaluate how the combination of value-sensitive design approaches and decentralized data-sharing technologies remains insufficient for guiding the binding of ex ante deployment decisions in governmental data-sharing contexts.

3 Literature review

This study follows an integrative literature review approach, guided by a semi-structured search strategy. Building on Snyder (2019), the literature review focuses on the theoretical state of knowledge, drawing on a qualitative conceptual analysis to examine how key concepts, values, stakeholder concerns, and justificatory arguments are articulated in the literature on Value Sensitive Design (VSD) and socio-technical governance.

The literature search and selection was conducted in three consecutive stages, corresponding to the analytical focus of the study. The process was iterative, with analytical insights informing subsequent inclusion decisions across three strands. The database-based screening and selection process is shown in Figure 1.

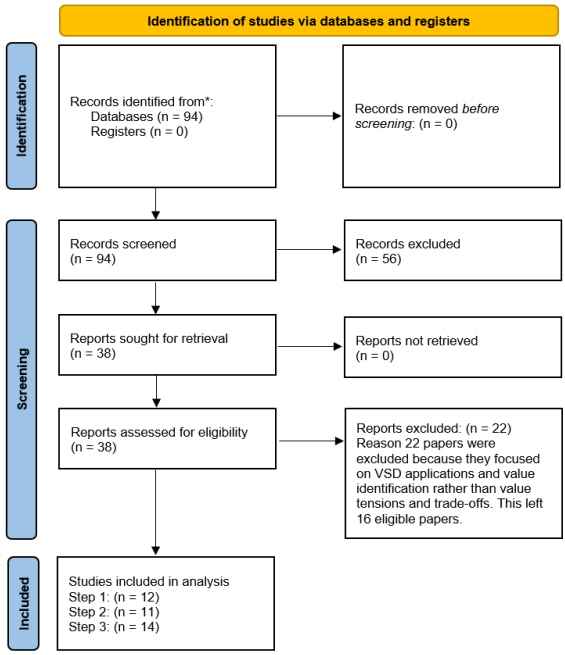

Figure 1: PRISMA Flow Diagram of the screening and selection process
Source: adapted from Page et al. (2021)

First, a search in Web of Science for "Value Sensitive Design" yielded 94 publications. Titles and abstracts were screened using analytically informed criteria for normative reflection on values, trade-offs and governance, resulting in a selection of 38 papers. A further assessment of normative relevance was conducted through full-text screening, focusing on the conceptualization of values, tensions, critiques of VSD, and institutional implementation. Papers were excluded if they primarily focused on technical design without normative reflection, on empirical applications without conceptual discussion, or on domains not directly related to socio-technical governance. This stage reduced the set to 16 papers. From this set, eight papers were identified as most directly relevant to the conceptual analysis of VSD, including the conceptualization of values, value trade-offs, and institutional considerations. The set was supplemented with four additional papers identified through theoretically overlapping sources and systematic snowballing. Snowballing involved backward and forward citation tracking via reference lists and Google Scholar. This process

aimed to identify conceptual contributions, critiques of VSD, and literature on governance and accountability in socio-technical systems and digital infrastructures.

Second, additional literature on governance, accountability, and socio-technical infrastructures was identified through targeted searches in Google Scholar, aligned with the second analytical focus of the study, using keyword combinations including "digital infrastructures", "data-sharing infrastructures", "public governance", "institutional commitments", and "responsibility allocation". Search terms were iteratively refined based on emerging conceptual themes and complemented by continued snowballing. Eleven papers informed this stage: four overlapped with the first set, two were identified through snowballing, and five had already been identified in the theoretical section.

Table 1: Summary of literature on analytical themes in VSD and socio-technical governance

Theme	Literature reference
Limitations of value-sensitive and procedural design approaches: Existing VSD approaches guide processes of value identification, specification, but do not establish how value trade-offs should be resolved in institutional decision-making.	Bovens (2007); Cenci and Cawthorne (2020); Friedman et al. (2006); Friedman et al. (2020); Friedman et al. (2021); Hendry et al. (2021); Jacobs and Huldtgren (2021); Manders-Huits (2011); Martin et al. (2023); Stahl (2025); Umbrello (2020); Van de Poel (2024)
Governance and accountability risks in decentralized data-sharing infrastructures: Literature on governance highlights how accountability, responsibility, and decision-making are distributed across socio-technical systems, where algorithmic regulation and governance by design shape institutional outcomes and create the problem of many hands.	Bovens (2007); Jacobs et al. (2021); Kaminski (2019); Lindell (2021); Löhr and Veluwenkamp (2025); Lüthi et al. (2021); Selbst et al. (2019); Stahl (2025); Thompson (1980); Umbrello (2020); Yeung (2018)
Absence of a structured mechanism for resolving value trade-offs: Across both strands of literature, there is no systematic approach for translating values into binding institutional commitments, resulting in unresolved value trade-offs in contexts where infrastructural design shapes governance outcomes.	Bovens (2007); Cenci and Cawthorne (2020); Jacobs and Huldtgren (2021); Jacobs et al. (2021); Jacobides et al. (2018); Kaminski (2019); Löhr and Veluwenkamp (2025); Lüthi et al. (2021); Manders-Huits (2011); Martin et al. (2023); van Ooijen et al. (2019); Selbst et al. (2019); Stahl (2025); Yeung (2018)

The final analysis synthesizes findings from fourteen papers selected through this iterative review process across the three strands. The analysis involved identifying and comparing recurring themes in the literature, including limitations of value-sensitive design approaches, governance and accountability risks, and the absence of

structured mechanisms for resolving value trade-offs. The key analytical themes identified through this process are summarized in Table 1.

3.1 Value Sensitive Design: the justification of value trade-offs

According to Hendry et al. (2021), VSD is characterized as a formative framework rather than a prescriptive ethical theory (Friedman et al., 2006). Unlike comprehensive ethical theories, it does not specify which values are binding in cases where commitments must be determined ex ante, how value conflicts ought to be resolved, or according to which justificatory standards such commitments can be justified (Friedman et al., 2021). Across the critical literature, three interrelated methodological limitations recur, which together point to a justificatory gap in VSD when applied to institutional decision-making: (i) the absence of explicit ethical grounding; (ii) the reliance on empirical convergence in value identification; and (iii) the absence of structured mechanisms for resolving value tensions.

First, the absence of ethical grounding implies that VSD does not articulate a principle determining when identified values generate binding obligations or how they ought to be hierarchized (Cenci & Cawthorne, 2020; Jacobs & Huldtgren, 2021; Manders-Huits, 2011). Second, the reliance on empirical convergence entails that stakeholder endorsement informs value identification within VSD processes (Friedman et al., 2020; Van de Poel, 2024). This does not justify why certain values prevail in conflicts (Jacobs & Huldtgren, 2021; Manders-Huits, 2011). Third, VSD lacks a systematic method for resolving value tensions and trade-offs (Cenci & Cawthorne, 2020). Without such standards, conflicts remain underdetermined at binding decision points, thereby weakening accountability (Bovens, 2007).

3.2 Structural risks in decentralized data-sharing infrastructures

According to Yeung (2018) and Kaminski (2019), decentralized data-sharing infrastructures serve as instrumental governance systems for the coordination, validation, and enforcement of data-driven policies or operational procedures. Any assessment of accountability must be sensitive to the interaction between infrastructural form and policy stakes involved. The same architectural design may appear functionally neutral in one context yet determine binding institutional outcomes in another. Bovens (2007) conceptualizes accountability as an "evaluative

concept" structuring the institutionalized relationship between an actor and a forum involving explanation, justification, and the consequences (p. 450). Design abstractions distribute responsibility among actors and contexts and embed institutional presumptions that affect the accountability conditions (Löhr & Veluwenkamp, 2025; Stahl, 2025; Umbrello, 2020).

Socio-technical governance scholarship identifies several structural risks that arise when governance functions are embedded in decentralized data-sharing infrastructures. These risks concern (i) shifts in justificatory assessment; (ii) the relocation of discretion to infrastructural design; (iii) abstraction effects operating at scale; (iv) the treatment of instrumental robustness as sufficient justification for deployment reflecting the distinction between instrumental and justificatory governance (Kaminski, 2019); and (v) the diffusion of responsibility across ecosystem actors. Table 2 provides an overview of how literature conceptualizes the limits of VSD (Section 3.1) alongside the risks of deploying decentralized data-sharing infrastructures (Section 3.2).

3.3 VSD limits in deploying decentralized data-sharing infrastructures

Taken together, VSD's justificatory limitations and the structural risks of decentralized data-sharing infrastructures interact and produce institutional effects. VSD operates as a formative framework that lacks determination justificatory criteria, while decentralized infrastructures embed decision logics within architectural configurations operating continuously at scale. When these dynamics converge, justificatory indeterminacy becomes stabilized within governance arrangements. VSD remains necessary for identifying values but is insufficient for guiding binding ex ante deployment decisions. A formative framework supports value identification but does not stabilize value trade-offs in governance architectures operating at scale.

Where decentralized infrastructures transform design choices into embedded and distributed institutional arrangements, justificatory underdetermination becomes institutionally operative. The fragmentation of accountability within ecosystems may not merely obscure responsibility but amplify systemic risks once decision logic becomes infrastructurally entrenched. When formative normativity is coupled with decentralized deployment, VSD cannot provide sufficient justification for binding

ex ante decisions. Table 3 summarizes how the justificatory limitations of VSD (Section 3.1) interact with the structural risks of decentralized deployment (Section 3.2) to produce institutional effects (Section 3.3).

Table 2: Literature mapping across the limitations of VSD and the risks of deploying DDSIs

Structural characteristic	Analytical finding	Governance implication	Associated literature
The absence of explicit ethical grounding	VSD does not specify when values generate binding institutional obligations	Normative justification of deployment decisions remains underdetermined	Cenci and Cawthorne (2020); Friedman et al. (2006); Hendry et al. (2021); Jacobs and Huldtgren (2021); Manders-Huits (2011); Martin et al. (2023); Stahl (2025); Umbrello (2020)
The reliance on stakeholder perspectives to identify values	Stakeholder agreement does not resolve normative priority in institutional conflicts	Institutional commitments may reflect stakeholder preference rather than normative justification	Cenci and Cawthorne (2020); Friedman et al. (2006); Friedman et al. (2020); Friedman et al. (2021); Hendry et al. (2021); Jacobs and Huldtgren (2021); Manders-Huits (2011); Martin et al. (2023); Umbrello (2020); Van de Poel (2024)
The absence of structured conflict resolution	VSD does not provide a method for resolving value trade-offs	Value trade-offs cannot be stabilized in governance architectures	Bovens (2007); Cenci and Cawthorne (2020); Friedman et al. (2021); Hendry et al. (2021); Jacobs and Huldtgren (2021); Manders-Huits (2011); Stahl (2025); Van de Poel (2024)
The embedding of justification in infrastructural design	Justificatory assessment is displaced from institutional deliberation to infrastructural design	Technical architecture determines normative commitments	Bovens (2007); Kaminski (2019); Löhr and Veluwenkamp (2025); Stahl (2025); Thompson (1980); Umbrello (2020); Yeung (2018)
The contextual misalignment between infrastructure and policy objectives	Operational discretion is embedded in design choices rather than institutional procedures	Design choices determine how institutional categories are applied in practice	Jacobs et al. (2021); Yeung (2018)
The abstraction effects of operating systems at scale	Design abstraction determines the structuring of distributive outcomes across governance contexts	Structuring shapes distributive outcomes in high-stakes policy domains	Jacobs et al. (2021); Lüthi et al. (2021); Selbst et al. (2019); Stahl (2025); Yeung (2018)
The relocation of discretion to design-level specification	Technical robustness substitutes normative justification in deployment decision	Procedural correctness replaces normative scrutiny of policy objectives	Bovens (2007); Kaminski (2019); Lindell (2021); Yeung (2018)
The diffusion of responsibility across ecosystem actors	Responsibility is fragmented across ecosystem actors, obscuring accountability	Accountability shifts to infrastructural design and configuration	Bovens (2007); Jacobs et al. (2021); Yeung (2018)

Table 3: Interplay of VSD limitations and structural risks in decentralized data systems

VSD limitation (Section 3.1)	Structural risk (Section 3.2)	Institutional effect (Section 3.3)
The absence of explicit ethical grounding	The embedding of justification in infrastructural design	Technical robustness may substitute for normative justification of deployment decisions
	The contextual misalignment between infrastructure and policy objectives	Deployment legitimacy may remain insufficiently scrutinized
The reliance on stakeholder perspectives to identify values	The abstraction effects of operating systems at scale	Underdetermined value choices become embedded in eligibility and classification structures
The absence of structured conflict resolution	The relocation of discretion to design-level specification	Value trade-offs become embedded in system architecture rather than resolved through institutional deliberation
	The diffusion of responsibility across ecosystem actors	Accountability becomes fragmented across organizational and technical layers

4 Discussion of findings

This paper contributes to debates in design ethics, algorithmic governance, and digital ecosystem theory by showing that formative ethical approaches reach institutional limits once embedded in infrastructural governance arrangements. Building on the three analytical themes identified in the literature review, we identify three contributions. First, we refine existing critiques of VSD. Prior scholarship highlights VSD's pluralism, and reliance on stakeholder perspectives and lack of explicit ethical grounding (Cenci & Cawthorne, 2020; Jacobs et al., 2021; Jacobs & Huldtgren, 2021; Löhr & Veluwenkamp, 2025; Stahl, 2025). We situate these concerns within infrastructural governance contexts and show that they become structurally consequential when embedded in infrastructures. Justificatory indeterminacy arises when value commitments are embedded in infrastructures, where instrumental robustness may substitute for publicly defensible normative justification (Kaminski, 2019). This clarifies the conditions under which VSD becomes normatively insufficient for deployment decisions.

Second, we extend socio-technical accounts of abstraction and scaling. Prior work shows how abstraction shapes institutional categories and embeds governance in computational architectures (Jacobs et al., 2021; Selbst et al., 2019). We show that

abstraction does not merely affect decision quality but stabilizes distributive outcomes within infrastructural arrangements. This reduces contestability and amplifies the persistence of underdetermined value trade-offs (Jacobs et al., 2021; Stahl, 2025).

Third, we shift the normative focus from artefact-level ethics to governance-level norm stabilization (Jacobides et al., 2018; Stahl, 2025). In digital ecosystems, normative outcomes emerge through interactions across actors rather than through isolated design choices. Artefact-level ethical design is therefore insufficient, and governance-level mechanisms are required to stabilize value priorities across institutional layers (Kaminski, 2019). In decentralized ecosystems, these dynamics are intensified by distributed responsibility, infrastructural rigidity, and limited ex post correction (Jacobs & Huldtgren, 2021; Stahl, 2025). We therefore argue for a governance-level normative framework for ex ante deployment decisions of decentralized data-sharing infrastructures.

5 Conclusions

This study examined whether VSD provides sufficient normative guidance for deploying decentralized data-sharing infrastructures in governmental ecosystems. Consistent with the analytical themes identified in Table 1, we identify three findings. When justificatory indeterminacy intersects with infrastructural stabilization, underdetermined value commitments become embedded as binding institutional arrangements operating at scale. Formative value articulation is thereby translated into governance architecture without prior normative fixation. VSD remains necessary but is structurally insufficient for guiding binding ex ante deployment decisions. Where infrastructures reorganize accountability and embed decision logics, deployment decisions cannot rely on value identification alone. Governance-level justification standards are required to stabilize value trade-offs before they are institutionalized through architectural design.

Future empirical research is required to operationalize governance-level justification criteria and design deployment review mechanisms capable of anticipating distributive and systemic effects in decentralized ecosystems. This should support normative frameworks with clear governance standards for ex ante deployment decisions, for example through impact assessments in specific ecosystem contexts.

References

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal*, *13*(4), 447-468.

Cenci, A., & Cawthorne, D. (2020). Refining Value Sensitive Design: A (Capability-Based) Procedural Ethics Approach to Technological Design for Well-Being. *Science and Engineering Ethics*, *26*(5), 2629-2662. https://doi.org/10.1007/s11948-020-00223-3

De Reuver, M., van Wynsberghe, A., Janssen, M., & van de Poel, I. (2020). Digital platforms and responsible innovation: expanding value sensitive design to overcome ontological uncertainty. *Ethics and Information Technology*, *22*(3), 257-267. https://doi.org/10.1007/s10676-020-09537-z

Friedman, B., Harbers, M., Hendry, D. G., van den Hoven, J., Jonker, C., & Logler, N. (2021). Eight grand challenges for value sensitive design from the 2016 Lorentz workshop. *Ethics and Information Technology*, *23*(1), 5-16. https://doi.org/10.1007/s10676-021-09586-y

Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design*. The MIT Press.

Friedman, B., Kahn, P., Borning, A., Zhang, P., & Galletta, D. (2006). Value Sensitive Design and Information Systems. In. https://doi.org/10.1007/978-94-007-7844-3_4

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2020). Value Sensitive Design and Information Systems. *The Ethics of Information Technologies*.

Gerdes, A., & Frandsen, T. F. (2023). A systematic review of almost three decades of value sensitive design (VSD): what happened to the technical investigations? *Ethics and Information Technology*, *25*(2), 26. https://doi.org/10.1007/s10676-023-09700-2

Hendry, D. G., Friedman, B., & Ballard, S. (2021). Value sensitive design as a formative framework. *Ethics and Information Technology*, *23*(1), 39-44. https://doi.org/10.1007/s10676-021-09579-x

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*, *39*(8), 2255-2276.

Jacobs, M., Kurtz, C., Simon, J., & Böhmann, T. (2021). Value Sensitive Design and power in socio-technical ecosystems. *Internet Policy Review*, *10*(3). http://dx.doi.org/10.14763/2021.3.1580

Jacobs, N., & Huldtgren, A. (2021). Why value sensitive design needs ethical commitments. *Ethics and Information Technology*, *23*(1), 23-26. https://doi.org/10.1007/s10676-018-9467-3

Kaminski, M. E. (2019). Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability. *Southern California Law Review*, *92*(6), 1529 - 1616.

Lindell, Y. (2021). Secure multiparty computation. *Communications of the ACM*, *64*(1), 86-96. https://doi.org/10.1145/3387108

Löhr, G., & Veluwenkamp, H. (2025). Rethinking philosophical methodology: Conceptual engineering meets Value Sensitive Design. *Metaphilosophy*, *56*(3-4), 328-343. https://doi.org/https://doi.org/10.1111/meta.12730

Lüthi, N., Matt, C., & Myrach, T. (2021). A value-sensitive design approach to minimize value tensions in software-based risk-assessment instruments. *Journal of Decision Systems*, *30*(2-3), 194-214. https://doi.org/10.1080/12460125.2020.1859744

Manders-Huits, N. m. (2011). What Values in Design? The Challenge of Incorporating Moral Values into Design. *Science and Engineering Ethics*, *17*(2), 271-287. https://doi.org/10.1007/s11948-010-9198-2

Martin, D. A., Clancy, R. F., Zhu, Q., & Bombaerts, G. (2023). Why do we Need Norm Sensitive Design? A WEIRD Critique of Value Sensitive Approaches to Design. *Global Philosophy*, *33*(4), 40. https://doi.org/10.1007/s10516-023-09689-9

Page M.J., McKenzie J.E., Bossuyt P.M., Boutron I., Hoffmann T.C., Muldow C.D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ;372:n71* https://doi.org/10.1136/bmj.n71

Selbst, A. D., Boyd, D. F., S.;, Venkatasubramanian, S.;, & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. Proceedings of the Conference on Fairness, Accountability, and Transparency, New York, NY, USA.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333-339. https://doi.org/https://doi.org/10.1016/j.jbusres.2019.07.039

Stahl, B. C. (2025). A framework for value sensitive design in socio-technical digital ecosystems. *European Journal of Information Systems*, 1-19. https://doi.org/10.1080/0960085X.2025.2595915

Thompson, D. F. (1980). Moral Responsibility of Public Officials: The Problem of Many Hands. *American Political Science Review*, *74*(4), 905-916. https://doi.org/10.2307/1954312

Umbrello, S. (2020). Imaginative Value Sensitive Design: Using Moral Imagination Theory to Inform Responsible Technology Design. *Science and Engineering Ethics*, *26*(2), 575-595. https://doi.org/10.1007/s11948-019-00104-4

Van de Poel, I. (2015). Conflicting values in design for values. In P. V. J van den Hoven, I van de Poel (Ed.), *Handbook of ethics, values, and technological design: sources, theory, values and application domains* (pp. 89-116). Springer.

Van de Poel, I. (2024). Design for Value. In A. Grunwald (Ed.), *Handbook of Technology Assessment* (pp. 301-309). Edward Elgar Publishing https://doi.org/https://doi.org/10.4337/9781035310685

van Ooijen, C., Ubaldi, B., & Welby, B. (2019). *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance*. O. W. P. o. P. Governance.

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, *12*(4), 505-523. https://doi.org/https://doi.org/10.1111/rego.12158

Tenth Anniversary of eHealth at the National Institute of Public Health

Živa Rant, Eva Štupica, Jure Janet, Miha Knavs, Lucija Tepej Jočić, Matjaž Malešič, Eva Hafner, Boško Simeunović, Dijana Čančarević, Simon Indihar, Hajdi Kosednar

National Institute of Public Health, Ljubljana, Slovenia
ziva.rant@nijz.si, eva.stupica@nijz.si, jure.janet@nijz.si, miha.knavs@nijz.si, lucija.tepej-jocic@nijz.si, matjaz.malesic@nijz.si, eva.hafner@nijz.si, bosko.simeunovic@nijz.si dijana.cancarevic@nijz.si, simon.indihar@nijz.si, hajdi.kosednar@nijz.si

DOI
https://doi.org/10.18690/um.fov.4.2026.51

ISBN
978-961-299-152-4

In 2015, the National Institute of Public Health of the Republic of Slovenia (NIJZ) assumed responsibility for managing the solutions of eHealth project. Since then, eHealth services have been continuously upgraded and improved. The use has steadily expanded, as has the range of content and functions available to patients and healthcare professionals. Despite the challenges NIJZ faced in introducing innovative solutions at the national level, the services have been successfully implemented. Today, eHealth services play a crucial role in supporting healthcare providers and facilitating electronic access to health records for the entire population. The most significant achievement of the Slovenian eHealth is its full population coverage, making it one of the few such examples worldwide. This paper describes the main characteristics and achievements of the eHealth solutions, including ePrescription, the zVEM patient portal, the zVEM+ healthcare professionals portal, eAppointment, the Central Registry of Patient Data, the eHeath Users Database, Teleradiology and TeleStroke. It also presents data on how these services are used.

Keywords:
National Institute of Public Health of the Republic of Slovenia, Healthcare digitalisation, ePrescription, zVEM patient portal, eAppointment, Central Registry of Patient Data

University of Maribor Press

1 Introduction

The eHealth project began in 2008 and ran until November 2015. It was based on the document eHealth2010 – Strategy for Informatization of the Slovenian Healthcare System 2005–2010 (2005), issued by the Project Unit of the Centre for Healthcare Informatics within the Ministry of Health of the Republic of Slovenia. The eHealth project was one of the largest public service sector digitalisation initiatives in Slovenia. It was led by the Ministry of Health of the Republic of Slovenia and co-financed by the European Social Fund. On 1 December 2015, the National Institute of Public Health of the Republic of Slovenia (NIJZ) assumed responsibility for managing the components of the eHealth system, while the Ministry of Health provided funding for its maintenance and further development. Since then, eHealth services have been continuously upgraded and developed. Use has steadily expanded, as has the range of content and functions available to patients and healthcare professionals. All components of the eHealth system operate at the national level and are used by all healthcare providers in Slovenia. The successful operation of the system is underpinned by two key elements: the Healthcare Databases Act – ZZPPZ (2015) and the Health Insurance Card Number, which provides a unique identifier for patients across all healthcare databases.

Despite the challenges NIJZ faced in introducing innovative solutions at the national level, they have been successfully implemented. Today, these eHealth services play a key role in supporting healthcare providers and enabling the entire population to access their health records electronically. According to the European Commission's Digital Economy and Society Index (DESI) 2024, Slovenia ranks eighth in the EU for access to electronic health records. The overall eHealth maturity score is 88%, compared to the EU Member State average of 79% (European Commission et al., 2024).

In 2017, ePrescription received the Award "Informacijska jagoda" for the best achievement in the field of the information society. In 2022, the zVEM portal received the Award for Current Achievements in the field of information society, and the zVEM mobile app received the "eNagrada" award.

2 Methodology

This paper analyses the functionality and use of the eHealth system to mark the 10th anniversary of eHealth at NIJZ. The research question is What are the eHealth solutions, and how have they evolved and been used over the past ten years?

A literature review of the relevant literature in this field was conducted (Rant & Rant, 2018; Rant, Kosednar, et al., 2025; Rant & Stanimirović, 2019, 2020, 2021, 2024a, 2024b; Rant et al., 2021, 2022a, 2022b; Rant et al., 2017; Rant et al., 2018; Rant et al., 2019; Rant et al., 2025; Stanimirović et al., 2022a, 2022b; Kosednar, 2025; Tepej Jočić, 2024, Janet & Štupica, 2025; Plevnjak & Stanimirović, 2024).

The study applied a case study research methodology (Kljajić Borštnar, 2022; Yin, 2018), which included an in-depth examination and analysis of the field. Project documentation, user guides and technical specifications were reviewed. An observational approach was used for in-depth examination, drawing on the experience and expertise of NIJZ experts responsible for managing the eHealth system. Usage statistics were retrieved from the business and administrative modules of the system and analysed.

The system's functionality and use were analysed between November 2025 and January 2026. Statistics from the business and administrative modules were retrieved and analysed in February 2026.

The observational approach applied in this study has several important limitations that should be explicitly acknowledged. The analysis relies on the direct experience and expertise of NIJZ experts, who are simultaneously system operators and researchers. This dual role may introduce potential bias, including selective interpretation of system performance, positive reporting bias, and limited critical distance from the observed processes. Furthermore, as the observations are not based on independent external evaluation, there is a risk that certain challenges or system limitations may be underreported or interpreted in a more favourable light.

In addition, the absence of a formalised observational protocol and the reliance on internal knowledge may affect the reproducibility and objectivity of the findings. While the approach provides valuable insider insights into system functionality and

development, the results should therefore be interpreted with caution and understood as context-dependent. Future research would benefit from the inclusion of independent evaluators and complementary methodological approaches to mitigate these limitations and strengthen the validity of the findings.

3 Results

3.1 Key solutions of the Slovenian eHealth system

The eHealth system includes twenty interconnected solutions. The most notable solutions are ePrescription, eAppointment, Central Registry of Patient Data, zVEM patient portal and zVEM+ healthcare professionals portal.

3.2 ePrescription

ePrescription was the first eHealth solution to be rolled out at national level. It was introduced at primary level throughout Slovenia in November 2015 and at secondary and tertiary levels in February 2016. It is used daily by healthcare providers and pharmacies and proved particularly important in ensuring patient access to medicines during the COVID-19 pandemic.

ePrescription is a national IT solution for the secure electronic prescription prescribing and dispensing of medicines. Electronic prescriptions (e-prescriptions) are generated in healthcare providers' (HCPs) local information systems and stored in the central electronic prescription record, which provides pharmacies with the information needed to dispense medicines to patients on the basis of e-prescriptions. Patients can access this information via the zVEM patient portal ("zVEM Portal", 2025). The portal displays prescribed medicines, including the prescriber and the date of prescription. It also displays medicine dispensations and how many remain for renewable prescriptions. Patients can collect medicines from the same prescription at different pharmacies across Slovenia.

In the first month of operation (December 2015) 772,217 e-prescriptions were issued and 1275 providers used ePrescription. The total number of prescriptions issued per year increased until 2019 and remained largely unchanged between 2019 and 2021. Since 2022, the total number of prescriptions has increased. In 2024, the

total number exceeded 16,000,000, and the same level was recorded in 2025 (Figure 1). ePrescription was used by 1500 healthcare providers; e-prescriptions were issued in 1092 organisations, and medicines were dispensed by 356 pharmacies.

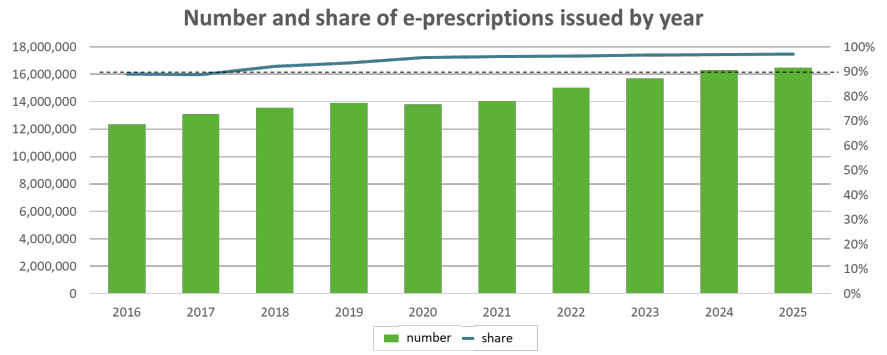

Figure 1: Number and share of e-prescriptions issued by year
Source: Own

The ePrescription objective was to have 90% of prescriptions prescribed electronically. Since 2020, more than 96% of prescriptions have been issued electronically, reaching 97% in 2025. A rate of 100% will never be achieved due to prescribing regulations, as some prescriptions must still be issued on paper, including emergency prescriptions, prescriptions issued during home visits, doctors' prescriptions for their own use, and prescriptions containing narcotic or psychotropic drugs (Rant & Stanimirović, 2024a).

The ePrescription solution represents good practice in the digitalisation of healthcare in Slovenia. During the COVID-19 pandemic, it ensured uninterrupted access to medicines, as doctors could prescribe medicines remotely and patients could collect them at any pharmacy in Slovenia using their health insurance card. This is also reflected in the successful use of this component: more than 97% of prescriptions are issued electronically, placing Slovenia among the global leaders.

In 2024, access to view e-prescriptions was extended to nurses. In 2025, a function was introduced to check medicines for substances prohibited for athletes, based on the List of Prohibited Substances and Methods for Athletes maintained by the Slovenian Anti-Doping Organisation (SLOADO). Further adjustments are planned

to meet new requirements introduced primarily by the Act Amending the Medicines Act (ZZdr-2), the Regulations on Classification, Prescribing and Dispensing of Medicines for Human Use, and the Healthcare Digitalisation Act *(Healthcare Digitalisation Act (ZDigZ)*, 2025).

3.3 zVEM Patient Portal

For patients, the most well-known solution of the system is the zVEM patient portal (the acronym zVEM means "health all in one place"). It provides secure and reliable access to data stored in the eHealth system and the Health Insurance Institute of Slovenia (HIIS) database via a single access point. From the patient's perspective, it represents an important milestone in the history of Slovenian healthcare and one of the most significant advances in recent decades. During the COVID-19 pandemic, it played an important role in informing the public and raising awareness.

The zVEM portal can be accessed using the national SI-PASS authentication service with suitable means of electronic identification, such as a qualified SIGEN-CA digital certificate, the smsPASS mobile identity, a biometric identity card via the eOsebna mobile application, or another type of digital identity. The portal was technically implemented at the completion of the eHealth project in November 2015. Full functionality, including user registration, has been available since the beginning of 2017. The zVEM mobile app was developed to meet the growing needs of users and has been in use since July 2021.

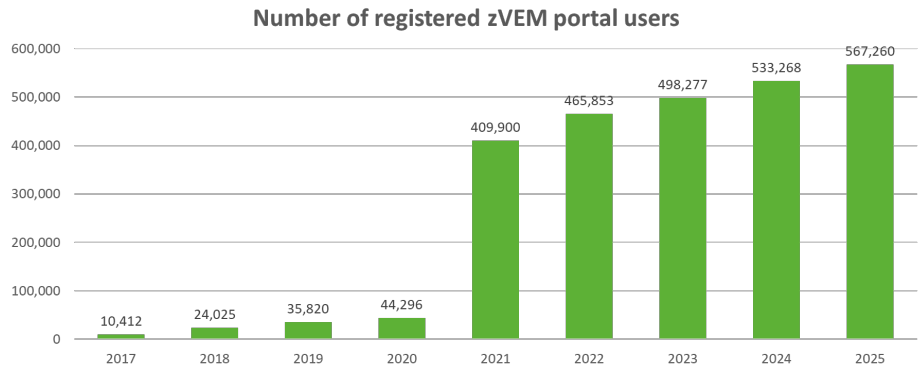

Figure 2: Number of registered zVEM portal users
Source: Own

By the end of 2025, around 570,000 users had registered of whom almost 380,000 used the mobile app. Including linked persons (i.e. children under 15, persons under guardianship and represented persons who authorized another person to access their data), the total number of users is around 740,000. Around 120,000 unique visits are recorded monthly. To facilitate the transition to digital services and assist the elderly in arranging digital identities and authorizations, the Ministry of Health, in cooperation with NIJZ, Ministry of Digital Transformation and Ministry of Public Administration, has established zVEM digi-points throughout Slovenia (*Digitočke za portal zVem*, 2026).

The portal developed significantly during the COVID-19 pandemic. During this period, the zVEM portal was used to establish the national vaccination appointment system, issue digital COVID certificates, and record COVID-19 test results for providers without their own information systems. The faster growth in the number of registered users in 2021 can be attributed to the introduction of the option to print EU Digital COVID Certificate in June 2021 and the launch of the zVEM mobile phone app in July 2021, when the number of users increased more than ninefold (Figure 2).

In 2025, the range of functionalities and available content expanded significantly, particularly as part of the Telehealth and Telemedicine project, funded by the EU through the Recovery and Resilience Plan. Patients can receive notifications via push notifications, emails and SMS. The portal also allows users to enter contact details, which are required for patient notifications and the national rollout of electronic communication with care teams.

The portal also provides access to HIIS content: electronic sick leave certificates (eBOL) and electronic blood donation certificates (ePODK), information on selected personal physicians, compulsory insurance status and insurance history, the costs of medicines and services, medical devices, cash benefits, the validity of the European Health Insurance Card, and A1 certificates.

A patient may authorize a more digitally skilled person to access their health data and perform most services on the zVEM portal, which significantly increases the accessibility of eHealth for vulnerable patient groups (Janet & Štupica, 2025). Another major innovation is the introduction of electronic communication with

healthcare teams through the zVEM portal, enabling patients to communicate with their doctors in a uniform way regardless of which provider within the public healthcare network they belong to.

3.4 National zVEM+ system

The national zVEM+ solution was launched in 2021 and is intended exclusively for professional users, i.e. staff working at healthcare providers. It enables the collection and processing of data, the submission of various reports, and queries based on basic patient demographics. During the COVID-19 pandemic, particularly important functions included the issuance of digital COVID certificates and the entry of COVID-19 test results into the Central Registry of Patient Data (CRPD). It also provides NIJZ, as the system operator, with access to data for public health activities, e.g. such as the continuous collection of mortality data (Rant et al., 2022a).

The Personal Medication Card (PMC) is one of the most widely used zVEM+ features. It is a document prepared by pharmacists in the zVEM+ or pharmacy apps. In doing so, they also use ePrescription data. The document has a standardised format and is intended for patients taking multiple medicines. It contains a list of the patient's medicines and instructions for their use. The PMC is available to patients electronically via the zVEM portal. The PMC has been available since March 2023; our data show a steady increase in the number of PMCs created, indicating that the function is being used increasingly by healthcare providers and patients (Janet & Štupica, 2025).

The zVEM+ portal also supports the Registry of Rare Non-Malignant Diseases (RRNB), managed by the Division of Paediatrics at the University Medical Centre Ljubljana.

Since 2025, zVEM+ has also been used as a communication hub for secure electronic communication between patients and their care teams, and between selected personal physicians and specialists regarding their patients. Through the zVEM portal, patients can contact their general practitioner, dentist, paediatrician or gynaecologist. In the future, communication will also be enabled with other care teams involved in patient care, such as registered nursing teams, community nursing services and specialist teams providing care for referred patients.

The zVEM+ portal includes care team management, which is a prerequisite for introducing targeted communication and other advanced features for team members involved in patient care. It allows the entering of zVEM authorisations in the CRPD.

3.5 eAppointment

eAppointment is a national information system for electronic referrals and appointment scheduling for healthcare services from primary to secondary and tertiary care, as well as within secondary and tertiary care. It is also used to monitor waiting times and waiting lists. The eAppointment system was introduced nationwide in 2016. On 10 April 2017, the Ministry of Health (MoH) placed electronic and paper referrals on an equal footing.

The eAppointment information system comprises three components: referrals, appointment scheduling and waiting lists. In their local information system, the doctor generates an electronic e-referral document (which may be either an e-referral or an e-work order), signs the document electronically and transmits it to the central eAppointment system. Appointments for referred healthcare services are scheduled on the basis of the e-referral document.

The system enables online appointment scheduling. Appointments may be scheduled by the doctor who issued the e-referral, by a nurse, or at the provider's infopoint. Patients can schedule appointments via the zVEM portal, via the eAppointment portal, or with the assistance of a call centre adviser. Patients may still contact the healthcare provider with whom they wish to schedule the selected healthcare service, as was possible before the introduction of eAppointment – by phone, email, regular mail or in person. In such cases, the referring provider (specialist) performing the service schedules the appointment.

The eAppointment system displays waiting lists or the first available appointment slots for many healthcare services, defined in the *Healthcare Service Code List* (*Šifrant vrst zdravstvenih storitev* – VZS). Waiting lists are reported for a defined set of healthcare services. In version 18.1 of the code list, the total number of healthcare service codes (VZS) is 2541. Waiting list monitoring is available for 1857 VZS. Patients can schedule appointments online themselves for 669 active (valid) VZS. Healthcare providers are responsible for reporting accurate data of available slots

and appointments. The central system displays the data reported by individual healthcare providers.

Each month, more than 400,000 e-referrals are issued and more than 700,000 e-appointments are scheduled. Since the introduction of eAppointment, more than 37 million e-referrals have been issued, more than 1200 healthcare providers deliver at least one service through the system.

Following the decline in 2020 and 2021, the number of issued e-referrals has increased again, exceeding 4 million per year. Since 2020, e-referrals have accounted for at least 95% of all referrals. In 2025, 97% of referrals were issued electronically (Figure 3).

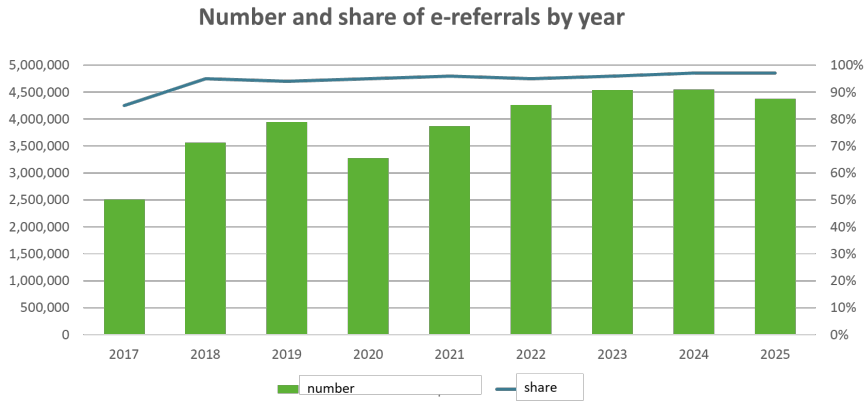

Figure 3: Number and share of e-referrals by year
Source: Own

Since August 2024, referrals have been valid until the completion of treatment, meaning that patients no longer need additional referral documents for diagnostics, follow-up examinations or procedures related to the same medical condition and type of healthcare service.

Push notifications in the zVEM mobile app allow patients to receive notifications from eAppointment about issued referral documents and their scheduled appointments. In this way, patients are informed in a timely manner about their

specialist appointments, which reduces missed appointments and raises awareness of patients' responsibilities.

The Rules on Appointment Scheduling and Management of Waiting Lists and Maximum Permissible Waiting Times (2024) introduce a new definition of a patient waiting beyond the maximum permissible waiting time (Kosednar, 2024). As a result, the methodology for calculating patients waiting beyond the maximum permissible waiting time had to be revised and implemented in the eAppointment system.

Since 1 September 2025, the eAppointment system has enabled the issuing and receipt of electronic work orders for physiotherapy. The gradual introduction of e-work orders for other services (for example community nursing and others) is also planned. With the physiotherapy e-work order, patients can schedule appointments with a physiotherapist online. First available slots for physiotherapy services are reported, and a national waiting list for patients has also been established.

Data from the first months after its introduction indicate an increase in the use of e-work orders. In September 2025, approximately 18,500 physiotherapy e-work orders were issued, rising to around 25,000 per month by January 2026. This trend shows a gradual transition to the electronic version and increasing integration of healthcare providers into the national appointment reporting system.

The eConsultation service has also been implemented as part of the eAppointment system. It enables electronic communication between healthcare professionals, specifically between a general practitioner and a specialist, allowing the general practitioner to obtain the specialist's opinion on a clinical question. The eConsultation service is available for all types of healthcare services.

3.6 Central Registry of Patient Data (CRPD)

The Central Registry of Patient Data is a unified system for collecting and exchanging health data on patients receiving healthcare in the Republic of Slovenia. The purpose of the CRPD is to enable the electronic exchange of health data between healthcare providers so that the data are available during patient care.

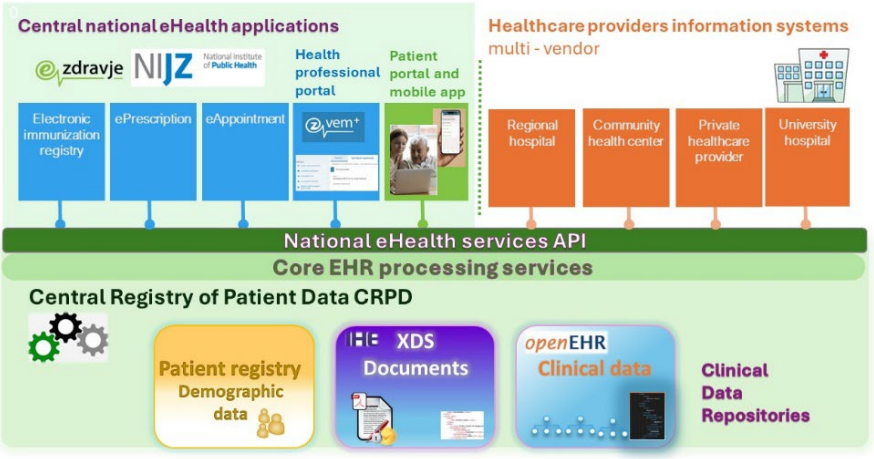

Figure 4: Core components of the CRPD
Source: Own

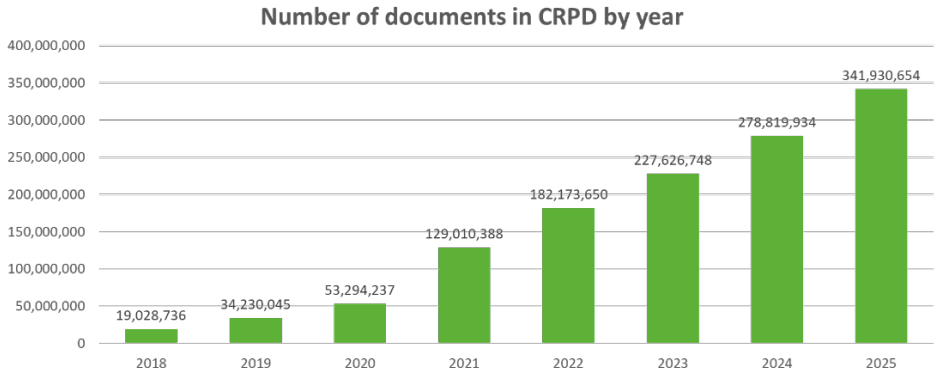

Figure 5: Number of documents in CRPD by year
Source: Own

The CRPD comprises three core system components (Figure 4). The Registry on Patient Demographics and Status data contains basic personal information such as name, surname, address, and health insurance details (Tepej Jočić, 2024). The document system contains electronic health records (e.g. pdf or xml documents). The clinical database contains structured health records in a format suitable for access and reuse. Access to the CRPD is restricted by technical mechanisms defined on the Rules on access to CRPD 2025). Healthcare providers access the CRPD

through their local health information systems or via the zVEM+ online portal. The CRPD data are made available to patients via the zVEM portal.

Under applicable legislation, all healthcare providers are required to submit data to the CRPD. Records are entered on a rolling basis. Approximately 200,000 new health records are entered into the CRPD each day, amount to approximately 5 million per month. From its launch until 31 December 2025, more than 330 million documents were available in the CRPD (Figure 5).

In 2023, the CRPD underwent a technological overhaul enabling the further development of advanced digital services. Significant functional expansions are expected through the "National EHR and unified data model" project, implemented under the Recovery and Resilience Facility (RRF).

3.7 eHealth Users Database (EueZ – security scheme)

The eHealth Users Database (EueZ) is the central authentication and authorisation service for the eHealth system components. Currently, eHealth system components are connected to the EueZ, and the components being developed within the RRP will soon also be integrated.

Almost 575,000 zVEM portal users have a profile in EueZ. The total number of EueZ users exceeds 590,000, including more than 2900 registered active healthcare providers.

Users' profiles contain the permissions and roles assigned to them within the organisations where they are employed. Most roles and permissions are assigned automatically based on rules derived from data on education and employment data recorded in the Register of Healthcare Providers and Healthcare Professionals (RIZDDZ), or from permissions defined on the HIIS professional card. Roles can also be assigned manually.

3.8 Teleradiology and TeleStroke

The two solutions are interconnected, as TeleStroke uses Teleradiology to transmit radiological images.

Teleradiology is used to transmit radiological images between HCPs. The system has been in place since the start of the eHealth programme and has evolved rapidly in terms of the number of transmissions and connected organisations. Most HCPs with a digital radiology image archive (PACS) connect it to the system. If not, they use Tele PACS to view and transmit images via Tele Portal. The number of monthly transmissions is currently approaching 16,000 (Figure 6). More than 70 organisations are already connected: two sevenths are hospitals, two sevenths health centres and three sevenths private HCPs.

The TeleStroke solution enables remote stroke management from the Division of Neurology in Ljubljana when a patient is treated in one of Slovenia's general hospitals without an on-call vascular neurologist. More than 1200 such cases are recorded each year. The patient is assessed using a web application; a standard questionnaire is completed, CT and/or MRI scans of the head are available for transmissions and viewing, and a report is prepared at the end of the procedure. Audiovisual patient management is also possible by means of diagnostic trolleys available in hospitals.

3.9 The empirical assessments of the adoption of eHealth services and qualitative data on user satisfaction

Kodermac and Peršolja (Kodermac & Peršolja, 2024) in their research about patient satisfaction with the eHealth system in Slovenia found that respondents were generally satisfied with the eHealth services (M = 4,18). Their satisfaction decreases only slightly with age. 62,1 % of respondents were satisfied or very satisfied with ePrescription, 62,4 % with eAppointment and 48 % with zVEM Patient Portal . Data from the Statistical Office of the Republic of Slovenia for the first quarter of 2025 show that 92% of individuals aged 16 to 74 use the internet, 27% of whom use it to access their own health records, while only 62% of the elderly (aged 65 to 74) use the internet several times a day. Among the 6% of those who have not yet used the internet, the elderly predominate (Statistical Office of the Republic of Slovenia, 2025).

4 Discussion

During the ten years of management by the NIJZ, the eHealth systems have made significant functional, organisational and user-related progress. The analysis shows that the system has gradually evolved from a project into a stable national infrastructure supporting data exchange, public health activities and patient involvement managing their own health. A notable acceleration in development was seen during the COVID-19 pandemic, when digital systems proved to be a key mechanism for ensuring continuity of care.

The greatest progress has been observed with the zVEM and zVEM+ portals. The development of new functionalities and the integration of HIIS content have enabled the zVEM patient portal to become a single-entry point for eHealth services and to significantly improve the transparency of health data. During the pandemic, it enabled users to view digital COVID certificates, test results and vaccination data. The surge in registered users in 2021 confirms that the system was capable of responding effectively to emergencies. At the same time, the zVEM+ portal enabled the real-time entry of vaccination and testing data into the CRPD, ensuring up-to-date data and supporting epidemiological surveillance. With the introduction of zVEM authorisations, the Slovenian eHealth system now fully meets the requirements of the European Health Data Space Regulation (EHDS) for the establishment of proxy services. In the future, zVEM+ will also serve as the main access environment for the National EHR and unified data model project, further strengthening its strategic role.

The ePrescription solution is also an important component of the digital transformation. During the pandemic, the use of remote electronic prescribing significantly improved patients' timely access to the medicines they needed. During the pandemic, remote prescribing significantly improved access to treatment, particularly for patients with chronic conditions and other vulnerable groups. The stable share of more than 97% of prescriptions issued electronically confirms a high level of acceptance among healthcare professionals and patients, as well as the system's organisational maturity. The successful implementation of the ePrescription system has also raised expectations regarding cross-border data exchange, highlighting the need for legal harmonisation, technical interoperability and standardisation of data models. Pilot data exchange with Croatia already

indicates the potential for Slovenia's broader integration into the European Health Data Space.

CRPD serves as the central data backbone of the eHealth system. The exponential growth in the number of documents and structured records confirm its key role in providing a comprehensive view of patients' medical records. Centralised access to data significantly contributes to the safety and quality of care, as it enables more informed clinical decision-making. At the same time, the growing volume of data raises questions about long-term architecture, infrastructure capacity, semantic interoperability, and secondary use of data for research and public health purposes.

The TeleStroke and Teleradiology systems demonstrate the important role of digital technologies in ensuring equitable access to specialist care across the country. TeleStroke directly affects stroke treatment outcomes, where time is a critical factor, while Teleradiology enables more efficient use of human and technical resources. Both systems demonstrate that digitalisation is not merely an administrative tool, but can directly influence clinical outcomes.

The eAppointment system has significantly improved the transparency of waiting times and the standardisation of waiting lists. The centralised display of first available appointments and the high share of electronically issued referrals indicate that the system is firmly established. However, ensuring the quality and timeliness of data remains a challenge, as the accuracy of reporting depends on healthcare providers. Changes in legislation and in the methodology for calculating patients waiting beyond permissible time limits required adjustments to the system, highlighting the need for a flexible architecture that enables the rapid implementation of regulatory changes.

A high level of personal data protection and security is essential for maintaining users' trust, particularly as the volume of digital communication and the exchange of sensitive health data continues to grow. Accordingly, cybersecurity and data protection are among the core principles of Slovenian eHealth. Several layers of security are ensured. All healthcare providers are connected to the secure communication network zNET. End users can only access the services by means of electronic identification. The eHealth Users Database provides centralised management of authentication and authorisation and ensures that the credentials of

health professionals are aligned with their actual employment and qualifications. Moreover, filtering of accessible content is implemented within eHealth applications. For example, access to CRPD is only granted to the patient's selected personal physician and the members of their care team. Moreover, the available content is filtered depending on the medical speciality (practice setting) so that the access to the most sensitive records (e.g. mental health records) is rigorously restricted (Tepej Jočić, 2021). Balancing data protection and accessibility is one of the greatest challenges of Slovenian eHealth. On the one hand, data availability is critical for patient safety and quality of care. On the other hand, nationwide access to EHR is often questioned in terms of privacy. The Slovenian eHealth system has embedded privacy by design. However, the governance of such a complex system is challenging and resource-consuming, and a vast proportion of issues and complaints are related to limited accessibility of personal data.

Despite these notable achievements, important challenges still remain. These include the need for additional investments in infrastructure, interoperability with European initiatives (e.g. the European Health Data Space), strengthening the digital competences of healthcare professionals and patients, and developing analytical tools for the secondary use of data. Although users' digital literacy has improved, it remains uneven, which may lead to inequalities in access to digital services.

In conclusion, the eHealth system in Slovenia has progressed from the implementation phase to a phase of system consolidation and further development. The COVID-19 pandemic acted as a catalyst for digital transformation, while future developments will largely depend on the successful implementation of projects under the Recovery and Resilience Plan and on the adoption of the Healthcare Digitalisation Act.

5 Conclusion

During the period of management by the NIJZ, the eHealth system in Slovenia progressed from the development and implementation phase to a phase of system stabilisation, expanded functionality and high user acceptance. Over the past ten years, it has become the core digital infrastructure of the Slovenian healthcare system, supporting data sharing, administrative processes, public health monitoring and, increasingly, active patient participation.

The analysis showed that one of the key turning points was the COVID-19 pandemic, which accelerated the use of already established systems (ePrescription, CRPD, eAppointment), and significantly increased the visibility and uptake of the zVEM portal and the zVEM+ system. Digital services have enabled continuity of care, secure data exchange, and rapid adaptation to emergency situations. This confirmed the strategic soundness of establishing nationally centralised systems based on a unified architecture and security framework.

A major achievement of the Slovenian eHealth system is the near-complete digitalisation of key processes (prescription issuing, referral issuing, electronic appointment scheduling), as well as the high level of population coverage. Today, the Central Registry of Patient Data is one of the largest national health data resources, opening significant opportunities for the development of analytics, research and data-driven decision-making in healthcare. At the same time, such a volume of data requires continued investments in security, interoperability, standardisation and the long-term sustainability of the infrastructure.

Despite the progress made, systemic challenges remain. These include ensuring adequate human and financial resources for further development, improving users' digital competences, reducing the digital divide among different population groups, and ensuring the quality and timeliness of data transmitted to central systems by healthcare providers. Particular attention will also need to be paid to alignment with European initiatives on cross-border health data exchange and the European Health Data Space.

Future development will be significantly shaped by the outcomes of projects under the Recovery and Resilience Plan and by the implementation of the Healthcare Digitalisation Act. Further development of systems such as National EHR, expanded electronic communication between patients and healthcare teams, and the wider use of electronic work orders will further strengthen the digital transformation of the Slovenian healthcare system.

It can be concluded that the eHealth system is indispensable for the functioning if the Slovenian healthcare system can no longer be imagined without the eHealth system. It is no longer just a supporting information tool, but a strategic

infrastructure that contributes to greater accessibility, safety, efficiency and quality of care, while strengthening the role of patients within the healthcare system.

References

Digitočke za portal zVem. (6 February 2026) NIJZ. https://nijz.si/obmocna-enota-koper/digitocke-za-portal-zvem

Drev, M., Tepej Jočić, L., & Kosednar, H. (2025). Nadzor nad obdelavo zdravstvenih podatkov = Oversight of health data processing. 579–584. https://nijz.si/wp-content/uploads/2025/09/Zbornik_Javno-zdravje-Most_2025.pdf

e-Zdravje2010 – Strategija informatizacije slovenskega zdravstvenega sistema 2005–2010 (2005). https://interoperable-europe.ec.europa.eu/sites/default/files/document/2014-12/e-Zdravje2010%20-%20Strategija%20informatizacije%20slovenskega%20zdravstvenega%20sistema%202005-2010.pdf

European Commission, Directorate-General for Communications Networks, C. a. T., & Invent, C. (2024). 2024 digital decade ehealth indicator study – Annex – Country factsheets. Publications Office of the European Union. https://doi.org/10.2759/276133

Janet, J., & Štupica, E. (2025). Pregled razvoja rešitev zVEM in zVEMplus ter predstavitev novosti. Overview of the development of zVEM and zVEMplus solutions and presentation of new features. 609–616.

Kljajić Borštnar, M. (2022). Znanstveno proučevanje družboslovnih pojavov.

Kodermac, N., & Peršolja, M. (2024). Zadovoljstvo pacientov z informacijskim sistemom eZdravje v splošni bolnišnici: presečna raziskava. Obzornik zdravstvene nege, In Press. https://doi.org/10.14528/snr.2024.58.4.3266

Kosednar, H. (2024). Dobrodošle nove možnosti eNaročanja. Digitalna preobrazba za dostopno, varno in kakovostno zdravstveno oskrbo: kongres MI'2024.

Kosednar, H. (2025). Vpliv projektov za digitalizacijo zdravstva iz Načrta za okrevanje in odpornost na javno zdravje v Sloveniji = The impact of healthcare digitalization projects from the recovery and resilience plan on public health in Slovenia. 617–624. https://nijz.si/wp-content/uploads/2025/09/Zbornik_Javno-zdravje-Most_2025.pdf

Plevnjak, N., Stanimirović, D. (2024). eNaročanje : pregled dosedanjih ugotovitev o delovanju informacijske rešitve in nadaljnji razvojni koraki. NOVIS: Novice, obvestila, vesti, informacije združenja: Glasilo Združenja zdravstvenih zavodov Slovenije, ISSN 1580-4917, 2024, letn. 51, št. 5/6, str. 38-41. https://www.zdrzz.si/Novis_1/

Portal zVEM. (2025). In.

Pravilnik o pooblastilih za obdelavo podatkov v Centralnem registru podatkov o pacientih. Uradni list RS, št. 51/16, 95/21 in 35/24; https://pisrs.si/pregledPredpisa?id=PRAV15410

Rant, m. Ž., & Rant, Ž. (2018). Projekt e-zdravje v Sloveniji - vsebinski in zgodovinski pregled. In Š. Adamič (Ed.), 30 let Slovenskega društva za medicinsko informatiko (pp. Str. 180-183). Slovensko društvo za medicinsko informatiko.

Rant, Ž., Čančarević, D., & Kosednar, H. (2025). Deset let eRecepta v Sloveniji = Ten years of ePrescription in Slovenia Javno zdravje – most do boljšega zdravja in odpornosti v skupnosti: NKJZ, https://nijz.si/wp-content/uploads/2025/09/Zbornik_Javno-zdravje-Most_2025.pdf

Rant, Ž., Kosednar, H., & Stanimirović, D. (2025). Novosti v nacionalni rešitvi eNaročanje. Human Being, Artificial Intelligence and Organization, Conference Proceedings, 785-796. https://doi.org/10.18690/UM.FOV.2.2025.60

Rant, Ž., & Stanimirović, D. (2019). Analysis of eHealth Solutions in Slovenia: A Usage Perspective. In.

Rant, Ž., & Stanimirović, D. (2020). Analiza uporabe rešitev e zdravja v Sloveniji : izzivi in prihodnje usmeritve. In 39th International Conference on Organizational Science Development (pp. 645-661). Maribor: University Press.

Rant, Ž., & Stanimirović, D. (2021). Presečna analiza uspešnosti in učinkovitosti rešitev eZdravja po petletnem obdobju : uspešnost in učinkovitost eZdravja. NOVIS : novice, obvestila, vesti, informacije združenja : glasilo Združenja zdravstvenih zavodov Slovenije. ISSN 1580-4917. - Letn. 48, Št. 1/2 (2021), 48(1/2), 23-25.

Rant, Ž., & Stanimirović, D. (2024a). Analiza razvojnih in uporabniških vidikov eRecepta v Sloveniji. Green and Digital Transition - Challenge or Opportunity : conference proceedings, 877-886. https://doi.org/10.18690/um.fov.3.2024.63

Rant, Ž., & Stanimirović, D. (2024b). Rešitev eRecept; Elektronski vir : pogled skozi prizmo koristi za pacienta in dodane vrednosti za zdravstveni sistem; The ePrescription solution : a view through the prism of benefits for the patient and added value for the healthcare system :: COBISS Plus. zbornik recenziranih prispevkov in povzetkov : 2. nacionalna konferenca javnega zdravja.

Rant, Ž., Stanimirović, D., & Janet, J. (2021). Novosti pri portalu za paciente zVEM in Centralnem registru podatkov o pacientu. In Spodbujeni s pandemijo v digitalno preobrazbo zdravstva : strokovno srečanje MI'2021 : v Termah Zreče in na daljavo, 4. in 5. november 2021 : zbornik povzetkov. - (pp. 23-23). Ljubljana: SDMI.

Rant, Ž., Stanimirović, D., & Janet, J. (2022a). FUNCTIONALITIES AND USE OF THE ZVEM PATIENT PORTAL AND THE CENTRAL REGISTRY OF PATIENT DATA. 35th Bled eConference: Digital Restructuring and Human (Re)action, BLED 2022 - Proceedings, 65-79. https://doi.org/10.18690/UM.FOV.4.2022.4

Rant, Ž., Stanimirović, D., & Janet, J. (2022b). Razvoj portala zVEM in Centralnega registra podatkov o pacientu. 41th International Conference on Organizational Science Development : society's challenges for organizational opportunities, 873-884. https://doi.org/10.18690/UM.FOV.3.2022.63

Rant, Ž., Stanimirović, D., Matetić, V., Indihar, S., Zidarn, J., Beštek, M.,…Živa Rant, D. S. V. M. S. I. J. Z. M. B. L. T.-J. A. Ž. (2017). eZdravje danes. Uporabna informatika, 25(3), str. 169-179.

Rant, Ž., Stanimirović, D., Tepej Jočić, L., Žlender, A., Gaspari, I., Božič, D.,…Zidarn, J. (2018). Rešitve e-Zdravja. In Š. Adamič (Ed.), 30 let Slovenskega društva za medicinsko informatiko (pp. Str. 184-190). Slovensko društvo za medicinsko informatiko.

Rant, Ž., Stanimirović, D., & Žlender, A. (2019). Nacionalni Portal zVEM v okviru eZdravja = National portal zVEM within eHealth. In Ekosistem organizacij v dobi digitalizacije (pp. Str. 873-884). Maribor: Univerzitetna založba Univerze.

Stanimirović, D., Drev, M., & Rant, Ž. (2022a). Digital Transformation as One of the Instruments for Overcoming the Public Health Crisis: The Role and Use of Ehealth Solutions during the Covid-19 Pandemic in Slovenia. Medicine, Law & Society, 15(1), 169-192. https://doi.org/10.18690/MLS.15.1.169-192.2022

Stanimirović, D., Drev, M., & Rant, Ž. (2022b). Ekspanzija digitalnih rešitev eZdravja v času pandemije covida-19 v Sloveniji. Isis, 31(6), 34-37.

Statistical Office of the Republic of Slovenia. (2025). Usage of internet in households and by individuals, 2025. Stat.si. https://www.stat.si/StatWeb/en/News/Index/13908

Tepej Jočić, L. (2021). *Impact of data protection regulation on Slovenian eHealth. Journal of Global Health, 11*, 03063. https://doi.org/10.7189/jogh.11.03063

Tepej Jočić, L. (2024). Vloga demografskih podatkov pri informatizaciji javnozdravstvene dejavnosti. *ISIS: Glasilo Zdravniške zbornice Slovenije, 2024*(3), 30–34.

Yin, R. K. (2018). Case study research and applications : design and methods (6th ed.). Sage.

Zakon o digitalizaciji zdravstva (ZDigZ). (2025). Uradni list

Zakon o zbirkah podatkov s področja zdravstvenega varstva (ZZPPZ). (2015). In. https://pisrs.si/pregledPredpisa?id=ZAKO1419

A Recent Review on Artificial Intelligence in Ports

Vítor Ribeiro, João Barata,
Paulo Rupino da Cunha, Jacinto Estima
University of Coimbra, CISUC/LASI, DEI, Coimbra, Portugal
vhribeiro@dei.uc.pt, barata@dei.uc.pt, rupino@dei.uc.pt, estima@dei.uc.pt

Ports are investing significantly in artificial intelligence (AI). However, research remains nascent and fragmented in this area, making it complex to develop integrated plans that leverage on data. This paper presents a systematic literature review of 28 recent publications that examine the value of AI in port's digital and green transformation. The work was conducted within the scope of a multiyear investment project on the transformation of ports and logistics. The taxonomy of publications outlines the main pillars of AI application in ports: operation efficiency, support and maintenance, scheduling and forecasting capacity, environmental policies, and surveillance and human safety. Our study reveals an increase in the scale of AI adoption, including the orchestration of seaport-dry operations, multimodal transportation integration, and analysis of environmental impacts. Our contribution includes near-term use cases and guidelines to address the challenges of AI adoption in mission-critical, large-scale logistics infrastructures while promoting sustainablility.

DOI
https://doi.org/10.18690/um.fov.4.2026.52

ISBN
978-961-299-152-4

Keywords:
artificial intelligence,
dry port,
machine learning,
seaport,
smart port,
sustainability

1 Introduction

Ports are the backbone of advanced logistics networks that support over 80% of the world's trade (Evmides et al., 2024; H. Xie & Ambrosino, 2025). On the one hand, seaports play a strategic role in establishing maritime transportation routes, which are increasingly constrained by demand for energy, valuable resources, and products that support economic development. On the other hand, dry ports, located inland and served primarily by trucks and trains, offer an effective means of reducing seaport congestion, improving container handling and storage efficiency, and contributing to local economies (Rekabi et al., 2025). Despite differences in location, size, or cargo type, all share common priorities: digital transformation, green practices, and sustainable management (Aerts & Mathys, 2024).

The term "smart port" has recently been used to describe the disruptive transformations enabled by digital technologies in these critical infrastructures for international trade (Li et al., 2023). Similar to other industry sectors, artificial intelligence (AI) is driving interest in digitalization of ports (Liu & Yuen, 2025). Grounded on the work of Russell & Norvig (1995), we define AI as the set of computational techniques (e.g., computer vision, machine learning, natural language processing) that capacitate systems to perform actions that typically require human intelligence (e.g., learning, decision-making, forecasting). The AI transition raises data governance challenges, including data quality and ownership management of shared data assets (Abraham et al., 2019). Moreover, complex platform ecosystems are necessary to coordinate the several stakeholders, enable data exchange, and foster integration (Hein et al., 2020). Lastly, AI-based solutions are delegating decision-making to automatic algorithms (Guo et al., 2025), requiring adjustments in port's organizational structures.

Relevant ports in Hamburg, Singapore, Valencia, or Rotterdam are leading the way in terms of AI adoption (Osundiran & Makgopa, 2025). Yet, "*there remains a gap, especially concerning the recent AI developments and their integration into port operations post-2022*" (Liu & Yuen, 2025). The work of Liu & Yuen (2025) is one of the most recent and comprehensive reviews of AI application in ports, detailing findings up to 2023. Our search revealed a major increase in publications in the past two years, motivating the team to extend prior research. Our Research Question (RQ) is: "*What are the emerging use cases of AI-based applications (post-2024) in port operations' Digital Transformation*

(DT), and how do they contribute to operational and sustainable improvements?" We frame our research through the DT process described by Vial (2019), aiming to understand how AI-based applications reshaped port's operations, contributing to automate operations, foster inter-organizational collaboration, and promote sustainable practices.

This research falls within the scope of the NEXUS – Innovation Agenda for Digital and Green Transition, a multiyear research and development project aimed at performing the digital and green transformation of ports and logistics in Portugal (NEXUS, 2026). Led by the Port of Sines and comprising 35 partners (e.g., universities, logistical operators, technology producers), NEXUS is developing an ecosystem of 28 innovative products, many of which are enabled by artificial intelligence, including federated applications and services and an open data collaboration platform. Understanding how AI is transforming ports is the main research objective of this paper, conducted in the form of a systematic literature review (Webster & Watson, 2002) of papers published since 2024, and a starting point for NEXUS intervention. Advances in AI over the past two years have created favorable conditions for summarizing knowledge on AI adoption in this essential sector of the economy, which is struggling to optimize operations and reduce environmental impacts (Rekabi et al., 2025).

The rest of this paper is structured as follows. Section 2 explains the research approach. The results of the bibliometric analysis and the systematic literature review are included in Section 3, followed by the taxonomy of recent use cases of AI adoption in ports (Section 4). The paper closes in Section 5, stating the main conclusions, limitations, and future work opportunities.

2 Research Approach

Our work began with a systematic literature review (SLR) using a concept-centric analysis (Webster & Watson, 2002), which was updated during the NEXUS project execution to ensure comprehensive coverage of the most relevant studies.

The current version of our sample was obtained via Scopus using the keyword combination ("artificial intelligence" OR "machine learning") AND ("smart port" OR "dry port" OR "seaport" OR "port 4.0"), yielding 210 results in the title,

abstract, and keywords fields. The initial sample of 210 works was used to construct bibliometric networks to identify relevant clusters and topics using the VoS Viewer Tool (van Eck & Waltman, 2010). We removed two duplicate publications and two non-English articles, reducing the initial sample to 208 papers. We included peer-reviewed journal publications, focusing on AI adoption in ports, published since 2024 (to provide the most recent discussion of the AI literature, since it is advancing extremely fast, with almost half of the papers being published from 2024 onwards and recent reviews reporting research until 2023 (Liu & Yuen, 2025)), resulting in a sample of 43 papers. Journals or publishers that did not meet the basic threshold for national rankings were also excluded from the sample. We screened the publications' title, abstract, and keywords, resulting in a final SLR sample of 28 articles selected for full-text reading and extraction of use cases, methodologies, and results of AI adoption in ports to perform the concept-centric analysis (Webster & Watson, 2002). Figure 1 introduces the SLR schema.

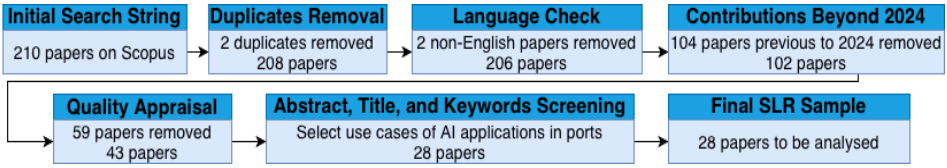

Figure 1: SLR Procedures on AI Applications
Source: Author's Own Work

We applied a concept-centric coding scheme, classifying studies by application domain and reported outcomes. The coding scheme was iteratively refined to align with the RQ, and ambiguous cases were reexamined to improve reliability. We iteratively reported SLR's insights during the NEXUS's project meetings, contributing to plan upcoming AI investments, anticipate implementation challenges, and forecast potential outcomes for each choice. As an example, NEXUS's deliverables included an automatic container damage identification tool and system to simulate and optimize berth allocation operations.

3 Results

This section details the findings for the AI-based solutions in port operations, including (1) a bibliometric analysis and (2) a SLR using a concept-centric analysis.

3.1 Bibliometric Analysis of Scopus Clusters

Figure 2 details the results of the bibliometric analysis for the 210 papers sample obtained in Scopus. It was a starting point to identify research priorities in the field.

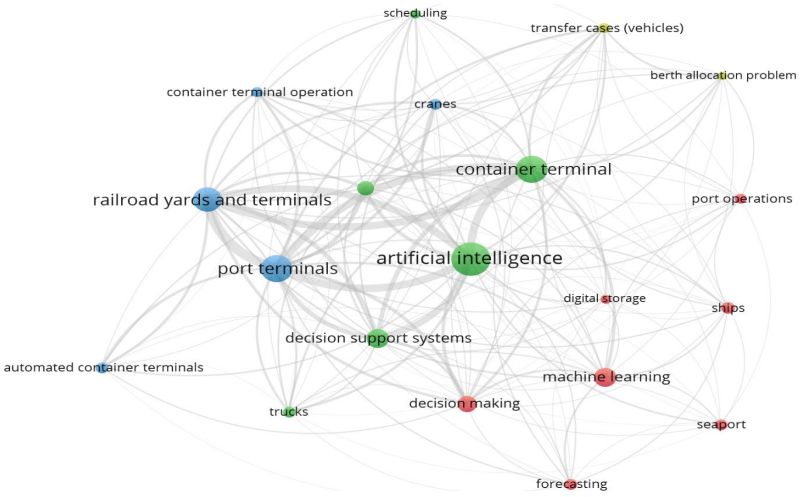

Figure 2: Bibliometric Analysis in Scopus (keywords co-occurrence with a minimum of 15 occurrences)

Source: Author's Own Work Using VoS Viewer

We identify four clusters in the bibliometric network in Figure 2, deploying the default VOSviewer resolution configuration to balance interpretability and granularity. The blue cluster (on the left part of Figure 2) represents port terminals and railroad yards, including automatic container terminals and corresponding operations (e.g., using cranes). The green cluster (at the center of Figure 2) highlights the role of AI in container terminal operations, addressing its deployment in decision support systems (DSSs) and scheduling. The red cluster (at the right part of Figure 2) details the use of machine learning (ML) sustained on digital storage to enable decision making and forecasting in port and ship-related operations. Lastly, the yellow cluster (at the right-top of Figure 2) focuses on the problem of berth allocation that often involves challenges related to transfer cases. The clusters reveal the multidimensional attributes of our sample, including both seaport and dry port settings, multimodal transport options, and the central role of cargo (e.g., minimizing movements) in AI models' concerns.

3.2 Systematic Literature Review Results

We use a concept-centric analysis to report the findings in our SLR (Webster & Watson, 2002), derived from our findings in the bibliometric clusters analysis, including AI applications for port operations. The studies included in our sample addressed relevant problems in port business. For example, (1) infrastructure management, (2) planning and scheduling, (3) traffic management, (4) environmental sustainability, (5) decision-making and forecasting in port operations, (6) asset and resource management, and (7) supply chain management. All these applications contributed to increase revenue, enhance resource and operational optimization, and promote competitiveness in the context of port operations.

In terms of infrastructure and management, we observe that machine learning (ML) algorithms operating on real-time data are increasingly used in port management, namely to overcome the limitations of rudimentary analytics. ML is used to study and predict risks of seaport disruptions induced by natural hazards (Khan et al., 2025) as well as from hazardous cargo operations (Khan et al., 2024). The recent study presented by Aslam et al. (2025) uses ML-based technology to predict faults in container handling equipment (e.g., inverters).

AI technology is applied to automate, streamline, and optimize planning and scheduling activities in the context of port operations. Berth allocation is a complex problem in port management operations, and AI is targeted as a potential enabler for more efficient and quicker solutions. In Lv et al. (2024), a deep reinforcement learning (DRL)-based model optimizes berth allocation operations. For the same problem, AI and mathematical models are applied to address complex berth allocation operations in inland water transport terminals, considering the challenges of the multiple types of handling devices, cargo, and barges (Pasha et al., 2025).

In Y. Xie et al. (2025), the researchers propose a predictive ML-based model for container out-terminal assignment that enables significant cost savings when compared with existing solutions. Also Yan et al. (2025) explore DRL algorithms to optimize container relocation operations. Another challenge that AI addresses is the quay scheduling. For example, You et al. (2025) exploit model-free DRL technology to address challenges in quay crane scheduling in terminal operations. Guo et al. (2025) develop and evaluate a novel algorithm (combining AI, flow modeling, and

optimization theories) to automatically define and optimize quay and berth machinery scheduling in inland river ports. Some ports require the combination and planning of multimodal transport networks. He et al. (2025) developed and tested a novel path optimization technique to optimize multimodal transport operations.

AI is revolutionizing traffic management within the seaport and surrounding areas. Pham et al. (2024) address the challenges posed by voluminous vessel flows in seaport zones, proposing a novel integrated maritime management system that predicts traffic conditions and identifies a secure path for incoming ships, while Duan et al. (2025) propose a novel approach to address vessel traffic flow and berth allocation in high-congestion ports with varying constrained channel widths. Ports can forecast vessel arrivals and estimate their paths, improving the existing solution for predicting ship arrival times (Saber et al., 2025). Zygouras et al. (2024) develop and evaluate a method that combines ML and advanced data techniques to estimate a vessel's complete route based on its characteristics and historical information. Trucks may also cause congestion in ports. In Silva et al. (2024), the authors explore ML techniques to propose a conceptual truck appointment management system that addresses potentially impacting events by promoting the necessary flux adjustments.

Seaports face increasingly stringent regulations and ambitious emission reduction targets. In Gulyás (2025), the authors discuss how AI/ML techniques can be used to reduce the effort required to create air emission inventories (EIs) while enhancing the accuracy of these valuable monitoring instruments. Energy consumption forecasting is fundamental to optimize energy production and minimize greenhouse gas emissions. Examples of studies in this line include Song et al. (2026), who discuss the use of machine learning to forecast energy demand in modern seaports, thereby improving efficiency, reducing operating costs, and advancing carbon neutrality. Also Conte et al. (2025) use artificial intelligence techniques to forecast and optimize energy use from ships at berth, utilizing a combination of sources (e.g., shore-side electricity, multi-energy storage, renewable generation), thus mitigating their greenhouse gas emissions. Zhang et al. (2025) contribute with an energy consumption and benefit assessment model for ship operations in container terminals that contributes to estimating efficiency. AI applications can estimate and predict air quality in ports, engaging managers to develop necessary actions. AI is used to predict air quality in areas surrounding ports. In this context, Bakary et al. (2024) propose an ML-based model to estimate greenhouse gas emissions in loading

and unloading operations located in quayside at seaports. Ultimately, ML, statistics, and geospatial analysis techniques are combined and applied to identify pollution hotspots in areas surrounding seaports (Cabrera-Ormaza et al., 2025).

AI can automate and augment decision-making in port operations, with a special emphasis on logistical planning. Siri et al. (2024) propose a novel Decision Support System (DSS)-based methodology to optimize train load planning in port container terminals. These new tools reveal the potential to change the role of humans in port operations, suggesting that training in more advanced DSS is necessary for operators and managers, which is not yet the norm. As AI becomes more relevant to support decision-making, human agents must be prepared to identify how to use it and the risks that may arise (e.g., poor data leading to incorrect recommendations).

AI can also be applied to the port's physical assets and resources, enabling infrastructure and equipment intelligence. Solutions target defects and faults in containers and port-related equipment. Convolutional neural networks, a subtype of ML algorithms, are used to develop a framework for automatically identifying and classifying visual defects in containers (Kuo et al., 2025). The work of Liao et al. (2024) proposes and tests an innovative approach to identify and diagnose faults in trolley mechanisms in ship-to-shore cranes, using knowledge-based reverse deep belief networks. Another possibility is to deploy AI-based systems to optimize the operationalization of port resources (e.g., cranes). In Yu et al. (2025), the authors apply computer vision techniques to estimate and monitor the swing angle of the portal crane's bucket grab, enhancing safety and security in port operations. Satellite imagery can be analyzed by AI-based algorithms to automatically estimate the number of containers in port storage infrastructures, significantly optimizing this operation while requiring fewer resources (Zhou & Wang, 2024).

AI streamlines inter-organizational collaboration with logistical chains. The combination of ML and Blockchain technology enhances security and information integration, improving traceability, logistics, and inter-organizational processes in port operations (Durán et al., 2024). For example, Rekabi et al. (2025) explore ML and IoT to create a novel dry port-seaport logistical network that assesses the impact of sustainable practices (e.g., CO2 emissions), enables real-time activity monitoring and prediction, and estimates demand forecasting and waste management.

Most of the analyzed AI-based solutions relied on near real-time, interoperable, and high-quality data collected from multiple devices and stakeholders (e.g., berth planning). The analysis of the results also highlights that most advanced AI implementations are applied in optimization operations (e.g., traffic management), while exploratory solutions target strategic planning and automation (e.g., risk assessment).

4 A Taxonomy of Use Cases for the NEXUS Agenda

Table 1 summarizes the findings for AI-enabled smart port implementation and their contribution to the United Nations' Sustainable Development Goals (SDGs) (United Nations, 2026), based on the insights of the SLR.

Table 1 extends the important work presented by Liu & Yuen (2025) to understand the most recent landscape of AI adoption in ports. Furthermore, it offers a practice-oriented taxonomy to deploy AI in optimizing port operations and promoting sustainability practices, focusing not only on the port itself, but also on the entire logistical chain of stakeholders. Lastly, it includes a reflection on the contribution of AI-based solutions to achieving SDGs. Our results have been iteratively updated over the past two years to ensure the adoption of state-of-the-art AI.

Our research contributes to the DT literature, demonstrating how the AI adoption is reshaping organizational capabilities by enabling automatic processes and data-driven decision making. In this context, reshaped inter-organizational data exchange networks become a priority to integrate decentralized information systems and streaming information dissemination. Consequently, digital governance structures are necessary to ensure accountability, coordination, and control in ports' multi-stakeholder setups.

The increased AI adoption requires a socio-technical transformation of organizational structures (Murire, 2024), in which technological solutions are blended with human roles and work practices to perform tasks and make decisions. Furthermore, the implementation of AI-based solutions supporting automatic decisions raises issues associated with accountability and risk (Cheong, 2024). Subsequently, ports must reorganize their organizational structures to include new roles, decision-making processes, and accountability for automated decisions.

Table 1: AI-Enabled Smart Port: A Taxonomy of Recent Use Cases

Scope	Use Case	SDGs	References
Port: terminal infrastructure and management	Energy optimization and forecasting Intelligent infrastructure maintenance [Data] Governance, legal, and ethical assessment Risk/hazards assessment and prediction Emissions inventory accuracy	SDG-7 SDG-9 SDG-13 SDG-14	(Aslam et al., 2025; Conte et al., 2025; Gulyás, 2025; Khan et al., 2024, 2025; Song et al., 2026; Zhang et al., 2025)
Port: cargo (e.g., containers) handling	Container relocation and stacking Container counting Berth planning Vessel berthing optimization Estimated time of arrival Throughput under supply chain disruptions Parking facility improvement Load/unload planning/sequencing Container location optimization	SDG-9 SDG-13	(Lv et al., 2024; Pasha et al., 2025; Saber et al., 2025; Siri et al., 2024; H. Xie & Ambrosino, 2025; Yan et al., 2025; Zhou & Wang, 2024)
Port: gate	Container identification, damage, and defect categorization Predict and/or reduce congestion	SDG-9 SDG-13 SDG-14	(Kuo et al., 2025; Y. Xie et al., 2025)
Port: quay cranes	Motion detection Machinery scheduling Fault diagnosis	SDG-9	(Guo et al., 2025; Liao et al., 2024; You et al., 2025; Yu et al., 2025)
Third party: Environmental quality	Environmental monitoring Pollutants prediction	SDG-9 SDG-13 SDG-14	(Bakary et al., 2024; Cabrera-Ormaza et al., 2025)
Third party: port network	Seaport – dry port cargo movements optimization Port–city integration Port selection	SDG-9 SDG-13 SDG-14	(Durán et al., 2024; He et al., 2025; Rekabi et al., 2025)
Third party: Transportation operators	Optimize energy in transport operations Truck routing optimization Autonomous navigation/driving Shipboard microgrids	SDG-9 SDG-13 SDG-14	(Silva et al., 2024; Zygouras et al., 2024)
Third party: Authorities	Safe maritime/ship traffic control Satellite-based bathymetry in seaports Small non-metallic vessels detection Inspections recommendations	SDG-9 SDG-14 SDG-16	(Duan et al., 2025; Pham et al., 2024)

Source: Author's Own Work

The deployment of AI-based solutions requires massive volumes of input data incoming from multiple infrastructures across the port (e.g., real-time data collection to predict traffic in the port (Pham et al., 2024)) and the distinct stakeholders in the logistical chain (e.g., data from vessels arriving at the port (Saber et al., 2025), optimizing multimodal transport routes (He et al., 2025)). As a consequence, ports

must deploy data governance practices to manage data sharing operations, meet data quality thresholds, and secure sensitive data (Abraham et al., 2019).

The deployment of AI can foster the "smart"- "green" port nexus (Dey et al., 2025), exploiting digital technologies to promote SDGs (United Nations, 2026). Regarding SDG7 – "*Affordable and Clean Energy*", AI technologies support ports in optimizing and predicting energy consumption in infrastructures and ships. Focusing on SDG9 – "*Industry, Innovation and Infrastructure*", AI is enabling the implementation of intelligent, safer, and resilient ports. For SDG 13 – "*Climate Action*", AI is enabling ports to make operations more efficient, promote environmental awareness, and reduce resource waste. Focusing on SDG 14 – "*Life Below Water*", AI-based solutions promote sustainable maritime practices that mitigate impact on seas and biodiversity.

Despite the inspirational works found in recent years advancing AI in ports, there are also relevant gaps in the literature that we confirmed in our field interventions in the NEXUS agenda. First, there is a lack of comprehensive evaluations of the business value of AI in ports. Most studies still address the operational level of decision-making, missing the analysis of the economic or societal benefits of AI (for example, quantifying waste reduction or the potential benefits to human safety during port operations). Second, while leading ports have already included AI in their investment agenda, it is not clear how the IT companies' ecosystem is incorporating AI into their offerings (e.g., terminal operating systems) and how AI use in daily operations is uncertain. Accurate recommendations require reliable data, which is not always available in port settings, and may also require data from third-party operators. Data integration between the port and other stakeholders is challenging. Third, there is a lack of studies focusing on agentic AI. Many operations may be delegated to intelligent agents (e.g., information requests to operators), but we found no studies that highlight agentic AI use cases or provide empirical examples. Future work is needed to shift AI's impact from near-term decisions (e.g., defect classification) to strategic value and autonomy in port operations. Given the diversity of ports and their networks, large-scale projects are necessary to obtain the data and accuracy levels critical to modern logistics networks. However, we gathered positive indications from existing studies and our contacts with port practitioners that the road to more effective AI use is already being developed.

5 Conclusion

This article presents the results of an SLR on AI adoption in port operations. We provide insights into field interventions within the NEXUS investment project. We systematize recent (2024-2026) use cases and suggestions for AI adoption in port operations, along with existing challenges, to further advance the adoption of technological solutions in the sector and promote sustainable practices.

Some limitations must be stated. First, our analysis was based on Scopus, limiting the validity of our findings. Future research can include other bibliographical databases (e.g., Web of Science). Second, the empirical insights for the AI-enabled smart port implementation focus on recent studies, aiming to capture the trends in this field and inform one of the largest European co-funded projects of port transformations. As a consequence, works published before 2024 were not included in the review. Upcoming research stages are aimed at extending the analytical synthesis to identify cross-cutting patterns and maturity levels of AI application in ports. Future work may address the gaps found, namely (1) evidence of IA impact in port strategies, (2) network use of AI requiring data lakes and trustworthy logistical operations data in a wider scale, and (3) prioritizing agentic AI to assist humans in port operations.

Acknowledgment

This work was supported by Portuguese Recovery and Resilience Plan (PRR) within the scope of the Agenda "NEXUS: Pacto de Inovação–Transição Verde e Digital para Transportes, Logística e Mobilidade" (Investment Project No. 53) under Grant C645112083-00000059. This work was also developed within the scope of the research undertaken by Research Unit–Centre for Informatics and Systems of the University of Coimbra–Grant UID/00326 - Foundation for Science and Technology (FCT), I.P.

References

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management, 49(January), 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Aerts, G., & Mathys, G. (2024). Discovering trends in the digitalization of shipping: An exploratory study into trends using natural language processing. Journal of Marine Science and Engineering, 12(4), 618.

Aslam, S., Navarro, A., Aristotelous, A., Garro Crevillen, E., Martnez-Romero, A., Mart\'\inez-Ceballos, Á., Cassera, A., Orphanides, K., Herodotou, H., & Michaelides, M. P. (2025).

Machine Learning-Based Predictive Maintenance at Smart Ports Using IoT Sensor Data. Sensors, 25(13), 3923.
Bakary, K., Mouhamadou, D., & Seydou, T. (2024). Contribution Based on Neurons Networks for the Prediction of Greenhouse Gas Emissions in a Handling Port. Asian Journal of Water, Environment and Pollution, 21(6), 261–269. https://doi.org/10.3233/AJW240094
Cabrera-Ormaza, M., Boente, C., Flor-Blanco, G., González-Garcia, N., Gallego, J. L. R., & Sierra, C. (2025). Seabed Pollution Near Industrialized Seaports: An Enhanced Multivariate Technique to Explore Trace Element Dynamics. Water, Air, and Soil Pollution, 236(11). https://doi.org/10.1007/s11270-025-08361-9
Cheong, B. C. (2024). Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. Frontiers in Human Dynamics, 6, 1421273.
Conte, F., Natrella, G., Sasso, D., Silvestro, F., D'Agostino, F., D'Antoni, F., & Merone, M. (2025). Optimizing green port operations with reinforcement learning and model predictive control. International Journal of Electrical Power & Energy Systems, 173, 111437.
Dey, S., Mishra, R., & Mukherjee, S. (2025). Supply Chain Resilience and Sustainable Digital Transformation with Next-Generation Connectivity in a Smart Port. Communications of the Association for Information Systems, 57(1), 51.
Duan, H., Miao, L., Jia, S., Zhang, C., & Lam, J. S. L. (2025). Integrated vessel traffic scheduling and berth allocation with restricted channel widths. European Journal of Operational Research.
Durán, C., Fernández-Campusano, C., Carrasco, R., & Carrillo, E. (2024). DMLBC: Dependable machine learning for seaports using blockchain technology. Journal of King Saud University - Computer and Information Sciences, 36(1), 101918. https://doi.org/https://doi.org/10.1016/j.jksuci.2024.101918
Evmides, N., Aslam, S., Ramez, T. T., Michaelides, M. P., & Herodotou, H. (2024). Enhancing prediction accuracy of vessel arrival times using machine learning. Journal of Marine Science and Engineering, 12(8), 1362.
Gulyás, F. Z. (2025). CAN AI ADDRESS THE SHORTCOMINGS OF THE CURRENT PORT EMISSIONS INVENTORY PRACTICES? International Maritime Transport and Logistic Journal, 14, 324–336.
Guo, H., Li, Y., Xie, S., & Wan, Y. (2025). Berth and Quay Machinery Scheduling Algorithm for Bulk Cargo Terminals in Small and Medium-Sized Inland River Ports. In Intelligent Transportation and Smart Cities (pp. 251–259). IOS Press.
He, J., Chen, W., Sun, J., & Zhu, L. (2025). Integrating graph neural networks and LSTM for path optimization in smart port multi-modal systems. Plos One, 20(12), e0336629.
Hein, A., Schreieck, M., Riasanow, T., Setzke, D. S., Wiesche, M., Böhm, M., & Krcmar, H. (2020). Digital platform ecosystems. Electronic Markets, 30(1), 87–98. https://doi.org/10.1007/s12525-019-00377-4
Khan, R. U., Yin, J., Ahani, E., Nawaz, R., & Yang, M. (2024). Seaport infrastructure risk assessment for hazardous cargo operations using Bayesian networks. Marine Pollution Bulletin, 208, 116966.
Khan, R. U., Yin, J., Asad, M., Alomayri, T., & Jameel, M. (2025). Predicting seaport disruptions from natural hazards using automated machine learning. Marine Pollution Bulletin, 221, 118477.
Kuo, S.-Y., Lin, P.-C., Huang, X.-R., Huang, G.-Z., & Chen, L.-B. (2025). Cad-Transformer: A CNN–Transformer Hybrid Framework for Automatic Appearance Defect Classification of Shipping Containers. IEEE Transactions on Instrumentation and Measurement, 74, 1–21. https://doi.org/10.1109/TIM.2025.3548214
Li, K. X., Li, M., Zhu, Y., Yuen, K. F., Tong, H., & Zhou, H. (2023). Smart port: A bibliometric review and future research directions. Transportation Research Part E: Logistics and Transportation Review, 174, 103098.
Liao, X., Ming, X., & Xia, M. (2024). KBRDBN: An Interpretable Deep Belief Network for the Fault Diagnosis of the Trolley Mechanism in Ship-to-Shore Cranes. IEEE Transactions on Instrumentation and Measurement, 73, 1–12. https://doi.org/10.1109/TIM.2023.3318717

Liu, X., & Yuen, K. F. (2025). A systematic review on artificial intelligence applications in seaports–a network analysis approach. Expert Systems with Applications, 289, 128309.

Lv, Y., Zou, M., Li, J., & Liu, J. (2024). Dynamic berth allocation under uncertainties based on deep reinforcement learning towards resilient ports. Ocean & Coastal Management, 252, 107113.

Murire, O. T. (2024). Artificial intelligence and its role in shaping organizational work practices and culture. Administrative Sciences, 14(12), 316.

NEXUS. (2026). NEXUS - Innovation Agenda for Digital and Green Transition. https://nexuslab.pt/

Osundiran, A. O., & Makgopa, T. (2025). An Examination of the Impact of Artificial Intelligence on Maritime Port Efficiency and Businesses. In Diversity, AI, and Sustainability for Financial Growth (pp. 255–288). IGI Global Scientific Publishing.

Pasha, A., Liu, J., & Rastogi, R. (2025). A mathematical model and a memetic algorithm for berth allocation at inland waterway terminals in developing countries. Engineering Applications of Artificial Intelligence, 152, 110843.

Pham, T.-A., Dang, X.-K., Koboević, Ž., Do, V.-D., & Pham, T.-D. A. (2024). Maritime data mining for marine safety based on deep learning: southern Vietnam case study. NAŠE MORE: Znanstveni Časopis Za More i Pomorstvo, 71(1), 21–29.

Rekabi, S., Sazvar, Z., Tavakkoli-Moghaddam, R., & Dolgui, A. (2025). Developing a green multi-modal dry port-seaport logistics network enhanced by the internet of things and machine learning. Computers & Industrial Engineering, 207, 111270.

Russell, S., & Norvig, P. (1995). A modern approach. In Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs (Number 27).

Saber, S. M., Thowai, K. Z., Rahman, M. A., Hassan, M. M., Bari, A. B. M. M., & Raihan, A. (2025). High-accuracy prediction of vessels' estimated time of arrival in seaports: a hybrid machine learning approach. Maritime Transport Research, 8, 100133.

Silva, M. R. F. da, Frazzon, E. M., & Silva, V. M. D. (2024). Design of flexible truck appointment system based on machine learning approach. International Journal of Logistics Systems and Management, 48(2), 244–266. https://doi.org/10.1504/IJLSM.2024.139958

Siri, S., Palmiere, A., & Ambrosino, D. (2024). Multi-Objective Optimization Methods for Train Load Planning in Seaport Container Terminals. IEEE Transactions on Automation Science and Engineering, 21(3), 3216–3228. https://doi.org/10.1109/TASE.2023.3277083

Song, P., Zhu, H., & Yang, Q. (2026). Short-term high-volatility power load forecasting in smart port energy systems using FeatureGating-BiLSTM enhanced by DualAttention mechanisms. Energy Conversion and Management, 348, 120664.

United Nations. (2026). United Nations Sustainable Development Goals. https://sdgs.un.org/goals

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), 118–144.

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly, 26(2), xiii–xxiii. https://doi.org/10.1.1.104.6570

Xie, H., & Ambrosino, D. (2025). Operations Research, Machine Learning, and Integrated Techniques for Decision Problems in the Seaside Area of Container Terminals. Operations Research Forum, 6(2), 54.

Xie, Y., Song, D.-P., Dong, J., & Feng, Y. (2025). Predicting out-terminals for imported containers at seaports using machine learning: Incorporating unstructured data and measuring operational costs due to misclassifications. Transportation Research Part E: Logistics and Transportation Review, 202, 104331.

Yan, Q., Song, R., Kim, K.-H., Wang, Y., & Feng, X. (2025). Optimizing container relocation operations by using deep reinforcement learning. Maritime Policy & Management, 52(8), 1288–1310.

You, S.-S., Cuong, T. N., Kim, H.-S., & others. (2025). Optimizing quay crane scheduling using deep reinforcement learning with hybrid metaheuristic algorithm. Engineering Applications of Artificial Intelligence, 143, 110021.

Yu, B., Wang, X., & Wang, H. (2025). A geometric model with stochastic error for abnormal motion detection of portal crane bucket grab. Engineering Applications of Artificial Intelligence, 139, 109481. https://doi.org/https://doi.org/10.1016/j.engappai.2024.109481

Zhang, Z., Ding, Y., Zhu, T., Chen, K., Wang, W., & Swee, S. Y. K. (2025). Digital twin-based energy efficiency evaluation for vessel operations in container terminals. Transportation Research Part D: Transport and Environment, 148, 104978. https://doi.org/https://doi.org/10.1016/j.trd.2025.104978

Zhou, B., & Wang, Z. (2024). An Automated Method for Container Counting in Satellite Images Based on Grid Analysis and Shadow Recognition. IEEE Access, 12, 110432–110446. https://doi.org/10.1109/ACCESS.2024.3440041

Zygouras, N., Troupiotis-Kapeliaris, A., & Zissis, D. (2024). EnvClus∗: Extracting Common Pathways for Effective Vessel Trajectory Forecasting. IEEE Access, 12, 3860–3873. https://doi.org/10.1109/ACCESS.2023.3349081

DOI
https://doi.org/10.18690/um.fov.4.2026.53

ISBN
978-961-299-152-4

Negotiating Usability and Security in Digital Transformation: A Qualitative Study of Everyday Work in Finnish Vocational Education

Anna-Liisa Ojala,[1] Tuomo Sipola,[2] Karo Saharinen[2]

[1] Jamk University of Applied Sciences, School of Professional Teacher Education, Jyväskylä, Finland
anna-liisa.ojala@jamk.fi
[2] Jamk University of Applied Sciences, Institute of Information Technology, Jyväskylä, Finland
tuomo.sipola@jamk.fi, karo.saharinen@jamk.fi

This study examines how employees in Finnish vocational education institutions experience digitally mediated work and balance tensions between usability and digital security. Based on 25 qualitative interviews and complementary observations, the study shows that everyday work is shaped by fragmented digital environments, where multiple systems, changing interfaces, and external collaborations increase the complexity of work. The findings demonstrate that security measures, such as access restrictions and authentication practices, both enable and constrain work, often introducing friction into everyday tasks. In response, employees develop informal practices, including workarounds and the use of non-official systems, to maintain workflow. The findings show that employees continuously balance usability and security in fragmented multi-system environments shaped by changing digital systems and organisational boundaries. The study suggests that usability and security should be understood as continuously negotiated aspects of everyday digitally mediated work.

Keywords:
digital transformation, usability–security tension, vocational education, shadow IT and workarounds, qualitative research

University of Maribor Press

1 Introduction

Digital transformation has reshaped everyday work across organisational settings, including educational institutions. In vocational education, work is increasingly mediated by a wide range of digital systems, services, and processes that structure how employees perform their tasks and interact with students, colleagues, and external stakeholders. Previous research has highlighted the importance of digital employee experience in understanding how workers interact with organisational technologies and digital environments (Moganadas & Goh, 2022). Studies on usability and security have shown a recurring tension between systems designed for efficient task completion and those enforcing strict security measures (Adams & Sasse, 1999; Norman, 2005a; Kainda et al., 2010; Agboola et al., 2024). While usability supports efficiency and user satisfaction, security measures such as authentication protocols and access restrictions may introduce friction into everyday work practices.

This tension is particularly relevant in vocational education institutions, where digital systems are used across pedagogical, administrative, and support functions and where work often involves collaboration with external actors such as guardians and public service providers. Vocational education also includes authentic work-based learning environments, such as training restaurants and workshops, where students interact with real customers and digital systems may involve handling customer data.

In addition, vocational education extends beyond educational institutions, as students complete practical training in workplaces and other organisations. This creates additional interfaces between organisational systems, practices, and data protection requirements. As vocational education is closely connected to working life, ongoing digitalisation across industries also shapes institutional digital infrastructures and employees' skill requirements (Busse et al., 2019).

Despite growing interest in digital transformation and employee experience, limited empirical research has examined how employees balance usability and digital security in everyday digitally mediated work. Previous research in vocational education has emphasised digital competences, including data protection and cybersecurity skills (Koramo et al., 2018; Paananen et al., 2023), while studies have also shown that digital transformation may be experienced as burdensome and contradictory

(Alanko-Turunen et al., 2024). This paper examines how employees in Finnish vocational education institutions experience digitally mediated work and balance usability and digital security in everyday practices. The study is based on qualitative interview data collected from three vocational education institutions and is guided by two research questions:

1. What kinds of digital systems and service environments shape everyday work in vocational education institutions?
2. How do employees negotiate tensions between usability and digital security in digitally transformed everyday work?

The study contributes to research on digital transformation by showing how fragmented system environments and security requirements shape everyday digitally mediated work in vocational education.

2 Theoretical background

Digital transformation has reshaped work environments through the increasing use of digital systems that mediate communication, information management, and task execution. These interactions form what has been described as the digital employee experience, referring to employees' engagement with organisational digital systems and processes throughout their work (Moganadas & Goh, 2022; Gheidar & ShamiZanjani, 2020). A key dimension of this experience is how usable employees perceive these systems to be.

Usability is commonly understood as the extent to which a system enables users to achieve their goals effectively, efficiently, and with satisfaction in a given context of use (ISO 9241-11:2018). In addition to effectiveness and efficiency, usability also includes dimensions such as learnability and memorability, which are particularly relevant in environments where employees must work with multiple systems and frequently changing interfaces (Kainda et al., 2010). These dimensions shape how smoothly employees can carry out their work and how easily they can adapt to new or evolving digital tools. In this study, usability is approached through the framework by Kainda et al. (2010), which conceptualizes usability through dimensions such as effectiveness, efficiency, satisfaction, learnability, and memorability. This framework

is particularly useful for examining how usability and security are intertwined in everyday work practices.

At the same time, digital work environments are increasingly shaped by requirements related to data protection, information security, and cybersecurity. Together, these requirements regulate how information is accessed, processed, shared, and protected in organisational digital environments (Maurer & Morgus, 2014; Althonayan & Andronache, 2018; Cains et al., 2022).

Previous research has consistently shown that usability and security are closely intertwined. Security mechanisms such as authentication procedures, access controls, and restrictions are designed to protect systems and data, but they may also complicate system use and introduce additional steps into work processes (Adams & Sasse, 1999; Norman, 2005b). Poor usability can lead users to bypass security measures, while overly strict security controls may hinder task performance and reduce system acceptance (Kainda et al., 2010; Lennartsson et al., 2021; Agboola et al., 2024). This creates a tension between usability and security, where improvements in one dimension may negatively affect the other.

Prior research has also identified the use of informal technologies and practices in organisational work. Shadow IT refers to systems or applications that are not officially approved by the organisation and often emerge when formal systems are perceived as restrictive or insufficient for everyday work (Silic & Back, 2014). Similarly, workarounds describe alternative ways of completing tasks when formal systems or procedures hinder work practices (Azad & King, 2008; Alter, 2014). While such practices may improve workflow efficiency, they may also introduce information security risks and reflect tensions between usability and security.

In complex organisational environments, such as vocational education institutions, this tension is further shaped by the diversity of roles, tasks, and system interfaces. Employees operate across pedagogical, administrative, and support functions, often in collaboration with external stakeholders. These conditions increase the likelihood that usability and security requirements come into conflict in practice. As a result, employees may use informal practices and non-official systems to maintain the flow of work when formal systems or security measures do not align with practical needs. This study examines how employees balance usability and security in everyday

digitally mediated work. Usability–security tensions emerge in situations where organisational requirements, work practices, and digital systems do not fully align.

3 Method

3.1 Data collection

The data consists of 25 thematic interviews (Amis, 2005), conducted in spring 2024 across three Finnish vocational education institutions operating in five locations. The total duration of the interviews was just under 22 hours, with individual interviews lasting on average slightly over 50 minutes. The interviews were conducted in Finnish and carried out by one or two members of the research team at a time. The interviews were mainly conducted individually; however, one was conducted as a small group interview with three participants. All participating institutions operated across multiple locations, and one of the institutions was a special needs vocational education provider.

The interviewees represented a range of organisational roles, including teaching staff, administrative personnel, IT and digitalisation specialists, student support services, and institutional leadership. This allowed the study to capture diverse perspectives on digitally mediated work across the organisation. The interview themes focused on employees' everyday work, the digital systems they use, and their experiences of usability and digital security. The interviews were semi-structured and followed a shared interview guide, while allowing flexibility for participants to describe their work practices and experiences in their own terms.

In addition to interviews, a small set of observations was conducted in selected learning environments within the participating institutions. The observations focused on everyday work practices and the use of digital systems in practice. These observations were used to support the interpretation of interview data, particularly in identifying situations where usability and security requirements come into tension.

All interviews were recorded and transcribed using the Whisper speech recognition system (Radford et al., 2023), implemented as a local solution so that the data did not leave the organisation's systems. The transcripts were subsequently reviewed and

edited manually. The data was anonymised during transcription to remove identifiable information related to individuals and organisations.

3.2 Data analysis

The data was analysed using qualitative thematic analysis (Braun & Clarke, 2006). The analysis proceeded in two phases. In the first phase, the interview transcripts were read and coded to identify recurring themes related to digital systems, work practices, and experiences of usability and security. The data was coded systematically following the research questions (Ruusuvuori et al., 2010). The coding process combined themes emerging from the interview data with concepts related to usability, security, workarounds, and shadow IT identified in previous research. Relevant excerpts were collected into an analysis document, where initial observations and interpretations were recorded. In addition, a table was constructed to map the digital systems mentioned in the interviews and the tasks associated with them. This supported the identification of patterns across roles and organisations.

In the second phase, the identified themes were further interpreted in relation to the concepts of usability and digital security. Particular attention was paid to situations where usability and security came into tension. The analysis focused on how employees described these situations and responded to them in everyday work. Through this process, the analysis identified patterns related to fragmented digital environments, security-related work constraints, and the use of informal practices and non-official systems to maintain workflow. These themes form the basis of the findings presented in the following section.

4 Findings

4.1 Fragmented digital work environment

The interview data shows that there is no single, shared digital work environment in vocational education institutions. Instead, employees work within a fragmented landscape of digital systems that varies across roles, tasks, and organisational contexts. This finding is in line with previous research highlighting the increasing complexity of digital environments in vocational education (e.g., Koramo et al., 2018; Paananen et al., 2023).

Interviewees described using a wide range of systems to carry out their work, including learning management systems, administrative platforms, communication tools, and specialised systems related to student support and external collaboration. The number and type of systems varied depending on the employee's role. Teaching staff emphasised learning platforms and communication tools, while administrative staff and IT specialists worked across multiple management and back-end systems. As a result, work is distributed across several digital systems, and even simple tasks may require the use of multiple platforms with different interfaces, logics, and requirements.

The data also shows that systems are not always well integrated. Employees described situations where the same information had to be entered into multiple systems or where systems did not communicate effectively with each other. This increases the complexity of work and creates uncertainty about where up-to-date information is located. These challenges require employees to develop practical ways of managing information across systems. The fragmented nature of the digital environment forms the context in which usability and security are experienced in practice. It also creates situations where employees must balance workflow efficiency, accuracy, and security-related organisational requirements.

4.2 Security measures as constraints on everyday work

Security-related requirements shape everyday work in vocational education institutions in multiple ways. The interview data shows that information security, data protection, and access control are embedded in digital systems and influence how employees can carry out their tasks. A central theme in the data concerns access rights and the management of information. Employees described situations where access to information is restricted in order to protect sensitive data, but where these restrictions may also complicate everyday work. Managing the balance between limiting access and enabling work was described as a continuous challenge. As one vice-rector noted, "The key issue is how to restrict information management and access rights in a way that does not prevent work tasks." This reflects the broader challenge of balancing usability and security in digital systems (Kainda et al., 2010; Agboola et al., 2024).

In practice, security measures such as access restrictions, authentication procedures, and system-specific permissions introduce additional steps into work processes. These measures are intended to ensure compliance and protect data, but they can also slow down work or make it more cumbersome. Employees described situations where they needed to log into multiple systems, manage different credentials, or navigate restrictions that limited their ability to access or share information when needed.

More concretely, the data provides several examples of how these security measures may constrain work in practice. Technical requirements such as VPN connections may, in some cases, prevent work altogether when systems fail. As one administrative staff member described: "The VPN didn't work during those in-between days… fortunately, I had taken paper documents, which I then archived. The system just didn't work… I simply couldn't do my job."

Similarly, multi-factor authentication (MFA), while recognised as an important security mechanism, was often experienced as frustrating and disruptive to workflow. A vocational teacher described the use of multi-factor authentication the following way: "Two-factor authentication has now been introduced. It's frustrating! The phone isn't necessarily right next to you, so you have to go and get it."

Access control mechanisms also emerged as a source of tension. While restricting access to information is essential for data protection, employees reported that access rights do not always align with practical work needs. In some cases, restricted access made it more difficult to obtain necessary information, while in others, broad access created challenges in managing large amounts of visible data. Many interviewees, particularly those working in administrative roles, reported that in some situations they had access to more information than necessary:

Interviewer: That was perhaps the next question — do you feel that you sometimes see too much information?

Interviewee: Yes, definitely. Yes, I do… although, of course, it's not really… well.

Interviewer: I suppose you also become somewhat desensitized to that information?

Interviewee: Yes. I mean, yes - it becomes normal that everything just sort of appears all at once.

At the same time, many of these professionals noted that it is part of employees' professional competence to manage this exposure to information without becoming overwhelmed.

The data also highlights situations where secure system use conflicts with the practical demands of work. Interview data and observations both point to such situations. For example, in some learning environments, such as simulated workplaces within educational settings (including settings like on-campus training restaurants or automotive workshops) digital systems may contain sensitive customer data. At the same time, the smooth running of work may require systems to remain easily accessible to multiple users. In such situations, practices that support workflow efficiency often conflict with data protection requirements and introduce security risks.

These findings indicate that security is not experienced as a separate layer of digital systems, but as something that directly shapes how work is carried out. Security measures can both enable and constrain work, and employees must continuously navigate these constraints in order to complete their tasks.

4.3 Informal practices and shadow IT

The data shows that employees do not rely solely on officially provided systems and practices in their work. Instead, they develop informal ways of working and use non-official tools to support task completion when formal systems do not adequately meet their needs. One form of this is the use of tools that are not part of the organisation's official system environment. As one vice principal described, their institutional work relies on a set of core systems, such as Microsoft 365 and the organisation's Moodle environment, which are largely under institutional control. At the same time, employees make extensive use of a wide range of third-party applications in teaching, communication, and everyday work. These systems fall outside institutional oversight, making it difficult to monitor their use or manage data within them. As the interviewee emphasised, this requires employees to develop an understanding of what kinds of information can be shared across different

platforms. Notably, the interviewee did not advocate for strict prohibition, but rather for informed and responsible use of systems. In this case, the institution also provided support for the use of non-core systems, viewing this as a more secure approach than attempting to prohibit their use altogether.

In addition, employees described using various communication tools to coordinate work, communicate quickly, or share information in situations where official systems were perceived as too slow, complex, or restrictive. The messaging application WhatsApp was widely mentioned in the data as an example of shadow IT. Although its use for work-related purposes was explicitly prohibited in two of the participating institutions, employees nevertheless used it in practice when necessary. As one teacher noted: "We are not allowed to use WhatsApp… however, teachers still use it via their personal phones to communicate with students, as it remains the most effective communication channel."

In addition to using non-official systems, some employees described adapting their use of digital systems and practices to better fit the demands of everyday work. Employees developed these adaptations when formal systems did not fully support everyday work tasks. These findings indicate that these informal practices are closely linked to the tension between usability and security. When formal systems are perceived as difficult to use or when security measures restrict access or slow down work, employees seek alternative ways to complete their tasks. In this sense, shadow IT and workarounds emerge when system design does not align with everyday work requirements. At the same time, these practices may introduce new risks. The use of non-official systems can challenge organisational control over data and communication, while workarounds may bypass intended security mechanisms. Employees balance workflow efficiency against data protection and information security requirements in these situations. This interpretation aligns with previous research on shadow IT and workarounds in organisational settings (Silic & Back, 2014; Alter, 2014).

4.4 External interfaces and system change

The data indicates that digitally mediated work in vocational education institutions is not limited to internal systems, but it extends to interactions with external actors and systems. Employees operate in environments where information must be shared

across organisational boundaries, including collaboration with students' guardians, workplaces where students complete their practical training, and public sector organisations such as healthcare providers when students require additional support. These external interfaces introduce additional complexity to digital work. The different organisations use different systems, follow different practices, and apply different requirements related to data protection and information sharing. As a result, employees must navigate not only in their own organisation's digital environment, but also those of external partners.

The findings also highlight the impact of ongoing system changes. Employees described situations where new systems are introduced, existing systems are updated, or processes are modified. These changes require continuous adaptation and learning, and they may temporarily disrupt established work practices. More concretely, interviewees described how multiple systems and digital processes had changed within a relatively short period of time, requiring them to continuously learn new functionalities alongside their regular work tasks. In some cases, this led to situations where not all information was recorded in systems, as employees did not have sufficient time to fully adopt new practices.

In addition, system updates and interface changes were reported to occur without prior notice, forcing employees to adapt abruptly to new layouts and functionalities. For example, interviewees described situations where familiar interface elements, such as buttons or menu structures in systems like learning management platforms or administrative tools, changed unexpectedly. This reduced memorability and increased cognitive load by previously learned routines.

Together, external interfaces and system changes increase the complexity of digital work and intensify the tension between usability and security. Employees must continuously adjust their practices in response to changing systems, evolving requirements, and the need to coordinate work across organisational boundaries. Technologies, organisational structures, and everyday work practices continuously shape how usability and security are realised in practice.

5 Discussion & Conclusion

The findings contribute to existing research on the relationship between usability and security by showing how this tension is experienced and managed in everyday work practices. Previous studies have demonstrated that security mechanisms may reduce usability and lead users to bypass or resist system constraints (Adams & Sasse, 1999; Kainda et al., 2010; Norman, 2005; Agboola et al., 2024). This study shows that usability and security emerge through everyday work practices, organisational requirements, and interactions with digital systems. Employees continuously balance workflow efficiency, access to information, and security requirements in their daily work.

In the studied setting, employees continuously balance the need to protect information with the need to perform their tasks efficiently. Security requirements are not experienced as separate from work but as embedded in everyday practices. The usability–security tension is situational and shaped by specific work tasks, organisational contexts, and system environments.

The findings also contribute to research on shadow IT and workarounds by showing how these practices emerge in response to mismatches between formal systems and everyday work requirements. Prior research has conceptualized shadow IT as the use of non-approved systems (Silic & Back, 2014) and workarounds as informal practices that enable task completion despite system constraints (Alter, 2014; Azad & King, 2008). The results of this study support these perspectives.

In this study, informal practices emerge when formal systems do not adequately support everyday work. Employees adapt their use of systems and introduce alternative tools to maintain workflow continuity.

The findings further show that digitally mediated work in vocational education is shaped by interactions across organisational boundaries and by continuous system changes. Employees must coordinate work with external actors, including guardians, training workplaces, and public sector organisations, which introduces additional complexity to digital work environments. This supports previous research of Gheidar & ShamiZanjani (2020) and Moganadas & Goh (2022), which emphasises

that digital employee experience is shaped by digital interactions between employees, technologies, and the organisational environment.

The findings also have practical implications for organisations managing digitally mediated work. Digital systems and security measures should support everyday work practices across organisational boundaries and multi-system environments. Particular attention should be paid to reducing fragmentation between systems, aligning access controls with practical work needs, and supporting secure use of third-party applications instead of relying solely on prohibitions. However, the study is limited to Finnish vocational education institutions and qualitative interview data, which provide contextual depth but do not support broad generalisation. Future research could examine similar usability–security tensions in other organisational settings and through longitudinal or comparative approaches.

This study contributes to research on digital work by showing how employees balance usability and security in complex multi-system environments. The findings demonstrate how fragmented digital infrastructures, organisational boundaries, and continuous system changes shape everyday digitally mediated work in vocational education. The findings underline the importance of designing and managing digital systems in ways that align with the realities of everyday work. In vocational education, this complexity is further intensified by the integration of work-based learning environments and cross-organisational collaboration.

References

Amis, J. (2005) Interviewing for case study research. In D. L. Andrews, D. S. Mason & M. L. Silk (Eds.) Qualitative methods in sports studies, Berg, pp. 104–138.

Adams, A. & Sasse, M. A. (1999) Users are not the enemy. Communications of the ACM, 42 (12), 40–46. https://doi.org/10.1145/322796.322806

Agboola, T. O., Adegede, J. & Jacob, J. G. (2024) Balancing usability and security in secure system design: A comprehensive study on principles, implementation, and impact on usability. International Journal of Computing Sciences Research, 8, 2995–3009.

Alanko-Turunen, M., Heinilä, H. & Vainio, V. (2024) Developing digitalizing practices in vocational education – Integrating digitalization and emotions [in Finnish]. eSignals Research, 5 (1). http://urn.fi/URN:NBN:fi-fe202402055586

Althonayan, A., & Andronache, A. (2018) Shifting from information security towards a cybersecurity paradigm. Proceedings of the 2018 10th International Conference on Information Management and Engineering, 68–79.

Busse, J., Lange, A. & Schumann, M. (2019) Effects of Digitalization on Vocational Education and Training: First Results of a Qualitative Study. In N. Pinkwart & J. Konert (Eds..) DELFI 2019, 67–72. Bonn: Gesellschaft für Informatik. https://doi.org/10.18420/delfi2019_206

Cains, M. G., Flora, L., Taber, D., King, Z. & Henshel, D. S. (2022) Defining cyber security and cyber security risk within a multidisciplinary context using expert elicitation. Risk Analysis, 42 (8), 1643–1669.

Gheidar, Y. & ShamiZanjani, M. (2020) Conceptualizing the digital employee experience. Strategic HR Review, 19 (3), 131–135.

Kainda, R., Flechais, I. & Roscoe, A. W. (2010) Security and usability: Analysis and evaluation. Proceedings of the 2010 International Conference on Availability, Reliability and Security, 275–282.

Koramo, M., Brauer, S., & Jauhola, L. (2018) Digitalization in vocational education [in Finnish]. Finnish National Agency for Education, Reports and surveys 9.

Lennartsson, M., Kävrestad, J. & Nohlberg, M. (2021) Exploring the meaning of usable security – A literature review. Information & Computer Security, 29 (4), 647–663. https://doi.org/10.1108/ICS-10-2020-0167

Maurer, T. & Morgus, R. (2014) Compilation of existing cybersecurity and information security related definitions. Washington: New America.

Moganadas, S. R. & Goh, G. G. G. (2022) Digital employee experience constructs and measurement framework: A review and synthesis. International Journal of Technology, 13(5), 999–1012. https://doi.org/10.14716/ijtech.v13i5.5830

Norman, D. A. (2005a) Human-centered design considered harmful. Interactions, 12 (4), 14–19. https://doi.org/10.1145/1070960.1070976

Norman, D. A. (2005b) The Design of Everyday Things. Revised edition. New York: Basic Books.

Paananen, H., Taivassalo, M., Raitanen, T., & Nieminen, A.-P. (2023). *Digital competence in vocational education – A guide for articulating competence* [in Finnish]. Finnish National Agency for Education.

About the authors

Anna-Liisa Ojala, PhD, Adjunct Professor (Title of Docent in Finland), works as a Senior Advisor at the Professional Teacher Education College at Jamk University of Applied Sciences, Finland. Her work focuses on researching and developing teachers' everyday work and examining the digitalisation of work and learning.

Tuomo Sipola, PhD, is a Principal Researcher at the Institute of Information Technology at Jamk University of Applied Sciences, Finland. His expertise includes artificial intelligence, machine learning, and cybersecurity, particularly the use of AI in cyber defence and digital security research.

Karo Saharinen, PhD, is a Senior Lecturer at the Institute of Information Technology at Jamk University of Applied Sciences, Finland. He is responsible for pedagogical development at the institute and works in the fields of cybersecurity, computer networks, and digital security education.

Summary

This study examined how employees in vocational education institutions experience and manage usability and security in digitally mediated work. The findings show that digital work is characterized by fragmented system environments in which employees operate across multiple platforms with varying functionalities and requirements. Security measures, such as access restrictions and authentication procedures, play a central role in shaping everyday work and may at times constrain task completion. In response, employees develop informal practices and rely on non-official systems to maintain workflow, particularly when formal systems do not adequately support their tasks. The findings also highlight that digital work extends beyond organizational boundaries and is continuously shaped by system changes and external collaboration. More broadly, the study suggests that vocational education institutions do not have a unified digital work environment. Instead, everyday digital work varies across roles, tasks, and organizational contexts, resulting in diverse and role-dependent employee experiences.

Sales and Operations Planning – Augmenting Existing Maturity Models with New Practices

Henrik J. Nyman,[1] Fabian Tschirch[2]

[1] Åbo Akademi University, Faculty of Human and Social Sciences, Turku, Finland
henrik.j.nyman@abo.fi
[2] Ruhr University Bochum, Chair for Industrial Sales and Service Engineering, Germany
fabian.tschirch@isse.ruhr-uni-bochum.de

Sales and Operations Planning (S&OP) is a vital mechanism for aligning operations with financial objectives. Despite its importance, traditional S&OP maturity models often lag behind advancements in data science, resulting in a "digital blind spot" for companies implementing S&OP. This paper investigates how existing maturity models should be complemented to incorporate recent advances in data science. We arrived at our results by first synthesizing five prominent maturity models into four core dimensions: process, people and organization, performance management, and IT and data. We then conducted a systematic literature review of 24 articles, detailing how emerging technologies affect S&OP maturity in these four dimensions. Ultimately, this research provides key considerations for aligning S&OP excellence with modern technological capabilities.

DOI
https://doi.org/10.18690/um.fov.4.2026.54

ISBN
978-961-299-152-4

Keywords:
sales and operations planning, S&OP, maturity model, probabilistic forecasting, machine learning, artificial intelligence

University of Maribor Press

1 Introduction

Sales and Operations Planning (S&OP) is a staple of the modern supply chain, serving as the primary mechanism for aligning demand, supply, and financial objectives (Danese et al., 2018; Seeling et al., 2022). It has also become a distinct research stream, featured in prominent supply chain journals (see, e.g., the special issue editorial by Jonsson et al., 2021). Historically, the purpose of S&OP has been to build a consensus demand forecast and to align supply, either in the form of purchases, production, or both, against demand. Recently, it has evolved into a more holistic business planning process, integrating financial considerations and strategic objectives with supply chain management (Nyman et al., 2025; Schlegel et al., 2021). A key feature of S&OP is its cross-functional nature, focusing on joint decision-making to meet financial goals (Oliva & Watson, 2011). To navigate process complexities, organizations have historically relied on maturity models (see, e.g., Danese et al., 2018). S&OP maturity models are frameworks designed to benchmark current capabilities and provide a structured roadmap for cross-functional integration. Like maturity models in general, S&OP maturity models describe different maturity stages that allow practitioners to assess their processes against desirable characteristics (e.g., Pöppelbuß & Röglinger, 2011). For S&OP, increased maturity signifies a shift from siloed, reactive behaviors toward integrated, proactive decision-making that enhances overall company profitability.

Despite the widespread adoption of maturity models, a critical gap has emerged in existing literature. One of the authors of this paper also works as a consultant, implementing S&OP in a corporate context. Based on his experience, one area in particular falls short: the way uncertainty is considered in planning. Most maturity models were developed when supply chain planning relied on manual inputs or, at best, statistical methods. More recently, supply chain planning has evolved to manage uncertainty through probabilistic forecasting, machine learning (ML), and AI (Hu et al., 2025; Kagalwala et al., 2025). We observe that a significant shortcoming of traditional approaches is their deterministic nature; while they acknowledge uncertainty, they offer little guidance on handling volatility. Here, probabilistic planning and AI's ability to process unstructured data offer significant improvements (Hu et al., 2025; Kagalwala et al., 2025). Consequently, we argue that existing maturity models fail to account for these rapid advances in data science.

This technological lag creates a "digital blind spot" where an S&OP process might be deemed mature by traditional standards while remaining technologically obsolete (Jazairy et al., 2025). While studies highlight the benefits of probabilistic forecasting (e.g., Hewage et al., 2025; Hu et al., 2025; Klug, 2011) and call for AI integration in supply chain management (Cohen et al., 2025), current maturity models lag behind. Consequently, we argue that maturity must be re-evaluated through the lens of advancements in data science in general, probabilistic forecasting, machine learning, and AI. As such, this paper will address the following research question: what are the key areas of existing S&OP maturity models that need to be complemented to incorporate recent advancements in data science?

In Section 2, we first synthesize prominent S&OP maturity models into a baseline representing the prevalent view on maturity. In Section 3, we describe how we performed a comparative analysis between this baseline and the capabilities of emerging technologies, specifically examining how advancements in data science, probabilistic planning, and AI/ML improve planning. The results are presented in Section 4. Finally, in Section 5 we present key considerations on complementing existing maturity models, aligning S&OP maturity with recent technological advancements.

2 Related literature

In this section, we review five prominent S&OP maturity models, selected for their academic or industrial influence, and highlight their key assessment criteria for S&OP maturity. These include four academic models and one from a commercial research lab (Gartner). After synthesizing the key findings from these models, we conclude with an overview of data science advancements that impact supply chain planning and S&OP.

2.1 S&OP maturity models

Grimson and Pyke (2007) provide an exploratory study that introduces a five-stage S&OP maturity model across five dimensions. The stages depict an organization with "no S&OP process", moving to "reactive", "standard", "advanced", and "proactive" S&OP. Their framework explicitly points toward profit optimization as the goal of S&OP, shifting the focus from simply balancing demand and supply

volumes to maximizing financial outcomes. The dimensions in the maturity model include meeting and collaboration practices, the organizational structure, process measurements, information technology, and S&OP plan integration. The study concludes that a seamless plan integration (both demand and supply) at the "proactive" stage forms the basis for profit optimization. The authors note that advanced analytics is needed to achieve this, "when optimization tools are available" (p. 336), noting that meeting practices, organization, and measure support higher maturity while IT plays a more supportive role if sufficient IT capabilities are not readily available.

Wagner et al. (2014) present a maturity model designed to help firms align operational and strategic plans. They highlight how S&OP can overcome functional silos that lead to suboptimal corporate plans and inventory discrepancies. Their model evaluates maturity across four dimensions: process effectiveness, process efficiency, people and organization, and information technology. Maturity across these dimensions is evaluated starting from "rudimentary", going to "reactive", "consistent", "integrated", and culminating in "proactive". Their results indicate that many organizations operate at a "reactive" maturity level, leaving significant room for improvement.

Danese et al. (2018) note that previous maturity models tend to be descriptive in nature. In other words, earlier models define the specific characteristics needed for any given maturity stage. Moving beyond descriptive models, the authors detail the required conditions and prerequisites for moving from one maturity level to another. Their maturity model consists of five stages across four dimensions: people and organization, process and methodologies, information technology, and performance measurement. They find that while early transitions follow a serial sequence, advanced transitions require parallel actions because dimensions become highly interdependent. A key finding is the growing criticality of the "people and organization" dimension, which requires more effort as maturity increases to ensure commitment and cross-functional alignment. The stages in the model go from "no S&OP" to "reactive", "standard", "advanced", and "proactive". The study notes that advanced S&OP stages are increasingly difficult to reach due to this rising complexity in human and organizational integration.

Vereecke et al. (2018) focus specifically on demand planning as a critical cornerstone of the broader S&OP process. They develop and validate a maturity assessment model comprising six dimensions: data management, forecasting methods, forecasting systems, performance management, organization, and people management. Their empirical study reveals that while demand data is generally well-managed, significant gaps remain in forecasting systems and performance tracking. Unlike some previous studies, they find a positive correlation between company size and demand planning maturity, likely due to larger firms having greater resources for formal systems and higher integration needs.

The Gartner S&OP maturity model (Tohamy et al., 2013) offers a five-stage framework across six dimensions: outcome, process focus, organization, metrics, time horizon, and technology. In the first stage, reacting, processes are siloed and focused on short-term fixes for supply shortages using inconsistent data. The second stage, anticipating, involves basic supply-demand balancing, typically with less focus on longer-term plans. By the third stage, integrating, cross-functional alignment matures to incorporate some financial targets and broader organizational participation. The fourth stage, collaborating, shifts the focus towards a supply response that is analysed for profitability. The S&OP process is still coordinated by the supply chain organization, but ownership resides with the business unit leaders that have profit and loss responsibility. Focus is on a planning horizon one to two years ahead. Finally, the fifth stage, orchestrating, actively manages trade-offs (e.g., between availability and capital tied to inventory). At this stage, scenario planning and advanced analytics becomes important, whereby alternative plans are analysed for profitability and risk. The S&OP process is coordinated either by the supply chain organization or finance.

2.2 Current state of maturity models: a synthesis

To synthesize common themes across maturity models, we cross-examined all the dimensions outlined for each maturity model described in Section 2.1. Table 1 details how we have synthesized these based on key features derived from the maturity models in extant literature (see the last column on the right in Table 1). Next, we will explain key characteristics of each dimension in more detail, detailing how we arrived at our synthesis of S&OP maturity.

Table 1: Maturity model dimensions

Grimson and Pyke (2007)	Wagner et al. (2014)	Danese et al. (2018)	Vereecke et al. (2018)	Gartner, Tohamy et al. (2013)	This paper
Meetings and collaboration S&OP plan integration		Process and metho-dologies		Process focus Time horizon Outcome	1) Process
Organization	People and organization	People and organization	People management Organization	Organization	2) People and organization
Measure-ments	Process effectiveness Process efficiency	Performance measurement	Performance management	Metrics	3) Per-formance management
Information technology	Information technology	Information technology	Data management Forecasting methods Forecasting systems	Technology	4) IT and data

1) Process

All five maturity models characterize S&OP as a structured, iterative, and cross-functional process typically following five core steps and meetings: data gathering, demand planning, supply planning, a pre-meeting, and an executive meeting. There is a consensus that the process evolves through sequential maturity stages, moving from a reactive, "siloed" state to an integrated, proactive state where plans are developed to optimize the overall profitability of the company. A key part of proactive planning is scenario planning, whereby alternative plans and related risks are considered to support better decision-making.

2) People and Organization

An important theme in previous maturity models is that "soft issues", including culture, commitment, and organizational structure, are significant drivers for high S&OP maturity. Previous models emphasize the need for a dedicated S&OP owner or team, clear roles and responsibilities, and the participation of top management in the process. They also stress the need for a widespread understanding of the process as a driver for profitable growth. At higher stages of maturity, the process ownership

resides with finance or business unit leaders, as opposed to the supply chain organization.

3) Performance management

Previous maturity models highlight performance measurement as vital for evaluating S&OP effectiveness. This can relate to the process, like measuring meeting attendance or meeting cadence, but most maturity models focus on the business metrics in the process. Mature organizations shift from using purely functional metrics (like capacity utilization or forecast accuracy) to KPIs that manage trade-offs between, for example, customer service levels, capital tied to inventory, and total supply chain costs. Ultimately, previous literature suggests that advanced S&OP should move toward profit optimization as the primary metric of success.

4) IT and data

Previous research stresses that while information technology is a necessary enabler for advanced S&OP, it is not a primary driver for initial success. Maturity in this dimension shifts from a reliance on disparate, manual spreadsheets to centralized, integrated systems and, with higher maturity, optimization tools that support scenario planning and automation of manual tasks (e.g., in forecasting).

2.3 Advances in data science that affect supply chain planning and S&OP

There have been numerous calls for improving practices in S&OP with the help of technology (see, e.g., Jonsson et al., 2021; Nicolas et al., 2021). While previous maturity models recognize this, the full implications of recent technology advancements on the S&OP process needs clarification (see, e.g., Nicolas et al., 2021). This involves the utilization of, for example, digital twins (i.e., a digital representation of a physical system), Advanced Planning Systems (APS), and big data analytics for supply chain planning (Sengupta & Dreyer, 2023; Wolfshorndl et al., 2020; Xu et al., 2021).

Probabilistic planning in supply chain management shifts the focus from single-point deterministic models to frameworks that explicitly account for uncertainty and risk (Hewage et al., 2025). In demand planning, this implies that a single-point value

for any given point in time is replaced with a demand range with associated probabilities, representing the breadth of potential outcomes for that same time bucket (Hu et al., 2025). Utilizing techniques like Monte Carlo simulation, probabilistic planning has also been applied to risk quantification in logistics network planning (Klug, 2011). All in all, probabilistic forecasting and planning allows organizations to model a spectrum of potential outcomes rather than relying on single-point estimates, thus mitigating inventory imbalances or service failures more efficiently than with previous deterministic approaches (see, e.g., Schmitt & Singh, 2009).

AI and ML enable organizations to move beyond traditional statistical models by capturing complex, non-linear relationships. In demand forecasting, ML enriches traditional methods by incorporating external factors like weather, social media sentiment, and economic indicators (Feizabadi, 2022; Kagalwala et al., 2025; Walter et al., 2025), boosting accuracy of "traditional" statistical forecasts (Mediavilla et al., 2022), and improving supply chain performance (see, e.g., Kagalwala et al., 2025). Beyond forecasting, AI-driven insights and real-time dashboards facilitate exception-based reviews. This shift allows human intervention to focus on high-impact anomalies rather than routine data processing (Jazairy et al., 2025; Walter et al., 2025).

3 Research design

To identify how S&OP maturity models should adapt to data science developments, we conducted a systematic literature review (SLR). This approach allowed us to synthesize existing knowledge in different domains to understand their impact on the S&OP process. Furthermore, since extending a maturity model requires a rigorous conceptual foundation (Becker et al., 2009), an SLR provided a methodologically sound basis for deriving our proposed adaptations. Our research was guided by established SLR frameworks in information systems research and followed three phases: 1) planning, 2) execution, and 3) analysis and synthesis (vom Brocke et al., 2009).

Drawing on our existing knowledge of S&OP maturity, we searched for literature at the intersection of new technology, data science, and supply chain management. The search string was iteratively refined to include specific topics like probabilistic

forecasting, with the review limited to peer-reviewed journal articles and conference proceedings. To ensure comprehensive coverage across relevant fields, we utilized Google Scholar Labs, which is an AI-powered extension of Google Scholar specifically designed to retrieve relevant articles. To capture recent technological advancements, the search focused on articles from 2020 onwards, with the exception of two older, highly relevant papers on probabilistic planning identified during the search refinement.

Conducted in December 2025, the initial search yielded over 60 articles. After filtering for peer-reviewed journals and conference proceedings and screening titles and abstracts for relevance, 45 articles remained. During this stage, topics like IoT and blockchain were removed as they were deemed less impactful on a holistic planning process like S&OP. For the remaining papers, content was analyzed for applicability to S&OP process, organizational structures, performance management, and IT & data; mathematical modeling papers without empirical data were excluded. Ultimately, 24 articles were retained for final analysis.

To systematically extract implications for S&OP maturity, both authors independently analyzed the final article set to enhance reliability and minimize bias. Based on the search strings used in data gathering, articles were categorized into three groups based on Section 2.3: supply chain planning and technology, probabilistic planning, and AI/ML. We used deductive coding to analyse content in each article, clustering the results in the four synthesized maturity dimensions outlined in Section 2.2. Table 2 includes example quotes from the articles to illustrate our coding. The "process" dimension focused on impacts to structure, decision-making, and cross-functional integration. The "people and organization" dimension addressed required changes in competencies, roles & responsibilities, and governance structures. Under "performance management," we captured necessary adaptations for process effectiveness, KPIs, and profit optimization. Lastly, "IT and data" considered implications on planning applications and integrated systems, automation, forecasting algorithms, predictive and prescriptive analytics, and scenario planning.

Table 2: Coding examples

Category	Example Quotes
Impact on process	"Data which, in the past, had been considered monthly according to the planning cycle were afterwards refreshed daily, hourly or even transmitted in real time, which ... ignited the idea of event-based planning." (Ohlson et al., 2022, p. 45) "Unlike point forecasting, probabilistic forecasting encompasses a spectrum of potential outcomes, each with associated probabilities ... This enables decision-makers to evaluate associated uncertainties and risks" (Hu et al., 2025, p. 9933)
Impact on people and organization	"Planning decisions can be augmented, where the planner is responsible for assessing the system recommendations ... Alternatively, it may be automatic or autonomous" (Cohen et al., 2026, p. 25) "LLM-based assistants can help planners test pricing scenarios, examine elasticity patterns ... via natural language prompts. ... these tools ... will increasingly automate routine decisions while humans guide strategic trade-offs." (Cohen et al., 2026, p. 25)
Impact on performance management	"in inventory management ... higher quantiles (e.g. the 95th percentile) might be important to ... maintain a high service level" (Hewage et al., 2025, p. 19) "LLMs address a different bottleneck: the gap between sophisticated optimization tools and the practitioners who use them." (Cohen et al., 2026, p. 7)
Impact on IT and data	"Analysis of data from social media, point of sales, and clickstream helps to improve ... demand forecasting accuracy" (Xu et al., 2021, p. 664) "Through our [probablistic planning] analysis, we quantified the disruption threat at various locations in the supply chain and developed risk profiles for each." (Schmitt & Singh, 2009, p. 1247)

4 Results

Table 3 outlines a summary of our key findings. It reviews 24 articles with a focus on SCP and technology (SCP), probabilistic planning (PP), and AI & ML, and cross-references these with how these advancements impact the four dimensions in our synthesized S&OP maturity model. Each S&OP maturity dimension is further broken down to the topics derived from the characteristics of each dimension outlined in Section 3.1. With a summary of the key results at hand, we will move to discussing implications in Section 5.

Table 3: Recent technology advancements and their impact on S&OP maturity

Category/ Reference	Impact on process	Impact on people and organization	Impact on performance management	Impact on IT and data
SCP and technology (SCP) (Dolgui & Ivanov, 2025; Gallego-García & García-García, 2020; Holmström & Romme, 2012; Nicolas et al., 2021; Schlegel et al., 2021; Sengupta & Dreyer, 2023; Wolfshorndl et al., 2020; Xu et al., 2021) **Probabilistic planning (PP)** (Hewage et al., 2025; Hu et al., 2025; Klug, 2011; Schmitt & Singh, 2009) **AI and ML (AI)** (Cohen et al., 2025; Feizabadi, 2022; Jazairy et al., 2025; Kagalwala et al., 2025; Kochakkashani et al., 2024; Mediavilla et al., 2022; Melançon et al., 2021; Ohlson et al., 2022; Tirkolaee et al., 2021; Vlachos & Reddy, 2025; Walter et al., 2025; Younis et al., 2022)	*Process structure:* **SCP:** Enables an event driven, as opposed to a monthly, process **SCP:** Drives exception management, rather than data review **PP:** Builds a stronger foundation for simulation and "what-if" analysis **AI:** Reduces the need for data preparation/more effective error mitigation *Decision making:* **SCP:** Enables augmented decisions rather than intuition-based assessment **PP:** Enables risk-aware decision-making **AI:** AI models non-linear relationships between data for better decisions *Cross-functional integration:* **SCP:** Provides a "single source of truth" **PP:** Allows functions to align on trade-offs	*Competencies:* **SCP:** Requires technical and data literacy as well as financial acumen **PP:** Requires capability to interpret and communicate probability distributions **AI:** Requires a better understanding of non-linear relationships between variables *Roles & responsibilities:* **SCP:** Planners become exception managers rather than data integrators **AI:** S&OP owners' responsibility move to a "decision-first" role using scenario-based planning *Governance structures:* **SCP:** Automated routine decisions complement augmented human decisions **PP:** Business targets are better aligned with risks, uncertainty, and opportunities **AI:** Algorithmic bias requires management	*Process effectiveness:* **SCP:** Improves communication between departments **PP:** Moves the process from static averages to stochastic variability **AI:** Large language models can enable native language access to complex data and models *KPIs:* **SCP:** Enables holistic, rather than local, optimization of business targets **PP:** Enables risk-adjusted performance metrics *Profit optimization:* **SCP:** Enables profit optimization as a metric **SCP:** Enables better incorporation of sustainability metrics **PP:** Enables analysis of the "cost of availability" (capital tied to inventory vs service level)	*Integrated systems:* **AI:** Enables task-specific agentic AI *Automation:* **SCP:** Automates routine tasks/enables less time on data preparation **AI:** Automating, e.g., data scraping, reports *Forecasting algorithms:* **SCP:** Enables external data integration (e.g., point of sales data, social media) **PP:** Enables the quantification of uncertainty **AI:** Can enrich forecasts based on external variables *Pred.& prescriptive analytics:* **SCP:** Enables early anomaly detection **PP:** Enables prescriptive risk mitigation **AI:** Enables resource allocation that aligns with financial & strategic goals *Scenario planning:* **SCP:** Enables what-if analysis/ objective trade-offs **PP:** Enables quantification of risks and opps.

5 Discussion and conclusion

In this section, we summarize key points from Table 2. This provides an answer to our research question, outlining the key areas of existing S&OP maturity models that need to be complemented to incorporate recent advancements in data science.

Advancing S&OP *process* maturity with the help of technology requires a shift towards a structure that prioritizes exception management and "what-if" simulations over traditional static monthly data reviews (Melançon et al., 2021). This enables more sophisticated decision-making, moving the organization away from intuition-based assessments toward risk-awareness (Hu et al., 2025). Ultimately, these capabilities enhance cross-functional integration by establishing a "single source of truth," allowing disparate teams to move past conflicting organizational goals and align on trade-offs (Jazairy et al., 2025; Schlegel et al., 2021). Further, modern technology can partly automate and thus lessen the need for a separate data gathering process step (Jazairy et al., 2025).

New technology also affects *organizational* aspects of S&OP; high process maturity might require new competences. Advances in data science demands technical literacy from the S&OP team (Cohen et al., 2025), wheres a shift in focus from demand and supply balancing to profit optimization requires financial acumen (Schlegel et al., 2021). In practice, teams must master the ability to understand fundamental principles of data science, prioritize topics that require human input, and communicate more complex probability distributions rather than just static figures (Hewage et al., 2025). Consequently, roles and responsibilities are redefined; planners increasingly focus on excpetions (Walter et al., 2025), while S&OP owners use scenario planning to prioritize decision-making over a traditional prediction-first mindset (Cohen et al., 2025). Supporting this is a governance structure that complements automated routine decisions with augmented human judgment (Sengupta & Dreyer, 2023).

By enhancing S&OP *process management and measures* with technology, KPIs detail operational trade-offs to optimize profit. This allows leaders to utilize sensitivity analysis and probabilistic planning to, for example, balance the "cost of availability" against capital constraints, ensuring decisions align with strategic priorities (Hewage et al., 2025; Kochakkashani et al., 2024).

Lastly, advances in data science mean that the *IT and data* S&OP maturity dimension enables a move from reactive planning to foresight. Disparate data streams can be consolidated, while automation eliminates the manual work of data scraping and report generation (Jonsson et al., 2021; Schlegel et al., 2021). This foundation allows forecasting algorithms to incorporate external signals and quantify uncertainty, moving beyond simple historical trends (Walter et al., 2025). When combined with predictive and prescriptive analytics, organizations can detect anomalies early and receive recommendations for resource allocation. These capabilities culminate in sophisticated scenario planning that can be supported by probabilistic analysis. This allows for "what-if" analyses that assess financial trade-offs with precision, transforming S&OP into a proactive process to support growth (Cohen et al., 2025).

This paper acts as a first steps towards complementing existing S&OP maturity models. A clear limitation of the research is the amount of articles included for review; the authors have a plan to continue the research, expanding the analysis to cover further relevant research.

References

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management: A Procedure Model and its Application. *Business & Information Systems Engineering*, *1*(3), 213–222. https://doi.org/10.1007/s12599-009-0044-5

Cohen, M. C., Dai, T., Perakis, G., Agrawal, N., Allon, G., Boute, R. N., Cachon, G. P., Chen, Z., Cohen, M. A., Cristian, R., Deshpande, V., de Véricourt, F., Fransoo, J. C., Gijsbrechts, J., Harsha, P., Hu, M., Keskinocak, P., Kwon, C., Lee, H. L., … Zhang, D. (2026). *Supply Chain Management in the AI Era: A Vision Statement from the Operations Management Community* (SSRN Scholarly Paper No. 5792542). Social Science Research Network. https://doi.org/10.2139/ssrn.5792542

Danese, P., Molinaro, M., & Romano, P. (2018). Managing evolutionary paths in Sales and Operations Planning: Key dimensions and sequences of implementation. *International Journal of Production Research*, *56*(5), 2036–2053. https://doi.org/10.1080/00207543.2017.1355119

Dolgui, A., & Ivanov, D. (2025). Internet of behaviors: Conceptual model, practical and theoretical implications for supply chain and operations management. *International Journal of Production Research*, *63*(1), 1–8. https://doi.org/10.1080/00207543.2024.2372008

Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications*, *25*(2), 119–142. https://doi.org/10.1080/13675567.2020.1803246

Gallego-García, S., & García-García, M. (2020). Predictive sales and operations planning based on a statistical treatment of demand to increase efficiency: A supply chain simulation case study. *Applied Sciences*, *11*(1), 233.

Grimson, J. A., & Pyke, D. F. (2007). Sales and operations planning: An exploratory study and framework. *The International Journal of Logistics Management*, *18*(3), 322–346. https://doi.org/10.1108/09574090710835093

Hewage, H. C., Rostami-Tabar, B., Syntetos, A., Liberatore, F., & Milano, G. R. (2025). A novel hybrid approach to contraceptive demand forecasting: Integrating point predictions with probabilistic distributions. *International Journal of Production Research*, 1–35. https://doi.org/10.1080/00207543.2025.2526169

Holmström, J., & Romme, A. G. L. (2012). Guest editorial: Five steps towards exploring the future of operations management. *Operations Management Research*, *5*(1), 37–42.

Hu, H., Wu, Y., Shi, Z., Qi, Y., Zhang, J., Han, S., Kang, N., Chen, J., & Wang, L. (2025). Explainable probabilistic forecasting for time series in supply chains: A latent auto-encoder approach. *International Journal of Production Research*, *63*(24), 9933–9953. https://doi.org/10.1080/00207543.2025.2541860

Jazairy, A., Shurrab, H., & Chedid, F. (2025). Impact pathways: Walking a tightrope—unveiling the paradoxes of adopting artificial intelligence (AI) in sales and operations planning. *International Journal of Operations & Production Management*, *45*(13), 1–27.

Jonsson, P., Kaipia, R., & Barratt, M. (2021). The future of S&OP: dynamic complexity, ecosystems and resilience. *International Journal of Physical Distribution & Logistics Management*, *51*(6), 553–565. https://doi.org/10.1108/ijpdlm-07-2021-452

Kagalwala, H., Radhakrishnan, G. V., Mohammed, I. A., Kothinti, R. R., & Kulkarni, N. (2025). Predictive Analytics in Supply Chain Management: The Role of AI and Machine Learning in Demand Forecasting. *Advances in Consumer Research*, *2*, 142–149.

Klug, F. (2011). Automotive supply chain logistics: Container demand planning using Monte Carlo simulation. *International Journal of Automotive Technology and Management*, *11*(3), 254. https://doi.org/10.1504/IJATM.2011.040871

Kochakkashani, F., Kayvanfar, V., & Baldacci, R. (2024). Innovative applications of unsupervised learning in uncertainty-aware pharmaceutical supply chain planning. *IEEE Access*, *12*, 107984–107999.

Mediavilla, M. A., Dietrich, F., & Palm, D. (2022). Review and analysis of artificial intelligence methods for demand forecasting in supply chain management. *Procedia CIRP*, *107*, 1126–1131. https://doi.org/10.1016/j.procir.2022.05.119

Melançon, G. G., Grangier, P., Prescott-Gagnon, E., Sabourin, E., & Rousseau, L.-M. (2021). A Machine Learning-Based System for Predicting Service-Level Failures in Supply Chains. *INFORMS Journal on Applied Analytics*, *51*(3), 200–212. https://doi.org/10.1287/inte.2020.1055

Nicolas, F. N. P., Thomé, A. M. T., & Hellingrath, B. (2021). Usage of information technology and business analytics within sales and operations planning: A systematic literature review. *Brazilian Journal of Operations and Production Management*, *18*(3). https://doi.org/10.14488/BJOPM.2021.023

Nyman, H. J., Tschirch, F., & Hongell, L. (2025). Linking strategy to tactical planning–using integrated business planning (IBP) as a strategy implementation process. *International Journal of Physical Distribution & Logistics Management*, *55*(11), 213–239.

Ohlson, N.-E., Riveiro, M., & Bäckstrand, J. (2022). Identification of tasks to be supported by machine learning to reduce sales & operations planning challenges in an engineer-to-order context. *SPS2022: Proceedings of the 10th Swedish Production Symposium*, 39–50.

Oliva, R., & Watson, N. (2011). Cross-functional alignment in supply chain planning: A case study of sales and operations planning. *Journal of Operations Management*, *29*(5), 434–448. https://doi.org/10.1016/j.jom.2010.11.012

Pöppelbuß, J., & Röglinger, M. (2011). What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management. *ECIS 2011 Proceedings*. European Conference on Information Systems (ECIS 2011). https://aisel.aisnet.org/ecis2011/28/

Schlegel, A., Birkel, H. S., & Hartmann, E. (2021). Enabling integrated business planning through big data analytics: A case study on sales and operations planning. *International Journal of Physical Distribution and Logistics Management*, *51*(6), 607–633. https://doi.org/10.1108/IJPDLM-05-2019-0156

Schmitt, A. J., & Singh, M. (2009). Quantifying supply chain disruption risk using Monte Carlo and discrete-event simulation. *Proceedings of the 2009 Winter Simulation Conference (WSC)*, 1237–1248. https://ieeexplore.ieee.org/abstract/document/5429561/

Seeling, M., Kreuter, T., Scavarda, L. F., Thomé, A. M. T., & Hellingrath, B. (2022). The role of finance in the sales and operations planning process: A multiple case study. *Business Process Management Journal*, *28*(1), 23–39. https://doi.org/10.1108/BPMJ-07-2021-0447

Sengupta, S., & Dreyer, H. (2023). Realizing zero-waste value chains through digital twin-driven S&OP: A case of grocery retail. *Computers in Industry*, *148*, 103890.

Tirkolaee, E. B., Sadeghi, S., Mooseloo, F. M., Vandchali, H. R., & Aeini, S. (2021). Application of Machine Learning in Supply Chain Management: A Comprehensive Overview of the Main Areas. *Mathematical Problems in Engineering*, *2021*(1), 1476043. https://doi.org/10.1155/2021/1476043

Tohamy, N., Tarafdar, D., Kohler, J., & Pukkila, M. (2013). *Introducing the five-stage sales and operations planning maturity model for supply chain leaders*. Gartner Inc.

Vereecke, A., Vanderheyden, K., Baecke, P., & Van Steendam, T. (2018). Mind the gap – Assessing maturity of demand planning, a cornerstone of S&OP. *International Journal of Operations and Production Management*, *38*(8), 1618–1639. https://doi.org/10.1108/IJOPM-11-2016-0698

Vlachos, I., & Reddy, P. G. (2025). Machine learning in supply chain management: Systematic literature review and future research agenda. *International Journal of Production Research*, *63*(16), 5987–6016. https://doi.org/10.1080/00207543.2025.2466062

vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. *ECIS 2009 Proceedings*. European Conference on Information Systems (ECIS 2009). https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1145&context=ecis2009

Wagner, S. M., Ullrich, K. K. R., & Transchel, S. (2014). The game plan for aligning the organization. *Business Horizons*, *57*(2), 189–201. https://doi.org/10.1016/j.bushor.2013.11.002

Walter, A., Ahsan, K., & Rahman, S. (2025). Application of artificial intelligence in demand planning for supply chains: A systematic literature review. *The International Journal of Logistics Management*, *36*(3), 672–719. https://doi.org/10.1108/IJLM-02-2024-0120

Wolfshorndl, D. A., Vivaldini, M., & De Camargo Junior, J. B. (2020). Advanced Planning System as support for Sales and Operation Planning: Study in a Brazilian automaker. *Global Journal of Flexible Systems Management*, *21*(S1), 1–13. https://doi.org/10.1007/s40171-020-00236-8

Xu, J., Pero, M. E. P., Ciccullo, F., & Sianesi, A. (2021). On relating big data analytics to supply chain planning: Towards a research agenda. *International Journal of Physical Distribution & Logistics Management*, *51*(6), 656–682.

Younis, H., Sundarakani, B., & Alsharairi, M. (2022). Applications of artificial intelligence and machine learning within supply chains: systematic review and future research directions. *Journal of Modelling in Management*, *17*(3), 916–940. https://doi.org/10.1108/JM2-12-2020-0322

IT TAKES TWO TO CO-CREATE - DIGITAL TWIN MECHANISM FOR VALUE CO-CREATION

ROBERT SCHMELZER

Chemnitz University of Technology, Chair of Business Information Systems – Business Process and Information Management, Chemnitz, Germany
robert.schmelzer@wirtschaft.tu-chemnitz.de

The convergence of physical assets and dynamic digital representations is reshaping how value is created in service ecosystems. Digital twins (DTs) are at the center of this development, yet a comprehensive understanding of how they influence the co-creation of value remains lacking. A systematic literature review of 53 peer-reviewed articles on DTs in service ecosystems was conducted. Drawing on service-dominant logic (SDL) as an analytical lens and expanding this perspective with Giddens' theory of structuration, the analysis identifies five mechanisms: real-time mirroring and synchronization, simulative anticipation and scenario experimentation, feedback-driven adaptive learning and reconfiguration, interactive co-creation and participatory design, and cross-boundary resource orchestration. Building on these two theoretical perspectives, a conceptual framework is developed that positions DTs as socio-material mediators of structuration in service ecosystems. In doing so, the SDL discourse is expanded by illustrating how digital technologies condition and transform resource integration, while offering implications for DT-based service design.

DOI
https://doi.org/10.18690/um.fov.4.2026.55

ISBN
978-961-299-152-4

Keywords:
digital twin,
value co-creation,
service-dominant logic,
service ecosystems,
structuation
resource integration

University of Maribor Press

1 Introduction

Technological advances in industry, energy, infrastructure and mobility are enabling increasingly dynamic and intelligent service offerings (Alter, 2020). These services emerge within service ecosystems, where stakeholders are connected through the exchange of services (Lusch & Nambisan, 2015). Digital twins (DTs) are establishing themselves as a key technology in these sectors, linking physical assets to digital representations via continuous data streams and enabling smart services through simulation and operational control (Enrique & Soares, 2022; Grieves & Vickers, 2017; Lim et al., 2020; Wache, 2023). Although Service-Dominant Logic (SDL) is often used to conceptualise value creation in such contexts (Schmidt et al., 2026), it only partially explains how DTs influence value creation processes.

Recent research, particularly in the context of digital platforms, has begun to examine how specific digital technologies shape interactions, coordination, and innovation (Schmidt et al., 2026). For instance, some scholars argue that DTs should be understood as resources used by interdependent actors to integrate information and co-create value within networks (Camposano et al., 2021). Other emerging frameworks draw on SDL to explain how DTs structure value-creation processes, such as adaptability and responsiveness, through resource integration in service design (Zarco, 2024). However, existing research remains fragmented and predominantly focuses on domain-specific approaches or isolated facets. However, a comprehensive conceptual explanation of how DTs enable value co-creation within service ecosystems remains lacking. In particular, it is unclear how DTs impose new conditions on value cp-creation and which facets are affected. This gap leads to the following research question (RQ): *How do digital twins shape value co-creation in service ecosystems?*

This paper addresses this gap by conceptualizing DTs as a factor that affects resource integration and value co-creation. Two interrelated objectives are pursued. First, the central mechanism through which DTs influence resource integration for value co-creation are identified. Second, a conceptual framework illustrating the ideal-typical integration of DTs is developed. This represents a first step toward extending SDL to account for artifact-mediated resource integration in value co-creation, with DTs as a paradigmatic case.

The remainder of this paper outlines the theoretical foundation, methodology and identifies five mechanisms through which DTs shape value co-creation. Finally, a conceptual framework is proposed, positioning DTs as socio-material mediators and conclude with implications, limitations, and future research.

2 Foundation

2.1 Service-Dominant Logic

According to the SDL conceptualization, value is not inherent in goods but emerges as the result of resource integration through the exchange of services between actors in service ecosystems (Lusch & Nambisan, 2015; Vargo & Lusch, 2016). Value co-creation is the collaborative process in which actors apply their knowledge and skills (operant resources) to integrate physical goods, data, and other resources (operand resources) in context, thereby realizing utility value (Lusch & Nambisan, 2015; Vargo & Lusch, 2016). Service ecosystems are accordingly self-regulating systems of resource-integrating actors coordinated by institutional arrangements. Vargo & Lusch (2016) distilled five axioms from this perspective, summarized in Table 1.

Table 1: The axioms of S-D logic according to Vargo & Lusch (2016)

Axiom	Description
1	Service is the fundamental basis of exchange.
2	Value is co-created by multiple actors, always including the beneficiary.
3	All social and economic actors are resource integrators.
4	Value is always uniquely and phenomenologically determined by the beneficiary.
5	Value cocreation is coordinated through actor-generated institutions and institutional arrangements.

Source: Vargo & Lusch (2016)

2.2 Technology and Resource Integration

Although SDL has traditionally emphasized human and organizational actors, the role of technology in value co-creation is increasingly recognized (Barile et al., 2021). Digital technologies shape processes of co-design, co-development, co-delivery, and co-learning across all levels of service ecosystems (Barile et al., 2020), yet existing analyses often treat technologies as enablers or platforms without explaining how specific artifacts influence the conditions under which resources become visible,

combinable, and coordinable (Barile et al., 2020). In data-intensive environments, digital artifacts possess material properties enabling actors to manage visibility and control resources across organizational boundaries (Balta et al., 2021; Herterich et al., 2016; Payne et al., 2021a). This allows digital technologies to become structure-forming element for resource combination (Adi et al., 2023). As AI-based systems increasingly sense, learn, and adapt autonomously, artifacts develop their own agency and participate in decision-making, therefore transforming service system configurations (Barile et al., 2021; Hwang et al., 2026; Mele & Russo-Spena, 2024; Payne et al., 2021b).

SDL alone does not fully explain how digital artifacts shape value creation. Structuration theory (Giddens, 1984) addresses this by conceptualizing structure as recursively produced through action. This aligns with SDL, where rules and norms both enable and are shaped by value creation (Vargo & Lusch, 2016). Here, structure corresponds to institutional arrangements (Axiom 5), while resource-integrating practices reflect agency (Axioms 1 and 3). While Giddens largely neglects technology, later work such as sociomateriality (Orlikowski, 2007) emphasizes the entanglement of technological artifacts and social practices, a link also reflected in SDL.

2.3 Digital Twins as Mediating Artifacts

DTs link physical entities with virtual counterparts, enabling data-driven services such as monitoring and control (Grieves, 2014; Tao et al., 2018). While definitions vary (Tuegel et al., 2011; van der Valk et al., 2020), what distinguishes DTs from other digital artifacts is the simultaneous combination of a persistent virtual representation, bidirectional synchronization, and embedded simulation. From an SDL perspective, DTs function as platforms for service innovation, enabling actors to integrate information, dynamically adapt to customer needs, and jointly provide new services (Camposano et al., 2021; Stoll et al., 2020; Zarco et al., 2024).

Combining these perspectives yields an integrated analytical framework. SDL establishes that value emerges through collaborative resource integration coordinated by institutional arrangements. Structuration theory, extended by the sociomaterial lens, accounts for the recursive relationship between digital artifacts and the structures they enact and transform; the technology literature specifies how

DTs mediate this integration by making resources visible, enabling virtual experimentation, and coordinating actors (Zarco et al., 2024; Zhu et al., 2026). Together, these perspectives capture both the collaborative nature of value co-creation and the structural role of DTs in reconfiguring resource integration in service ecosystems.

3 Research Approach

A systematic literature review guided by Webster & Watson (2002) was conducted to identify and classify prior SDL studies examining how DTs shape value co-creation in service ecosystems. Following Paré et al. (2015), this study adopts a mapping approach to identify concept characteristics and derive underlying mechanisms.

The search was conducted across five relevant databases (see Figure 1). To ensure conceptual fit, articles were included if they (1) address DTs in service ecosystems, (2) align with the chosen conceptual definition of DTs, and (3) provide sufficient detail on resource integration, value co-creation, and institutional arrangements. The query combined "digital twin" and "service ecosystem" with related terms and was limited to English-language, peer-reviewed literature, resulting in 174 initial articles. After excluding duplicates, the remaining articles were screened for relevance, first based on titles and abstracts, and finally on the full text (Webster & Watson, 2002).

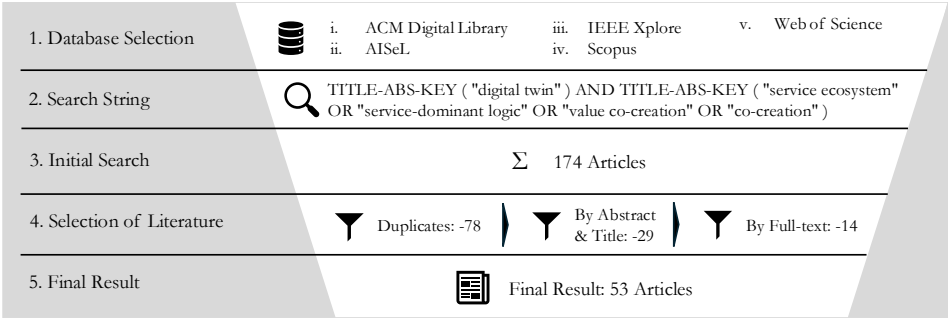

Figure 1: Overview literature research
Source: Own

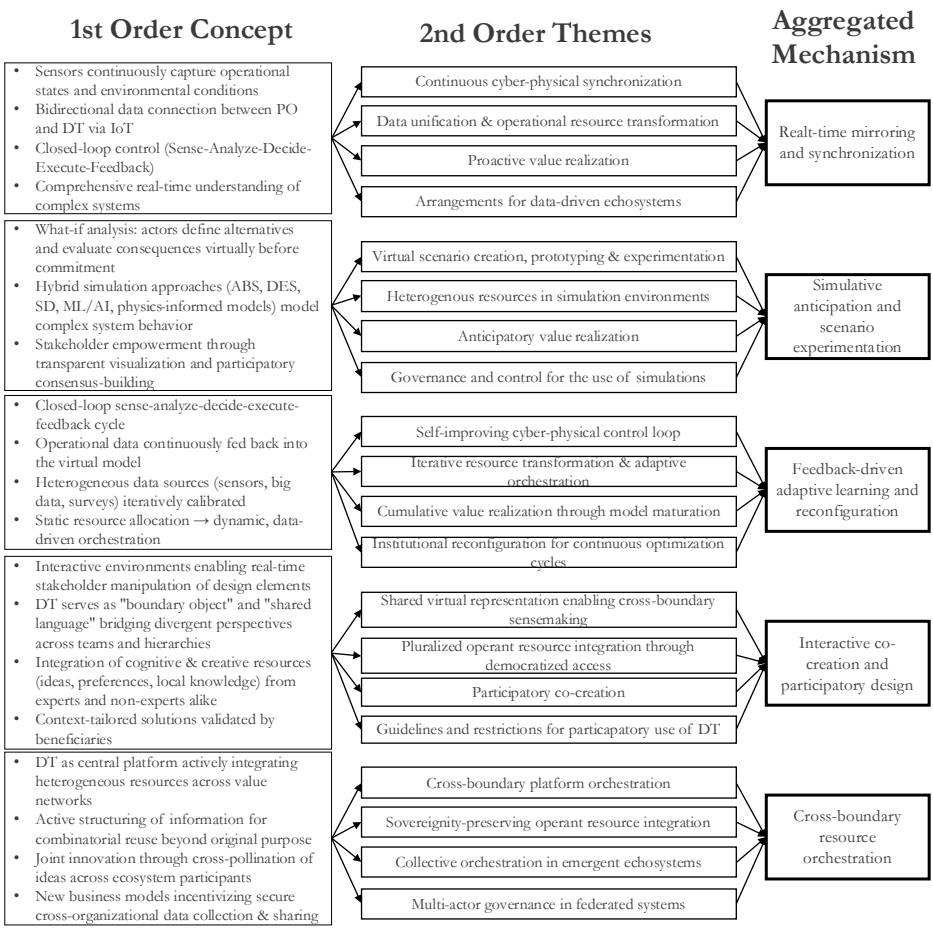

Figure 2: Data structure following the approach of Gioia et al. (2013)
Source: Own

This study adopts a multi-level coding approach inspired by Gioia et al. (2013), using SDL axioms as sensitizing devices to examine how resource integration, value creation, and institutional arrangements manifest in DTs. In the aggregation phase, structuration theory and its sociomaterial extension (Section 2) guided the organization of emerging themes along the duality of structure and actor practises. This two-layered strategy allowed us to ground first-order concepts in SDL vocabulary while situating the aggregated mechanisms within the broader dynamics of how digital artifacts condition and transform structural arrangements of service ecosystems.

Each article was systematically coded and reviewed, yielding first-order concepts that describe DTs within service ecosystems (Gioia et al., 2013). In a subsequent step, these codes were compared and iteratively refined. By continuously relating them to existing literature, particularly SDL concepts such as resource integration (Vargo & Lusch, 2016), common second-order themes were interpretively derived. Guided by the duality of structure and actor practices as an interpretive frame, captured themes were then aggregated into overarching mechanisms. Each mechanism is traceable to a distinct cluster of second-order themes (see Figure 2), that illuminate the role of DTs within service ecosystems and their interdependencies. Figure 2 shows the resulting data structure and presents a selective subset of first-order concepts.

4 Digital Twin SDL Supporting Mechanism

4.1 Real-time mirroring and synchronization

Real-time data acquisition and processing form the cornerstone of DT functionality, enabling the continuous mirroring and synchronization of physical systems within service ecosystems (Tao et al., 2018). Heterogeneous data is continuously integrated and synchronized with the twin, while advanced implementations offer bidirectional synchronization (van der Valk et al., 2020). This shifts resource integration from sequential, isolated data exchange to continuous, cross-domain access. Available data becomes a shared operational resource that multiple actors and systems can utilize simultaneously. Furthermore, situational awareness and accelerated decision-making are enhanced (Madeira et al., 2022). The mechanism also provides the foundational infrastructure for resource integration (Axiom 3) by transforming distributed data into a shared, dynamically accessible resource across the ecosystem.

4.2 Simulative anticipation and scenario experimentation

DTs function as anticipatory experimentation environments that enable actors to construct, manipulate, and evaluate alternative future states before intervening in the physical world (Verdugo-Cedeño et al., 2024). They integrate real-time and historical data, analytical models, and domain knowledge into prospective representations in which varying parameters and consequences can be explored (Meierhofer et al., 2020). Decision-making thus shifts from ex post reaction to ex ante evaluation. Actors can experiment with configurations in a low-risk setting, anticipate trade-offs

prior to committing resources, and embed simulation-based evidence into organizational practices, fostering more proactive forms of coordination and governance (Bertoni, 2023). In SDL terms, this mechanism extends resource integration (Axiom 3) into the temporal domain, making potential future resource configurations evaluable. By embedding simulation-based evidence into decision practices, it simultaneously reshapes the institutional arrangements (Axiom 5) through which value co-creation is coordinated.

4.3 Feedback-driven adaptive learning and reconfiguration

As adaptive learning systems, DTs establish recursive feedback structures that continuously integrate operational data, user interactions, and service outcomes into digital models, transforming static representations into dynamically evolving systems (Zarco, 2024). This manifests as operational closed-loop adaptation, where sensory data informs model updates, and as stakeholder-driven reconfiguration, where user feedback shapes service adaptation (Cao & Ming., 2025). In both forms, DTs connect observation, simulation, and intervention within recursive learning loops, actively contributing to shaping the systems they represent. Value creation thus shifts from discrete optimization toward continuous adaptation, giving rise to new institutional requirements such as adaptive governance structures and frameworks for algorithmic decision-making in real-time contexts (Schultheiß et al., 2023). This mechanism enhances the recursive nature of resource integration (Axiom 3) by enabling continuous feedback, experimentation, and adaption. As a result, Axiom 4 becomes more operationally generative, since beneficiaries experimental value assessments can more directly inform subsequent service configurations and institutional arrangements (Axiom 5).

4.4 Interactive co-creation and participatory design

DTs function as co-creative environments in which multiple stakeholders collaboratively explore and refine solution concepts within a shared digital representation (Trizio et al., 2024). Rather than relying on sequential feedback, actors directly intervene in the design artifact by modifying parameters, testing alternatives, and observing effects in real-time. In doing so, DTs externalize diverse forms of knowledge and render them actionable through joint activity, shifting participation from passive involvement to active co-design of both problem definitions and

solution spaces (Stoll et al., 2020). Through iterative cycles of interaction, evaluation, and adaptation, actors align their perspectives and converge on contextually grounded solutions, co-creating value not only in the outcome but throughout the process itself (Meske et al., 2021). So, from an SDL point of view, this mechanism puts Axiom 2 into action. Beneficiaries and other stakeholders become active co-creators of value instead of being passive recipients. It also illustrates how resource integration (Axiom 3) can proceed through synchronous joint activity rather than prior aggregation, with the shared representation serving as the medium through which diverse operant resources are combined.

4.5 Cross-boundary resource orchestration

Functioning as integrative infrastructures, DTs actively structure, coordinate, and recombine heterogeneous resources across organizational, lifecycle, and systemic boundaries (Meierhofer et al., 2021). They aggregate distributed data and models into a shared representation while coordinating resource flows and actor contributions within the ecosystem (Nguyen et al., 2024; Trizio et al., 2024), thereby promoting cross-lifecycle coordination and transparency (Martinelli et al., 2024). Shared value emerges through the coordinated contributions of multiple actors, enabling new services and business models (Ricci et al., 2021) and improving decision quality through integrated, context-specific information (Meierhofer et al., 2020). Taken together, the five mechanisms constitute an interconnected configuration: M1 establishes the continuously updated data foundation, M2 projects real-time states into alternative futures, M3 feeds operational outcomes back into the model, M4 opens the shared representation as a collaborative design space, and M5 extends these dynamics across organizational and lifecycle boundaries. From a SDL perspective, M5 provides the infrastructural precondition for service ecosystems in which resource integration (Axiom 3) and value co-creation (Axiom 2) extend beyond single organizations, coordinated through shared institutional arrangements (Axiom 5).

5 Digital Twins as Mediator of Structuration in Service Ecosystems

Drawing on the SDL and structuration-theoretical foundation from Section 2, this study conceptualizes how DTs facilitate value co-creation in service ecosystems. The duality of structure and actor practices serves as the analytical lens: institutional

arrangements (Axiom 5) correspond to structure, while resource-integrating and service-exchanging practices (Axioms 1 and 3) reflect agency. The sociomaterial extension of structuration theory (Orlikowski, 2007) captures the constitutive role of technological artifacts in this recursive process. DTs are argued represent a particularly radical form of this entanglement by materializing structure into a persistent, shared, and manipulable digital representation, making the duality observable and accelerating its dynamics.

Five mechanisms emerge through which DTs mediate structuring in service ecosystems. Two of these operate along the conditioning dimension, linking structure to action. Real-time mirroring and synchronization (M1) render institutional arrangements visible and accessible, transforming distributed data into a dynamically usable resource that multiple actors can integrate simultaneously (Axiom 3). Simulative anticipation (M2) extends this logic into the temporal domain, enabling actors to construct, evaluate, and compare potential future states before committing resources. As a result, decision-making shifts from reactive responses toward more forward-looking assessments, introducing a prospective dimension to resource integration. Figure 3 illustrates these relationships.

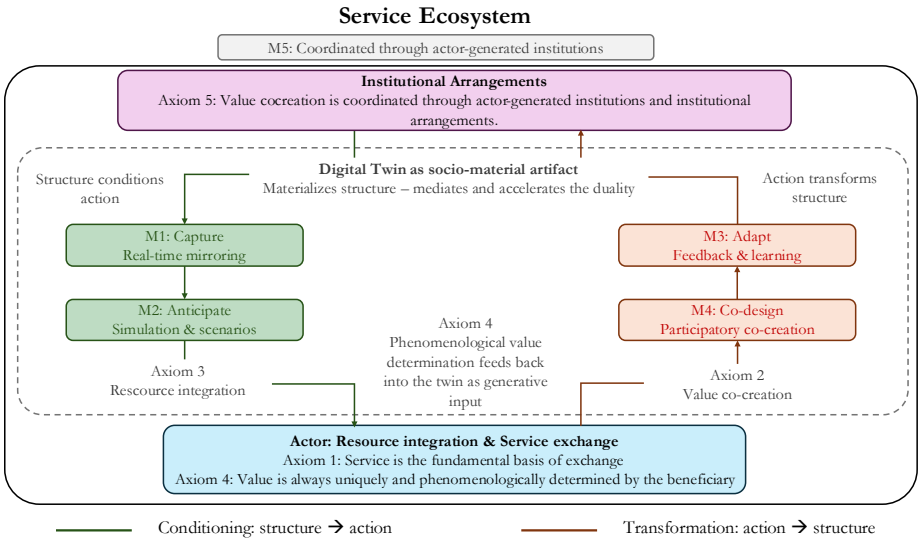

Figure 3: The digital twin as a socio-material mediator of structuration in service ecosystems
Source: Own

Two additional mechanisms operate along the transformative dimension. Feedback-driven adaptation (M3) establishes recursive learning loops through which operational data, user interactions, and service outcomes continuously reconfigure the structural conditions for subsequent action. Interactive co-creation (M4) enables multiple actors to jointly explore and iteratively refine solution designs within a shared digital representation, co-creating not only value but also the institutional arrangements coordinating future value creation (Axiom 2). Through M3 and M4, the beneficiary's phenomenological determination of value (Axiom 4) becomes operationally generative. Hence, experience-based assessments are not merely evaluative endpoints but transformative inputs that, mediated by the DT's feedback structures, reshape the institutional fabric of the ecosystem.

Finally, cross-boundary resource orchestration (M5) functions as an overarching mechanism that scales the structuring process beyond individual organizations by actively structuring, coordinating, and recombining heterogeneous resources across organizational, lifecycle, and system boundaries. This mechanism creates the infrastructural prerequisite for what Vargo & Lusch (2016) describe as service ecosystems, namely relatively self-contained and self-adapting systems of resource-integrating actors linked by shared institutional arrangements. It thereby extends the scope of the duality from the intra-organizational to the ecosystemic realm.

Overall, this conceptualization suggests that the DT is not merely a tool within an existing ecosystem but a sociomaterial artifact embedded in the structuring process. It materializes structure, renders the duality observable and actionable, and compresses Giddens' (1984) process of social reproduction into a near real-time, data-driven cycle, transforming value co-creation from being episodic and implicit into being continuous, explicit, and reflexively controlled.

6 Discussion & Conclusion

A systematic literature review of 53 articles and an SDL-oriented analysis identify five mechanisms derived from 20 second-order themes (see Figure 2), which are integrated into a conceptual framework grounded in SDL and structuration theory.

The results consolidate fragmented perspectives on DTs through the lens of SDL, extending prior work that treats DTs mainly as resources within service ecosystems.

While earlier studies remain largely case-specific (Meierhofer et al., 2021; Zarco, 2024), the developed framework synthesizes findings from diverse application contexts, including manufacturing, energy, and infrastructure, into a domain-independent set of mechanisms, thereby moving beyond case-specific accounts. By embedding technically described DT functionalities within the duality of structure and actor practises, the framework connects implementation-level descriptions with SDL´s value-oriented vocabulary. Adopting a structuration-theoretical perspective further reveals a generative dimension often overlooked: DTs not only enable resource integration but also generate, reinforce, and transform norms, rules, and practices. In this sense, DTs emerge as sociomaterial, structure-forming artifacts that actively shape value co-creation.

A key implication concerns the relationship between SDL axioms and the identified mechanisms. While the framework builds on resource integration, value determination, and institutional arrangements (Axioms 3–5), Axioms 1 and 2 remain underexplored. The predictive and autonomous capabilities of DTs may render beneficiaries more passive, however, can also empower actors as co-creators depending on how environments are designed. Future research should therefore examine how DTs reshape value determination.

This study has several limitations. DT implementations vary widely, and it cannot be fully ensured that all analyzed cases reflect core DT characteristics. In addition, DTs are often intertwined with other technologies, making it difficult to isolate a technological scope. This challenge was addressed by focusing on phenomena attributable to core DT capabilities. Moreover, SDL was not originally developed as a theory for a technological context, which required interpretive linking between mechanisms and SDL concepts. Methodologically, the coding process was conducted by a single researcher, which may introduce individual interpretive bias. To counteract this, SDL axioms were used as sensitizing devices to guide the coding process. The complete coding scheme is provided in an online appendix (link) to enable transparency.

Future studies can validate the identified relationships using approaches such as case study research. Questions of power asymmetries and governance remain open, particularly how control over DTs and data access shape actor relationships. The dynamic nature of service ecosystems also calls for longitudinal studies examining

how these mechanisms evolve in response to technological and market developments.

In particular, future research could apply the proposed mechanisms in empirical settings to test their analytical utility and refine the boundaries between conceptually adjacent mechanisms such as M3 and M4. Additionally, configurational or process-oriented approaches could help formalize the interdependencies among the analytically distinct but mutually reinforcing mechanisms.
Overall, this study advances the understanding of DTs as both components of and influential elements within service ecosystems. By identifying DT-specific mechanisms and embedding them within the duality of structure and actor practises, it connects technical implementation with SDL´s value-oriented perspective.

References

Adi, S. P. S., Ghinahana, S., Yudianto, B., & Wibowo, A. J. I. (2023). Institutions, technology and resource integration in the value co-creation process: a study of service-dominant logic in higher education. *Journal of Applied Research in Higher Education 16(5),* 1864–1883. https://doi.org/10.1108/jarhe-03-2023-0104

Alter, S. (2020). Making Sense of Smartness in the Context of Smart Devices and Smart Systems. *Information Systems Frontiers, 22(2),* 381–393. https://doi.org/10.1007/s10796-019-09919-9

Balta, M., Valsecchi, R., Papadopoulos, T., & Bourne, D. (2021). Digitalization and co-creation of healthcare value: A case study in Occupational Health. Technological Forecasting and Social Change. *Technological Forecasting and Social Change, 168*, 120785. https://doi.org/10.1016/j.techfore.2021.120785

Barile, S., Grimaldi, M., Loia, F., & Sirianni, C. (2020). Technology, Value Co-Creation and Innovation in Service Ecosystems: Toward Sustainable Co-Innovation. *Sustainability, 12(7),* 2759. https://doi.org/10.3390/su12072759.

Barile, S., Bassano, C., Piciocchi, P., Saviano, M., & Spohrer, J. (2021). Empowering value co-creation in the digital age. *Journal of Business & Industrial Marketing, 39(6),* 1130-1143. https://doi.org/10.1108/jbim-12-2019-0553

Bertoni, M. (2023). Towards Digital Immersive Experiences for Collaborative Value Co-creation in Design. In: Camarinha-Matos, L. M., Boucher, X., & Ortiz, A. (Eds.), *IFIP Advances in Information and Communication Technology. Collaborative Networks in Digitalization and Society 5.0, 688,* 193–206. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42622-3_14

Camposano, J. C., Smolander, K., & Ruippo, T. (2021). Seven Metaphors to Understand Digital Twins of Built Assets. *IEEE Access, 9*, 27167-27181. https://doi.org/10.1109/ACCESS.2021.3058009

Cao, Y., & Ming, X. (2025). Digital Twin-Driven Low-Carbon Service Design and Modularization in Central Air Conditioning Ecosystems: A Multi-Criteria and Co-Intelligence Approach. *Sustainability, 17(21),* 9877. https://doi.org/10.3390/su17219877

de Trizio, F., Sciddurlo, G., Cianci, I., Piro, G., & Boggia, G. (2024). Optimizing Key Value Indicators in Intent-Based Networks through Digital Twins aided service orchestration mechanisms. *Computer Communications*, *228*, 107977. https://doi.org/10.1016/j.comcom.2024.107977

Enrique, D.V., & Soares, A.L. (2022). Cognitive Digital Twin Enabling Smart Product-Services Systems: A Literature Review. In: Camarinha-Matos, L.M., Ortiz, A., Boucher, X., & Osório, A.L. (Eds.) Collaborative Networks in Digitalization and Society 5.0. PRO-VE 2022. *IFIP Advances in Information and Communication Technology, 662,* 77-89. Springer, Cham. https://doi.org/10.1007/978-3-031-14844-6_7

Giddens, A. (1984). The constitution of society: Outline of the theory of structuration. Polity Pr.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods, 16(1),* 15–31. https://doi.org/10.1177/1094428112452151

Grieves, M. (2014). Digital Twin: Manufacturing Excellence through Virtual Factory Replication. *Digital Twin White Paper.*

Grieves, M. W., & Vickers, J. (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Kahlen, J., Flumerfelt, S., & Alves, A. (Eds.) *Transdisciplinary Perspectives on Complex Systems,* 85–113. Springer, Cham. https://doi.org/10.1007/978-3-319-38756-7_4

Herterich, M., Eck, A., Uebernickel, F. (2016). Exploring How Digitized Products Enable Industrial Service Innovation - an Affordance Perspective. *Research Papers.*

Hwang, B., Puntha, P., & Jitanugoon, S. (2026). Human-centric AI: transforming service ecosystems for inclusivity and excellence in resource-constrained contexts. *Journal of Service Theory and Practice*, *36 (3),* 357-411. https://doi.org/10.1108/jstp-01-2025-0017

Lim, K., Zheng, P., & Chen, C.-H. (2020). A state-of-the-art survey of digital twin: techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing, 31(6)*, 1313–1337. https://doi.org/10.1007/s10845-019-01512-w

Lusch, R, & Nambisan, S. (2015). Service Innovation: A Service-Dominant Logic Perspective. *MIS Qarterly, 39 (1)*, 155-175. https://doi.org/10.25300/misq/2015/39.1.07

Madeira, R. N., Santos, P. A., Java, O., Priebe, T., Graça, E., Sárközi, E., Asprion, B., & Gómez, R. P.-B. (2022). Towards Digital Twins for Multi-Sensor Land and Plant Monitoring. *Procedia Computer Science, 210*, 45–52. https://doi.org/10.1016/j.procs.2022.10.118

Martinelli, M., Pietri, M., Rossini, E., Picone, M., & Mamei, M. (2024). Digital Twin Driven Collaboration in Industry 5.0. In: *2024 32nd International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE),* 104–109. https://doi.org/10.1109/WETICE64632.2024.00027

Meierhofer, J., Schweiger, L., Lu, J., Züst, S., West, S., Stoll, O., & Kiritsis, D. (2021). Digital Twin-Enabled Decision Support Services in Industrial Ecosystems. *Applied Sciences, 11(23),* 11418. https://doi.org/10.3390/app112311418

Meierhofer, J., West, S., Rapaccini, M., & Barbieri, C. (2020). The Digital Twin as a Service Enabler: From the Service Ecosystem to the Simulation Model. In: Nóvoa, H., Drăgoicea, M. , & Kühl, N. (Eds.), *Lecture Notes in Business Information Processing. Exploring Service Science, 377*, 347–359. Springer International Publishing. https://doi.org/10.1007/978-3-030-38724-2_25

Mele, C., & Russo-Spena, T. (2024). Agencement of onlife and phygital: smart tech–enabled value co-creation practices. *Journal of Service Management, 36(2),* 217-240. https://doi.org/10.1108/josm-03-2023-0113

Meske, C., Osmundsen, K., & Junglas, I. (2021). Designing and Implementing Digital Twins in the Energy Grid Sector. *MIS Quarterly Executive, 20(3).*

Nguyen, D., Dao, M., & Zettsu, K. (2024). Digital Twin Orchestration: Framework and Smart City Applications. Proceedings of the CEUR Workshop Proceedings 2022.

Orlikowski, W. J. (2007). Sociomaterial Practices: Exploring Technology at Work. *Organization Studies, 28(9)*, 1435–1448. https://doi.org/10.1177/0170840607081138

Paré, G., Trudel, M.-C. , Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management 52 (2),* 183–199. https://doi.org/10.1016/j.im.2014.08.008

Payne, E., Dahl, A. J., & Peltier, J. (2021a). Digital servitization value co-creation framework for AI services: a research agenda for digital transformation in financial service ecosystems. *Journal of Research in Interactive Marketing, 15(2),* 200-222. https://doi.org/10.1108/JRIM-12-2020-0252

Payne, E., Peltier, J., & Barger, A. (2021b). Enhancing the value co-creation process: artificial intelligence and mobile banking service platforms. *Journal of Research in Interactive Marketing, 15(1),* 68-85. https://doi.org/10.1108/JRIM-10-2020-0214

Ricci, A., Croatti, A., & Montagna, S. (2021). Pervasive and Connected Digital Twins - A Vision for Digital Health. *IEEE Internet Computing, 26*, 26-32. https://doi.org/10.1109/MIC.2021.3052039

Schmidt, R., Alt, R., & Zimmermann, A. (2026). Value co-creation on networked platforms - A service-dominant logic analysis of the Alexa Assistant ecosystem. *Electronic Markets, 36(1),* 1. https://doi.org/10.1007/s12525-025-00812-9

Schultheiß, A., Polovoj, E., Dolanovic, S., & Gutsche, K. (2023). Digital Service Twin - Design Criteria, Requirements and Scope for Service Management. In: Alfnes, E., Romsdal, A., Strandhagen, J.O., von Cieminski, G., & Romero, D. (Eds.), *IFIP Advances in Information and Communication Technology. Advances in Production Management Systems. Production Management Systems for Responsible Manufacturing, Service, and Logistics Futures, 690*, 302-316. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43666-6_21

Stoll, O., West, S., Gaiardelli, P., Harrison, D., & Corcoran, F. J. (2020). The Successful Commercialization of a Digital Twin in an Industrial Product Service System. In: Lalic, B., Majstorovic, J., Marjanovic, U., von Cieminski, G., & Romero, D. (Eds.), *IFIP Advances in Information and Communication Technology. Advances in Production Management Sytems. Towards Smart and Digital Manufacturing,* 592, 275-282). Springer International Publishing. https://doi.org/10.1007/978-3-030-57997-5_32

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, *94(9-12),* 3563–3576. https://doi.org/10.1007/s00170-017-0233-1

Tuegel, E. J., Ingraffea, A. R., Eason, T. G., & Spottswood, S. M. (2011). Reengineering Aircraft Structural Life Prediction Using a Digital Twin. *International Journal of Aerospace Engineering, 2011(1)*, 1–14. https://doi.org/10.1155/2011/154798

van der Valk, H., Haße, H., Möller, F., Arbter, M., Henning, J.-L., & Otto, B. (2020). A Taxonomy of Digital Twins. Proceedings of the American Conference on Information Systems (ACMIS 2020).

Vargo, S. L., & Lusch, R. (2016). Institutions and axioms: An extension and update of service-dominant logic. *Journal of the Academy of Marketing Science, 44(1),* 5–23. https://doi.org/10.1007/s11747-015-0456-3

Verdugo-Cedeño, M., Jaiswal, S., Ojanen, V., Hannola, L., & Mikkola, A. (2024). Simulation-Based Digital Twins Enabling Smart Services for Machine Operations: An Industry 5.0 Approach. *International Journal of Human–Computer Interaction*, 40(20), 6327–6343. https://doi.org/10.1080/10447318.2023.2254607

Wache, H. (2023). It's Time to Smarten Up! - A Framework for Building Smart Service Systems. Proceedings of the American Conference on Information Systems 2023 (ACMIS 2023).

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26(2)*, xiii–xxiii.

Zarco, C. G. (2024). Digital Twin and Services: A Theoretical Development Framework. 2024 IEEE Technology and Engineering Management Society (TEMSCON LATAM), 1-6. https://doi.org/10.1109/TEMSCONLATAM61834.2024.10717667

Zhu, L., Liu, S., Wang, Q., & Zhang, J. Z. (2026). Value Co-Creation and Resource Integration in Supply Chain Innovation: An Evolutionary Game Perspective. *Sage Open, 16(1).* https://doi.org/10.1177/21582440261423746

HELPFUL ENTERPRISE ARCHITECTURE ELEMENTS FOR ORGANIZATIONAL NETWORKS

ARJEN MARIS,[1] PASCAL RAVESTEIJN,[2]
STIJN HOPPENBROUWERS,[3] JOS VAN HILLEGERSBERG[1]

[1] University of Twente, Enschede, the Netherlands
a.maris@utwente.nl, j.vanhillegersberg@utwente.nl
[2] HU University of applied sciences, Utrecht, the Netherlands
pascal.ravesteijn@hu.nl
[3] HAN University of applied sciences, Arnhem, the Netherlands
stijn.hoppenbrouwers@han.nl

Healthcare organizations increasingly collaborate within networks of private and public partners. To do so, data is recorded, stored and exchanged with help of a complex landscape of information systems (IS). While IS landscapes could boost digital innovation, they also generate significant administrative burdens. Enterprise Architecture (EA) supports organizations in managing IS landscapes. However, current EA practices do not fit well on the dynamics of networks, and research on EA for networks is scarce. Therefor the research objective is to determine what EA elements help to manage an IS landscape within healthcare organizations that collaborate in a network structure. Three group sessions were conducted using a collaboration engineering approach to refine network characteristics and usefull EA elements. The findings indicate that 'network motive' and 'network governance mechanisms' are the most important network characteristics, while 'EA goals' and 'EA processes' are the most helpful for network architecture (EA for a network).

DOI
https://doi.org/10.18690/um.fov.4.2026.56

ISBN
978-961-299-152-4

Keywords:
enterprise architecture, network governance, collaboration engineering, group support system, healthcare

University of Maribor Press

1 Introduction

Globally, life expectancies are rising (OECD, 2025b). As a result, the number of people with two or more chronic illnesses (multimorbidity) increases (Palladino et al., 2016). To assist persons with multimorbidity, healthcare providers must work with a wide range of professions to keep providing the best possible care for the best economic value, and they need to collaborate as efficiently as possible (OECD, 2025a).

Healthcare organizations collaborate in a variety of networks, including service networks, supplier networks, and integrated healthcare networks (Gremyr et al., 2025). For this study, a network consists "of three or more legally autonomous organizations that work together to achieve not only their own goals but also a collective goal" (Provan & Kenis, 2008). In case of healthcare, members of such networks trying to collaborate to provide the best possible care might include healthcare insurers, regional and national authorities, health quality inspectors, other healthcare providers, and of course the patient and their relatives (Maris, 2022).

To collaborate, communicate, and be accountable, healthcare organizations record, store, and exchange data using a complex landscape of information systems (IS) (Maris, 2022). While the use of digital technologies is gaining traction (Sienkiewicz-Małyjurek & Zyzak, 2025) and IS landscapes can stimulate digital innovation, such as Artificial Intelligence efforts (Zahlan et al., 2023), their complexity also creates significant administrative burdens (Michel, 2017). Such as multiple registrations, time-consuming apps, continuous authentication, and standardized checklists instead of patient-centred registrations.

Enterprise Architecture (EA) helps with managing IS landscapes. Although EA has been beneficial for a diversity of enterprises, current EA products and EA processes fail to fit well on the dynamics of Network Governance (NG) contexts (Lapalme et al., 2016). In addition, research on (the usability of) EA within NG contexts is scare (Maris et al., 2023; Maris, Hoppenbrouwers, & Van Hillegersberg, 2023). Therefore, our research question focused on: *What elements of enterprise architecture help to manage an information system landscape within healthcare organizations operating within a network governance structure?*

The next section of this paper describes the research's core concepts, 'Enterprise architecture elements' and 'network governance', followed by a description of the research method in Section 3. The study's initial findings are reported in the results section. Conclusions, recommendations for future research, and limitations are discussed in the last section.

2 Theoretical perspective

2.1 Enterprise architecture elements

Originally, EA was intended to visualize the current and future state of the whole enterprise information system landscape (Kotusev, 2016). These visualizations resulted in descriptions of how to change the information system landscape from a current state to a desired state (a roadmap). Subsequently EA processes, such as change management, also became part of EA (Spewak & Tiemann, 2006) and EA research (Kotusev, 2017). Bernard (2012, p. 31) states that EA has become *"a strategy and business-driven activity that supports management planning and decision-making by providing coordinated views of an entire enterprise."*

Although many different EA frameworks and models have been used in the public sector (Ansyori et al., 2018), and specifically in healthcare (da L. Júnior et al., 2021), standard EA frameworks and models are not always successful in organizations. In most cases only a couple of elements of the EA frameworks and models that were applied fit the context instead of the whole framework or model (Kotusev, 2018). In addition, several issues had to be taken into account when applying EA for network organizations (Lapalme et al., 2016) and especially in the healthcare sector, where executives and other stakeholders have limited awareness of the existence and value of EA (Ajer & Olsen, 2018).

Maris, Hoppenbrouwers, and Van Hillegersberg (2023), partially based on Gebre-Mariam and Bygstad's (2019) work, distinguish five EA elements: EA products, EA goals, EA contractual processes, EA relational processes, and EA competencies (Figure 1). A variety of EA products, including the roadmap to the desired state, define and illustrate the information system landscape's existing and desired states. With these products, the organization hopes to achieve one or more EA goals. Most stated EA goals fit into one of two basic categories: the first focuses on business and

information technology alignment; the second focuses on the extent of flexibility and agility, which enables information management and the organization to adapt.

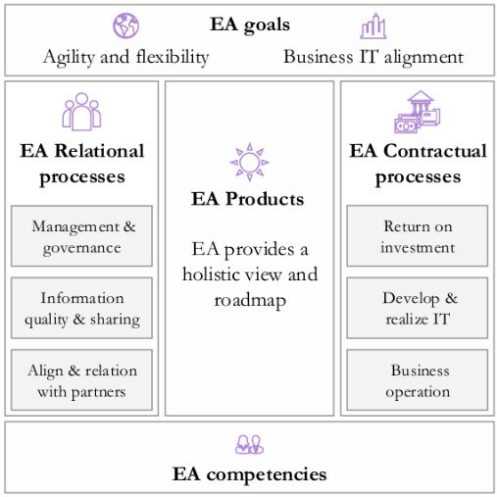

Figure 1: EA elements
Source: Authors, based on Maris, Hoppenbrouwers and Van Hillegersberg (2023)

EA contractual processes refer to information management practices, such as development, realisation, and maintenance of information technology (business operation), and its financial aspects. EA relational processes, on the other hand, have a foundation in human interaction. The emphasis is on information governance, including who is in charge of what and when, what information must be elicited and shared, what level of quality is required, and how to further align with partners and relationships. These elements can only be realized if the organization possesses the right EA competencies, which might include effective communication, considering systems and environments, and managing large groups (Lapalme, 2012).

2.2 Network governance

A network could be an institutional arrangement, a social system of collaborating individuals, a physical system of logistic partners, or more technical oriented like information and communication technology (Sienkiewicz-Małyjurek & Zyzak, 2025). In this study a network represents an institutional collaboration of three or more legally autonomous organizations (Provan & Kenis, 2008).

To effectively govern a network, Provan and Kenis (2008) introduced three network governance structures: shared governance; a network administrative organization; and lead governance. Hastings et al. (2014) distinguished six healthcare governance mechanisms based on 113 studies: (1) shared governance; (2) Magnet accreditation (Magnet® status by the American Nurses Credentialing Center); (3) professional development initiatives; (4) quality-focused initiatives; (5) reorganization of health services (e.g., moves to team-based care, private vs. public organizations); and (6) funding models. Ferlie & McGivern (2010) discuss a range of established network governance mechanisms, including the degree of network regulation (Perri et al., 2006). Maris et al. (2023) describe network characteristics based on 80 studies. They combine governance structure (lead- to shared governance) and governance regulation style (relational, trust based versus contractual based governance), to form four network governance mechanisms. Inspired by Perry et al. (2006) those mechanisms were named: Managed care network (MCN); Coordinated network (CN); Learning and informational network (LIN); and Procurement network (PN). Adding to this Maris et al. (2023) proposed five network characteristics: motive, dimensions, perspective, history, and capabilities (Figure 2).

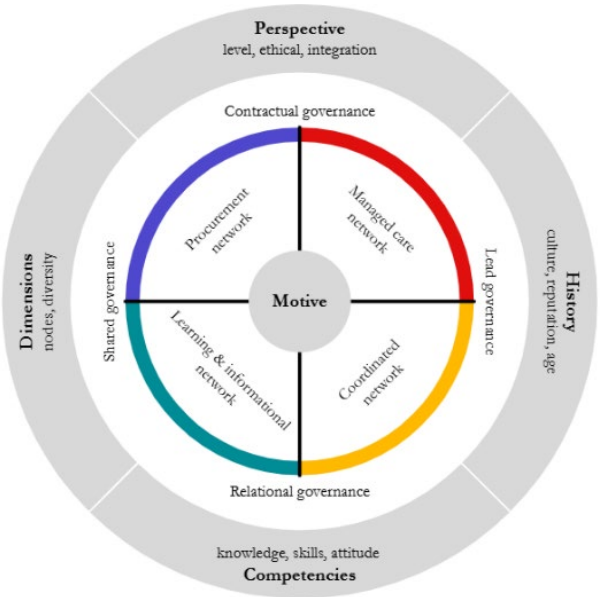

Figure 2: Network characteristics
Source: Authors, based on Maris et al. (2023) and Perri et al. (2006)

Motive is central in Figure 2. All parties could strive towards a shared network motive, but also have their own motives to join the network. The four network governance mechanisms are placed directly around the motives. The other network characteristics are situated as the third layer of in Figure 2. *Dimensions* relates to the number of parties acting in the network, where diversity among the parties plays a role. *Perspectives* considers diverse vantage points for engaging the network, such as personal, organizational, and societal network perspective, or ethical perspectives (to what extent do the motives of the network participants diverge/is there overlap). *History* represents the age of a network, since the effectiveness of a network should increase the longer it has existed, and a network's reputation (external image) and culture (internal image) should grow as it matures. *Competencies* pertain to the knowledge, skills and attitudes that network participants need to organize the network and act within it.

3 Method

Social science researchers often use qualitative research techniques such as individual interviews, participant observation, and engaging groups in various ways (Plummer-D'Amato, 2008). To determine what EA elements could help to manage an IS landscape in healthcare organizations operating within a NG structure, the input and perspectives of EA practitioners with experience in network environments is seen as a necessity. By involving multiple respondents simultaneously and discuss with them literature' findings, the respondents can enrich the findings in a way that might not be possible in one-on-one interviews (Plummer-D'Amato, 2008). Frey and Fontana (1991) have distinguished four well known group session types in social science: (1) focus groups; (2) brainstorming; (3) nominal and Delphi groups; and (4) natural / formal field interview. To support these group session types, group (decision) support systems can help to be more effective and efficient than pen and paper techniques (Briggs et al., 2001; Fjermestad & Hiltz, 2000; Jaclyn Lee, 2015; van Hillegersberg & Koenen, 2016).

3.1 Group Support Systems

For this study, three group sessions were conducted using a collaborative engineering technique. Since 2001 a combination of group sessions with the support of group support systems (GSS) is known as collaboration engineering (Briggs et al.,

2003; de Vreede & Briggs, 2019). According to Briggs et al. (2001), there are three steps that must be considered in order to ensure the quality of a group session that is supported by a GSS: (1) Tool selection, (2) Configuration, and (3) Establishing the session's protocol.

There are many digital systems that can support group sessions, but only a few are used for research aims (Fjermestad & Hiltz, 2000). MeetingWizard and Spilter are two systems mentioned as GSS in more recent studies, for example Namagembe et al. (2023) and Veen and Heuts (2024). Both systems were applicable for this study, however Spilter (https://spilter.nl) came out as the best option based on continuity, functionality and user interface.

3.2 Configuration

The second step was to configure Spilter to create a pattern of interaction. The group session was set up in three rounds in order to develop common knowledge about, and verify, the theoretical findings with the participants. One concentrated on the EA elements identified in the literature. Another concentrated on the network characteristics. The third round explored possible connections between network characteristics and EA elements. The respondents were able to view their own and each other's responses at each round. The notion was that if someone was having trouble providing an example or response, that this insight could provide inspiration.

The first two rounds were constructed in three phases. To gather all potentially relevant ideas and examples about the network characteristics and EA elements, the first phase involved brainstorming with pre-designed rubrics in line with the literature's findings (synthesized in Figures 1 and 2). Additionally, a 'no category' rubric was provided to allow the respondents to make improvements on rubric level. Findings from the first phase were to be discussed and categorized with all respondents during this second phase. In the last phase, the respondents were asked to score the second phase's results (enriched and new categories) according to how the participants perceived the importance for network architecture on a 4-points Likert scale: (1) not important; (2) could be of use; (3) important; and (4) essential. In addition, (-) no opinion was provided.

The results of the first and second rounds were imported into a multiple-choice assessment, to answer this question: "Which network characteristics are connected to which EA elements?" The respondents were asked to check the boxes that, in their opinion, represented the links among the network characteristics and the EA elements.

3.3 Session's protocol

To set up the session's protocol, three conditions must be satisfied to conduct a GSS session successfully (Briggs et al., 2003): (A) conceptual reduction; (B) facilitation skills; and (C) building blocks.

3.1.1 Conceptual reduction

When it comes to keeping the protocol and tool's usability as simple as possible, the respondents have been guided through the process of using the Spilter tool. They were informed about the goal, and set-up of the session, the use of the Spilter tool during the session from the time they were invited until the final results were shared, as well as the core concepts of the rounds.

3.1.2 Facilitation skills

It is hard to simultaneously combine GSS experience and presenter skills (Briggs et al., 2003). Therefore, two people facilitated the meetings to handle any potential technical issues and maintain a clean procedure. The researcher only introduced the theoretical notions to minimize the potential for confirmation bias. The other individual, who did not conduct the research and began with the same conceptual knowledge as the respondents, facilitated the discussions and categorizations.

3.1.3 Building blocks

The Spilter tool works with a library of reusable items for idea generation, clustering, scoring, appointing and rigorous research. With help of a Spilter specialist a selection of those items was used to set up the rounds, guaranteeing the use of standard items and minimizing the time expended for the session.

The networks of danw.nl, knvi.nl, and LinkedIn were used to select and get in touch with the participants. The selection criteria for participants required experience as an enterprise architect or information manager in a network setting, preferably in the healthcare sector. A test session with four volunteers was conducted in January 2026 to perfect the setup's validity. Adjustments were made based on the responders' feedback and the experiences of the facilitators. In February, the first group session was conducted with university researchers in the field of EA, network governance and/or healthcare. Two more group sessions were held in March. All spoken content shared in these sessions was recorded using the Spilter tool and a notebook. A report was created following each session and returned to respondents for verification. The results presented in this article are based on these reports.

4 Results

A total of 17 respondents from 14 organizations across the Netherlands participated in the three group sessions (Table 1). The first group (#1. University researchers) included a total of four researchers. The second (#2 Practitioners) and third (#3 Practitioners) groups consisted of six and seven EA practitioners with organizational network experience, respectively. All respondents had knowledge of or experience with the healthcare sector.

Table 1: Respondents

Group:	# respondents	Average work experience
#1. University researchers	4	6 years (0 - 10 years)
#2. Practitioners	6	11 years (2 - 25 years)
#3. Practitioners	7	7 years (3 - 12 years)

Knowledge of EA for networks	Insufficient	Reasonable	Sufficient	Expert
#1. University researchers	1	1	2	0
#2. Practitioners	0	0	4	2
#3. Practitioners	0	2	4	1

Governance mechanism[1]:	MCN	CN	LIN	PN
#1. University researchers	1		3	
#2. Practitioners	3.5		2	0.5
#3. Practitioners	2.5	1	1	1.5

Source: Authors

[1] (MCN) Managed care network, (CN) Coordinated network, (LIN) Learning and informational network, and (PN) Procurement network

The participants indicated how well they understood EA for networks, network characteristics, and EA elements. The first group gave their level of knowledge with the relevant topics a fair to sufficient rating. Some respondents of the second and third group indicated that they were experts. All respondents were requested to think of a network based on their personal experiences to collect relevant thoughts and examples about the network characteristics and EA elements. The sector to which this network had to belong was not restricted in advance. Based on the respondent's descriptions, the networks were classified as one of the four network governance mechanisms of Maris et al. (2023). All participants were able to use the six network characteristics and five EA elements to define their networks and corresponding EA, no new EA element was proposed in any of the sessions. An extra network characteristic was only added by the third group.

Table 2: Value of network characteristics and EA elements for network architecture[2]

Network characteristics	Motive	NG mechanism	Funding streams	Competencies	Perspective	History	Dimension
#1. University researchers	3,75	3,25	-	2,75	3,00	2,75	2,00
#2. Practitioners	3,83	3,67	-	3,00	2,33	2,83	2,33
#3. Practitioners	4,00	3,57	3,29	3,14	3,00	2,43	3,14
Total average	**3,88**	**3,53**	3,29	3,00	2,76	2,65	2,59
Standard deviation	0,33	0,51	0,76	0,87	0,97	0,79	0,94

EA elements	EA goals	EA relational processes	EA competencies	EA products	EA contractual processes
#1. University researchers	4,00	3,00	2,50	2,75	3,00
#2. Practitioners	3,33	3,67	3,33	3,00	2,50
#3. Practitioners	3,71	3,29	3,43	3,57	2,71
Total average	**3,65**	**3,35**	3,18	3,18	2,71
Standard deviation	0,61	0,79	0,88	0,88	0,69

Source: Authors

[2] Scores are averages on a scale of 1 to 4: (1) not important, (2) could be of use, (3) important, and (4) essential.

At the end of the first and second rounds, participants were asked to each score how they perceived the importance of network characteristics and EA elements for network architecture (Table 2).

Their scores collectively rated network motive and NG mechanisms as the most important network characteristics for network architecture, while EA goals and EA relational processes were rated as the most important EA elements. By the end of the second round the group of respondents had developed common knowledge about the EA elements and network characteristics. With that, they were able to identify whether or not there might be a link between the two (Table 3).

Table 3: Links between Network characteristics and EA elements[3]

	NG mechanism	Motive	Perspective	Competencies	Dimension	History	Funding streams
EA relational processes	**82%**	59%	53%	41%	53%	71%	29%
EA goals	59%	**88%**	71%	41%	35%	35%	29%
EA competencies	47%	47%	41%	**88%**	47%	29%	0%
EA contractual processes	**76%**	41%	35%	29%	41%	29%	86%
EA products	41%	41%	41%	41%	35%	41%	43%

Source: Authors

The links between NG mechanisms and EA processes, motive and EA goals, and network competencies and EA competencies are mentioned the most by the respondents. It became evident which EA elements could be of use in a network environment by combining the value of network characteristics and EA elements with the link score (Table 4). The scores are calculated by multiplying the link score by the average of the network characteristics score and the EA elements score.

The findings demonstrate that the NG mechanisms are best linked with EA processes (relational processes, 2,83; contractual processes, 2,38). Furthermore, the strongest link is between network motives and EA goals, with a score of 3,32 out of 4, and EA competencies link the most with network competencies with a score of

[3] Scores represent the percentage of respondents that identified the combination.

2,72 out of 4. Additionally, there seems to be a strong link between financial flows and EA contractual processes, however, only the third group scored this additional network characteristic. As a result, the reported funding streams' numbers are based on the number of respondents in that group.

Table 4: Combination of value and links

	NG mechanism	Motive	Competencies	Perspective	Dimension	History	Funding streams
EA relational processes	**2,83**	2,13	1,31	1,62	1,57	2,12	0,95
EA goals	2,11	**3,32**	1,37	2,26	1,10	1,11	0,99
EA competencies	1,58	1,66	**2,72**	1,22	1,36	0,86	0,00
EA products	1,38	**1,45**	1,27	1,22	1,02	1,20	1,38
EA contractual processes	**2,38**	1,36	0,84	0,97	1,09	0,79	2,57

Source: Authors

5 Conclusion

The research aim was to determine what EA elements help to manage an IS landscape within healthcare organizations operating within a network governance structure. To refine the characteristics of an organizational network and to link them to usefull EA elements, a group session based on a collaboration engineering approach was designed and tested. In addition, three sessions were conducted.

The findings indicate that network governance mechanisms and network motive are the most important network characteristics for setting up EA within networks, while EA goals and EA processes, particularly relational processes, are the most helpful in a network context. These empirically grounded findings suggest that healthcare managers and enterprise architects should focus on how EA processes within networks are organized in line with the network motive and EA goals, instead of designing and building EA products.

These results help to consider why EA products and EA processes might not fit well with the dynamics of NG contexts, even though they do not provide clarity on potential differences in the emphasis on EA elements per governance

mechanism. Because a network's governance mechanism will likely evolve over time, clarity on this subject could help to position EA as an adaptive governance instrument, rather than a static design method for improving inter-organizational collaboration, which should enhance digital innovation capabilities across public healthcare networks. More generally, this promises to be key aspect of public management in the future. Therefore, we recommend further research on the EA elements primarily associated with each governance mechanism.

Acknowledgements

We would like to acknowledge J. van Staveren (Spilter), S. Wagemaker (Spilter), G. Ongena and E. Lindquist for their commitment and contribution to this research project.

References

Ajer, A. K., & Olsen, D. (2018). *Enterprise architecture challenges: A case study of three Norwegian public sectors*. *26*, 17.

Ansyori, R., Qodarsih, N., & Soewito, B. (2018). A systematic literature review: Critical success factors to implement enterprise architecture. *Procedia Computer Science, The 3rd International Conference on Computer Science and Computational Intelligence (ICCSCI 2018) : Empowering Smart Technology in Digital Era for a Better Life*, *135*, 43–51. https://doi.org/10.1016/j.procs.2018.08.148

Bernard, S. A. (2012). *An Introduction to Enterprise Architecture: Third Edition*. AuthorHouse.

Briggs, R. O., De Vreede, G.-J., Nunamaker, J. F., & Tobey, D. (2001). ThinkLets: Achieving predictable, repeatable patterns of group interaction with group support systems (GSS). *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, 9. https://doi.org/10.1109/HICSS.2001.926238

Briggs, R. O., Vreede, G.-J. D., & Nunamaker Jr., J. F. (2003). Collaboration Engineering with ThinkLets to Pursue Sustained Success with Group Support Systems. *Journal of Management Information Systems*, *19*(4), 31–64. https://doi.org/10.1080/07421222.2003.11045743

da L. Júnior, S. H., Silva, F. Í. C., Albuquerque, G. S. G., de Medeiros, F. P. A., & Lira, H. B. (2021). Frameworks, methodologies and specification tools for the enterprise architecture application in healthcare systems: A systematic literature review. *International Conference on E-Health Networking, Application Services*, 1–7. https://doi.org/10.1109/HEALTHCOM49281.2021.9398916

de Vreede, G.-J., & Briggs, R. O. (2019). A Program of Collaboration Engineering Research and Practice: Contributions, Insights, and Future Directions. *Journal of Management Information Systems*, *36*(1), 74–119. https://doi.org/10.1080/07421222.2018.1550552

Ferlie, E., & McGivern, G. (2010). Networks in health care: A comparative study of their management, impact and performance. *Academia.Edu*. https://www.academia.edu/download/31109588/ SDO_FR_08-1518-102_V01.pdf

Fjermestad, J., & Hiltz, S. R. (2000). Group Support Systems: A Descriptive Evaluation of Case and Field Studies. *Journal of Management Information Systems*, *17*(3), 115–159. https://doi.org/10.1080/07421222.2000.11045657

Frey, J. H., & Fontana, A. (1991). The group interview in social research. *The Social Science Journal*, *28*(2), 175–187. https://doi.org/10.1016/0362-3319(91)90003-M

Gebre-Mariam, M., & Bygstad, B. (2019). What can enterprise architecture do for healthcare? A framework of antecedents and benefits. *Electronic Government*, *15*(2), 213–239. https://doi.org/10.1504/EG.2019.098672

Gremyr, I., Colldén, C., Hjalmarsson, Y., Schirone, M., & Hellström, A. (2025). Networks for healthcare delivery: A systematic literature review. *Journal of Health Organization and Management*, *39*(9), 36–53. https://doi.org/10.1108/JHOM-09-2023-0262

Hastings, S. E., Armitage, G. D., Mallinson, S., Jackson, K., & Suter, E. (2014). Exploring the relationship between governance mechanisms in healthcare and health workforce outcomes: A systematic review. *BMC Health Services Research*, *14*(1), 479. https://doi.org/10.1186/1472-6963-14-479

Jaclyn Lee, M. S. (2015). *A computer-based group discussion support tool for achieving consensus and culture change using the organisational culture assessment instrument (OCAI): An action design research study*. https://doi.org/10.3990/1.9789810978068

Kotusev, S. (2016). The History of Enterprise Architecture: An Evidence-Based Review. *Journal of Enterprise Architecture*, *12*(1), 29–37.

Kotusev, S. (2017). Enterprise architecture: What did we study? *International Journal of Cooperative Information Systems*, *26*(04), 1730002. https://doi.org/10.1142/S0218843017300029

Kotusev, S. (2018). TOGAF-based Enterprise Architecture Practice: An Exploratory Case Study. *Communications of the Association for Information Systems*, 321–359. https://doi.org/10.17705/1CAIS.04320

Lapalme, J. (2012). Three schools of thought on enterprise architecture. *IT Professional*, *14*(6), 37–43. https://doi.org/10.1109/MITP.2011.109

Lapalme, J., Gerber, A., Van der Merwe, A., Zachman, J., Vries, M. D., & Hinkelmann, K. (2016). Exploring the future of enterprise architecture: A Zachman perspective. *Computers in Industry, Special Issue on Future Perspectives On Next Generation Enterprise Information Systems*, *79*, 103–113. https://doi.org/10.1016/j.compind.2015.06.010

Maris, A. (2022). Enterprise architecture for network organizations: A research design to investigate what elements of EA help to reduce administrative burdens in healthcare. *Proceedings of the 35th Bled eConference*, 717–731. https://irf.fhnw.ch/server/api/core/bitstreams/ede56e55-0223-416f-bc2c-081cf1b23fe6/content#page=731

Maris, A., Hoppenbrouwers, S., Ravesteijn, P., & Van Hillegersberg, J. (2023). Network governance in healthcare: A systematic literature review. *Proceedings of the 19th European Conference on Management Leadership and Governance*, *19*, 249–257. https://doi.org/10.34190/ecmlg.19.1.1798

Maris, A., Hoppenbrouwers, S., & Van Hillegersberg, J. (2023). Enterprise architecture in healthcare networks: A systematic literature review. *Communications of the IIMA*, *21*(1). https://doi.org/10.58729/1941-6687.1446

Michel, L. (2017). *Into the black box of a nursing burden, a comparative analysis of administrative tasks in French and American hospitals*. https://tel.archives-ouvertes.fr/tel-01762491/document

Namagembe, F., Nakakawa, A., Tulinayo, F., Proper, E., & Overbeek, S. (2023). Towards an e-government enterprise architecture framework for developing economies. *Complex Systems Informatics and Modeling Quarterly*, *2023*(35), 30–66. https://doi.org/10.7250/csimq.2023-35.02

OECD. (2025a). *Does Healthcare Deliver?: Results from the Patient-Reported Indicator Surveys (PaRIS)*. OECD Publishing. https://doi.org/10.1787/c8af05a5-en

OECD. (2025b, September). *Life expectancy*. With Minor Processing by Our World in Data. "Life Expectancy at Birth Period Tables" [Dataset]. OECD, "Life Expectancy at Birth" [Original Data]. https://archive.ourworldindata.org/20250909-093708/grapher/life-expectancy-at-birth-oecd.html

Palladino, R., Tayu Lee, J., Ashworth, M., Triassi, M., & Millett, C. (2016). Associations between multimorbidity, healthcare utilisation and health status: Evidence from 16 European countries. *Age and Ageing*, *45*(3), 431–435. https://doi.org/10.1093/ageing/afw044

Perri, P., Goodwin, N., Peck, E., & Freeman, T. (2006). *Managing Networks of Twenty-First Century Organisations*. Springer. https://doi.org/10.1057/9780230286115

Plummer-D'Amato, P. (2008). Focus group methodology part 1: Considerations for design. *International Journal of Therapy & Rehabilitation*, *15*(2), 69–73. https://doi.org/10.12968/ijtr.2008.15. 2.28189

Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory*, *18*(2), 229–252. https://doi.org/10.1093/jopart/mum015

Sienkiewicz-Małyjurek, K., & Zyzak, B. (2025). Digital Technologies in Public Administration Networks: Systematic Literature Review and Research Avenues. *Administration & Society*, *57*(8–9), 1051–1087. https://doi.org/10.1177/00953997251369090

Spewak, S., & Tiemann, M. (2006). Updating the Enterprise Architecture Planning Model. *Journal of Enterprise Architecture*, 11–19.

van Hillegersberg, J., & Koenen, S. (2016). Adoption of web-based group decision support systems: Experiences from the field and future developments. *International Journal of Information Systems and Project Management*, *4*(1), 49–64.

Veen, H. C. J. van der, & Heuts, L. F. (2024). National Risk Assessment on Money Laundering for the Netherlands 2023. *Den Haag*. https://repository.wodc.nl/handle/20.500.12832/3403

Zahlan, A., Ranjan, R. P., & Hayes, D. (2023). Artificial intelligence innovation in healthcare: Literature review, exploratory analysis, and future research. *Technology in Society*, *74*, 102321. https://doi.org/10.1016/j.techsoc.2023.102321

A Framework for Evaluating the Sustainability of IoT Smart City Projects

Hans-Dieter Zimmermann

OST Eastern Switzerland University of Applied Sciences, School of Management, St.Gallen, Switzerland
hansdieter.zimmermann@ost.ch

Although Internet of Things (IoT) technologies are essential for developing sustainable smart cities, systematically evaluating their multidimensional sustainability benefits is a significant challenge. synthesizes existing approaches and proposes a comprehensive, integrated evaluation framework. The proposed framework combines life cycle assessment (LCA) for environmental impacts, cost-benefit analysis (CBA) for economic justification, and multi-criteria decision analysis (MCDA) to integrate stakeholder perspectives and social dimensions. Ultimately, this hybrid approach provides city planners with a rigorous, transparent, and practical tool to ensure that IoT deployments deliver absolute sustainability gains rather than merely relative efficiencies.

DOI
https://doi.org/10.18690/um.fov.4.2026.57

ISBN
978-961-299-152-4

Keywords:
IoT,
smart cities,
sustainability assessment,
evaluation framework,
triple bottom line

University of Maribor Press

1 Introduction

1.1 Motivation

Sustainability has emerged as a critical megatrend driven by global challenges, including climate change and resource scarcity. The goal is to balance current and future environmental, social, and economic needs. In this context, the circular economy is increasingly recognized as a vital tool for achieving broader sustainability goals, including the United Nations' Sustainable Development Goals. Digitalization, specifically the application of Internet of Things (IoT) technologies, is a fundamental driver of this transformation and the technological backbone of sustainable smart cities (Ahad et al., 2020; Zimmermann, 2025). Several cities and communities in various countries are implementing IoT-based systems and applications to support different aspects of smart cities, thereby embracing the smart city vision.

However, although IoT technologies are widely deployed, it remains difficult to systematically evaluate whether they genuinely deliver on their promises of multidimensional sustainability (Bibri & Krogstie, 2017). Current evaluations often rely solely on technical metrics, failing to capture complex trade-offs, hidden ecological costs (e.g., e-waste and energy consumption), and neglected social factors across the triple bottom line (Ahvenniemi et al., 2017; Pirson, 2023; Shao et al., 2023). The motivation for developing a specialized evaluation framework stems from several key issues:

Complex Interdependencies and Trade-offs: The environmental, economic, and social dimensions of sustainability are deeply interconnected and often create conflicting objectives (Bibri & Krogstie, 2017; Yigitcanlar et al., 2018). These trade-offs are discussed in detail in Section 3.4.

Hidden Environmental Costs and Rebound Effects: IoT infrastructure itself generates ecological burdens - including energy consumption and e-waste - and rebound effects may offset efficiency gains (Ahad et al., 2020; Pirson, 2023). These issues are examined in Section 3.1.

Neglected Social Dimensions: The social dimension - including digital inclusion, citizen participation, and privacy - remains underrepresented in current evaluations (Ahvenniemi et al., 2017; Berville et al., 2025; Sharma et al., 2023), as discussed in Section 3.3.

This multidisciplinary and transformational perspective is not merely theoretical; the author's involvement in the 'IoT Sustainability Lab' project demonstrates the necessity and complexity of evaluating IoT projects from a practical perspective (Wissenschaftsverbund, n.d.).

To address these gaps, this study proposes an integrated evaluation framework that moves beyond isolated metrics. By combining Life Cycle Assessment (LCA) for environmental impacts, Cost-Benefit Analysis (CBA) for economic justification, and Multi-Criteria Decision Analysis (MCDA) to integrate stakeholder perspectives and social dimensions, the framework rigorously assesses the true ecological, economic, and social impacts of IoT smart city projects (Akande et al., 2019; Ipsen et al., 2019; Yigitcanlar et al., 2018).

However, achieving this rigorous assessment requires more than methodological integration; it necessitates multidisciplinary convergence. While Information Systems (IS) focuses on data flows and business logic, engineering expertise is imperative for assessing hardware reliability, sourcing materials, improving energy efficiency, and managing end-of-life. Absent this engineering perspective, the evaluation remains "hardware-blind," disregarding the material costs and physical risks that determine long-term viability (Baumgartner et al., 2025). Additional disciplines - including cybersecurity, law, ethics, human-computer interaction, and domain-specific knowledge (e.g., energy, buildings, transportation) - must also be integrated. This multidisciplinary requirement reflects the theory of digital transformation (DT), where IoT serves as the physical enabler of strategic transformation (Vial, 2019). DT theory concepts, such as Dynamic Capabilities, Sociotechnical Systems Theory, and Business Model Innovation, provide a strategic lens for evaluating IoT's broader implications.

Related work has addressed aspects of this challenge, but with notable limitations. Cavalieri et al. (2024) present a survey-based socio-ecological assessment tool for manufacturing supply chains, but deliberately omit the economic dimension and rely

on organizational theories rather than systemic life-cycle models. Baltuttis et al. (2022) focus exclusively on the economic value of industrial IoT solutions, distinguishing between direct "frontstage" and indirect "backstage" value along B2B2C value chains. However, they do not address environmental or social impacts. Lanfranchi (2025) examines the intersection of IoT and environmental sustainability but overlooks the broader Triple Bottom Line framework and lacks empirical grounding through field studies. These limitations underscore the need for the integrated, multidimensional approach proposed in this study.

We address the following research question: How can these distinct methodologies be combined into a single, integrated framework that provides methodological rigor while remaining practically applicable for city stakeholders?

The study proceeds as follows: Section 2 provides the theoretical foundations, Section 3 discusses the sustainability dimensions, Section 4 reviews methodological approaches, Section 5 presents the integrated framework, and Section 6 offers a discussion and future research directions.

1.2 Methodology

To provide a comprehensive and interpretive overview of the current state of knowledge regarding the sustainability evaluation of IoT smart city projects, this study employs a critical narrative literature review (Paré et al., 2015; Tate et al., 2015) as a basis for systematically developing an integrated evaluation framework through conceptual framework building. The research process was operationalized into distinct phases to ensure methodological transparency:

Literature synthesis and problem identification: A narrative review approach was deliberately chosen, as it allows for critical synthesis of fragmented findings in the dynamic and emerging field of IoT sustainability. Additionally, the interdisciplinary nature of the topic - spanning technology, ecology, economics, and social sciences - further necessitated this flexible approach to map the landscape of isolated technical metrics and identify core multidimensional gaps.

Requirement elicitation and theoretical grounding: The framework's design decisions were directly informed by gaps identified in the literature. Theoretically, it is grounded in the Triple Bottom Line (TBL) approach (integrating environmental, economic, and social dimensions) and Socio-Technical Systems Theory (acknowledging the interplay between human, organizational, and technological layers). From this foundation, the core requirement was derived: a hybrid evaluation model capable of resolving socio-technical and economic-ecological trade-offs.

Framework construction and integration logic: The theoretical components were operationalized through methodological triangulation. LCA, CBA, and MCDA were conceptually integrated not as parallel tools, but as sequential building blocks, designed to satisfy the requirements elicited in the previous phase.

The literature search followed a broad, exploratory, deep review approach rather than a systematic review methodology. Accordingly, no PRISMA protocol was applied, and no formal inclusion or exclusion criteria were documented. The search was conducted across multiple databases, including SciSpace, Google Scholar, and SciSpace Full Text, to ensure both breadth and depth of coverage. Multiple query combinations were employed, encompassing key terms related to IoT, smart cities, sustainability, and evaluation methods. No year filter was applied, keeping the search open-ended across all available publication years. Additionally, snowballing was applied by tracing references from key publications to identify further relevant sources that may not have surfaced through database searches alone.

A systematic literature review was deliberately avoided because its rigid protocol and quantitative data analysis focus are less suited to a highly dynamic and interdisciplinary research field currently in the conceptual design stage (Grant & Booth, 2009). The methodological flexibility of a narrative approach allows for dynamic refinement of search parameters and theoretical focal points during the iterative research process. This flexible form also enables critical synthesis of fragmented findings (Snyder, 2019), which is essential for "gap spotting" and legitimizing new research questions (Alvesson & Sandberg, 2011). Ultimately, this review serves as a strategic guide to establish a solid theoretical foundation for the proposed integrated framework rather than acting as a purely statistical, retrospective inventory.

2 Background and Theoretical Foundations

2.1 Smart Cities and IoT Technologies

Smart cities use digital technologies and data analysis to optimize urban services and infrastructure. IoT forms the technological backbone of this by equipping physical objects with sensors, actuators, and communication capabilities. These networked systems enable the continuous collection, transmission, and analysis of data on urban processes in real time (Ahad et al., 2020; Bibri & Krogstie, 2017).

The integration of IoT with artificial intelligence (AI) expands the possibilities for predictive analytics and autonomous decision-making in urban systems. Alwar et al. (2024) demonstrate how the combination of IoT and AI can support sustainable smart city development, with the technologies working synergistically to maximize efficiency gains. The technological infrastructure typically comprises several layers: the sensor layer (data collection), the network layer (data transmission), the platform layer (data processing and storage), and the application layer (services and user interaction) (Prasad et al., 2025).

2.2 Sustainability Concepts and the Triple Bottom Line

The concept of sustainability in smart cities is based on the triple bottom line approach, which integrates environmental, economic, and social dimensions (Krúpová et al., 2025). These three pillars are interdependent and must be considered as a whole to enable a comprehensive sustainability assessment. The environmental dimension focuses on minimizing environmental impacts, including greenhouse gas emissions, resource consumption, and waste generation. The economic dimension encompasses cost-efficiency, return on investment (ROI), and long-term economic viability. The social dimension addresses aspects such as quality of life, social justice, inclusion, and citizen participation (Ahvenniemi et al., 2017; Akande et al., 2019). Shao et al. (2023) emphasize that a sustainable smart city framework must give equal consideration to all three dimensions in order to make trade-offs transparent and enable balanced decisions. The theoretical foundation also requires consideration of planetary boundaries and the question of whether IoT solutions lead to absolute or merely relative sustainability improvements - a topic

further explored in the future research agenda (Section 6.1) (Ahad et al., 2020; Pirson, 2023).

3 Sustainability Dimensions in IoT Smart City Projects

The environmental dimension is most comprehensively quantified in the literature, often using life cycle assessments (LCAs) (Baumgartner et al., 2025; Berville et al., 2025; Ipsen et al., 2019). On the one hand, IoT systems offer significant environmental benefits by optimizing resource use and reducing greenhouse gas emissions (Berville et al., 2025; Ipsen et al., 2019). On the other hand, they themselves cause environmental damage through the energy consumption of the infrastructure (data centers, networks) and the generation of electronic waste (Ahad et al., 2020; Pirson, 2023). Systemic analyses demonstrate that, in addition to technical components, peripheral factors (e.g., materials used) and rebound effects must be considered to accurately determine the ecological footprint. (Duygan et al., 2022; Kluczek et al., 2021; Ragab et al., 2023).

From an economic perspective, IoT projects entail high initial investments in hardware and installation, as well as ongoing operating costs, which pose significant barriers to adoption, particularly for municipalities (Bibri & Krogstie, 2017; Yigitcanlar et al., 2018). A justifiable assessment over the entire lifecycle must balance these costs against operational savings - such as those achieved through smart building control or route optimization (Bouramdane, 2023; Duygan et al., 2022). In addition, economic benefits that are difficult to quantify, such as innovation drivers and improved service quality, must be included in the assessment (Shamsuzzoha et al., 2021; Yigitcanlar et al., 2018).

The social dimension has been the least systematically addressed in the literature to date, yet it is crucial for the long-term acceptance of smart city projects (Duygan et al., 2022). Key positive aspects that must be evaluated include social justice, digital inclusion, citizen participation, and improvements in quality of life (Ahvenniemi et al., 2017; Bouramdane, 2023). At the same time, this dimension requires a critical evaluation of social risks, including the digital divide, privacy concerns, and potential surveillance (Akande et al., 2019; Berville et al., 2025; Sharma et al., 2023).

In practice, the three dimensions of sustainability are characterized by complex interdependencies (Bibri & Krogstie, 2017; Yigitcanlar et al., 2018). A highly economically efficient solution can be ecologically detrimental (e.g., due to increased energy consumption), while ecologically beneficial systems often fail due to economic barriers (Bibri & Krogstie, 2017; Yigitcanlar et al., 2018). To make these inherent trade-offs transparent and enable balanced decision-making, integrated assessment approaches are essential (Shao et al., 2023; Kramers et al., 2014).

4 Methodological Approaches to Sustainability Assessment

To operationalize the multidimensional sustainability dimensions discussed in Section 3, an integrated methodological triangulation is essential. The proposed framework relies on three primary pillars, each carefully selected to address a specific sustainability dimension while compensating for the inherent limitations of the others:

LCA for Ecological Precision: LCA is a standardized method (ISO 14040/14044) used to quantify the ecological dimension. In the context of IoT, it evaluates direct and indirect environmental impacts from production through disposal, identifying ecological "hotspots". While highly precise, its limitation lies in its almost exclusive focus on the environment and difficulties in capturing systemic rebound effects.

CBA for Economic Assessment: To address the economic dimension, CBA quantifies high initial hardware investments and long-term operational savings in monetary terms. It provides decision-makers with a clear economic justification through metrics like 'return on investment' (ROI). However, CBA struggles to monetize non-market-based benefits, such as environmental or social effects, and neglects distributional impacts.

MCDA for Social Integration: To capture the underrepresented social dimension and resolve the complex trade-offs between ecology and economy, MCDA methods (e.g., AHP, TOPSIS) are applied. MCDA excels at combining quantitative and qualitative data and systematically incorporating diverse stakeholder perspectives into the evaluation.

The Hybrid Integration Logic: As no single method is sufficient, the proposed framework's hybrid integration logic sequentially combines these three pillars. This triangulation yields a significantly more comprehensive and robust assessment than isolated metrics. The following chapter operationalizes this integration logic into a structured, applicable framework.

5 Framework Development for Sustainability Assessment

In the following, we develop an integrated framework for evaluating the multidimensional sustainability of IoT projects in smart cities. This section outlines how to structure a holistic assessment that combines methodological rigor with practical applicability. To this end, the subsequent four elements are particularly formulated:

1. Conceptual Foundations: This includes the definition of basic principles such as the triple bottom line approach, life cycle orientation, and the systemic perspective.
2. Structural components: The structure of the framework is divided into seven building blocks, ranging from goal setting and hybrid method selection to monitoring.
3. Implementation Steps: A practical guide with six concrete phases for applying the framework to real-world projects.
4. Indicators and Metrics: A detailed set of specific environmental, economic, social, and integrative KPIs for measuring sustainability impacts.

5.1 Conceptual Foundations

A robust evaluation framework for IoT smart city projects requires a multidimensional foundation. It is primarily based on the triple bottom line approach, which treats the environmental, economic, and social dimensions as equally important (Ahvenniemi et al., 2017; Shao et al., 2023; Krúpová et al., 2025). Also essential are a strict life-cycle orientation (from production to disposal) and a systemic perspective that captures direct, indirect, and systemic feedback loops within the overall urban system (Ahad et al., 2020; Berville et al., 2025; Duygan et al., 2022; Ipsen et al., 2019). Furthermore, the framework must incorporate multiple

stakeholder perspectives and be adaptable to the city's specific local context (Akande et al., 2019; Bouramdane, 2023; Ferro-Escobar et al., 2022; Quijano et al., 2022).

5.2 Structural Components

The proposed framework consists of seven key components: (1) Objectives and Scope Definition: Clarification of system boundaries, objectives, and stakeholder interests, (2) Method Selection and Integration: The targeted combination of LCA, CBA, and MCDA (as detailed in Section 4) to ensure a holistic assessment across all dimensions, (3) Data collection and management: Identification of necessary data, establishment of standards, and handling of uncertainties, (4) Indicator system: Definition of key performance indicators (KPIs) and reference values, (5) Evaluation and analysis: Practical implementation of analyses (including sensitivity tests) as well as identification of synergies and trade-offs, (6) Interpretation and communication: Derivation and presentation of recommendations for action tailored to the target audience, (7) Monitoring and adaptive control: Continuous monitoring, establishment of feedback loops, and adaptation of the system.

5.3 Implementation Steps

For practical implementation, the framework is implemented in six consecutive phases: (1) Preparation: Interdisciplinary team formation, stakeholder mapping, and resource planning, (2) Methodology design: Adaptation of evaluation methods, development of data collection tools, and pilot testing, (3) Data collection: Collection of quantitative data (sensors, LCA databases) and qualitative data (interviews, surveys), (4) Analysis and evaluation: Conducting method-specific evaluations (LCA, CBA, MCDA) and synthesizing the final results (Bibri, 2019; Ipsen et al., 2019; Yigitcanlar et al., 2018), (5) Interpretation: Stakeholder workshops to discuss results and prepare reports, (6) Implementation and monitoring: Implementation of recommendations for action and continuous, adaptive learning.

5.4 Indicators and Metrics

The framework uses a robust system of SMART indicators (Specific, Measurable, Achievable, Relevant, Time-bound), which should be developed through a participatory process (Akande et al., 2019; Quijano et al., 2022). These metrics are divided into:

Environmental: e.g., CO_2 eq emissions per functional unit, primary energy demand (kWh), e-waste generation rate (kg/year), recycling rate (%) (Ahad et al., 2020; Berville et al., 2025; Kluczek et al., 2021).

Economic: e.g., total life-cycle costs (CAPEX/OPEX), net present value (NPV), return on investment (ROI), payback period (years) (Bibri & Krogstie, 2017; Bouramdane, 2023; Duygan et al., 2022).

Social: e.g., digital inclusion index, citizen participation rate (%), data protection compliance score, accessibility coverage (%) (Ahvenniemi et al., 2017; Berville et al., 2025; Bouramdane, 2023).

Integrative: e.g., urban resilience index, system scalability score, overall stakeholder satisfaction rating (Bouramdane, 2023).

The selection and weighting of indicators should be carried out in a participatory manner, involving relevant stakeholders, to ensure legitimacy and acceptance (Akande et al., 2019; Bouramdane, 2023). Quijano et al. (2022) emphasize the importance of KPI-driven evaluation frameworks that enable replicability and comparability across different cities.

6 Discussion and Future Research Directions

This section focuses on critically reflecting on the sustainability assessment of IoT projects and translating theoretical insights into practice.

Regarding challenges and limitations, evaluating IoT projects is complicated by limited data availability throughout the lifecycle (Ahad et al., 2020; Ipsen et al., 2019) and the difficulty of defining system boundaries. Specific methodological hurdles

include capturing systemic effects across the lifecycle (Ahad et al., 2020; Pirson, 2023), the monetization of non-market-based benefits for cost-benefit analyses (Yigitcanlar et al., 2018), and the systematic integration of social indicators (Ahvenniemi et al., 2017; Berville et al., 2025; Sharma et al., 2023).

Best practices and success factors: Successful evaluations require early integration into project planning (Quijano et al., 2022) and systematic stakeholder engagement in interdisciplinary teams (Ahvenniemi et al., 2017; Akande et al., 2019; Bouramdane, 2023; Kramers et al., 2014; Ipsen et al., 2019). As detailed in this paper, successful evaluations require methodological pluralism - specifically the hybrid combination of LCA, CBA, and MCDA - to make trade-offs transparent and deliver robust results.

Implications for policy and practice: At the policy level, regulatory frameworks, standardization, and capacity building for municipalities are required (Quijano et al., 2022). In practice, a shift in mindset toward "digital sobriety" (conscious, needs-based use of technology) is required (Berville et al., 2025), as well as the consistent alignment of IoT systems with the principles of the circular economy (Kluczek et al., 2021; Berville et al., 2025; Krúpová et al., 2025). To operationalize this alignment in practice, decision-makers can apply models such as the Digital Function for Circular Economy (DF4CE) framework. This framework explicitly maps how digital functions, such as data collection and analysis, enable specific circular strategies, such as recycling, reuse, and reduction (Zimmermann, 2025).

Based on the identified limitations, the literature outlines the following key research needs:

Social indicators: Development of standardized methods for measuring social impacts such as digital inclusion and data protection (Ahvenniemi et al., 2017; Sharma et al., 2023; Berville et al., 2025). Furthermore, future investigations must address critical socio-technical challenges, such as data privacy, system interoperability, and mechanisms to encourage diverse stakeholders to securely share information across integrated data platforms (Zimmermann, 2025).

Long-term effects and planetary boundaries: Systematic longitudinal studies on rebound effects and the integration of planetary boundaries to measure absolute rather than merely relative sustainability gains (Ahad et al., 2020; Pirson, 2023).

Technological and methodological advancements: Optimization of integrated evaluation frameworks (Kramers et al., 2014; Shao et al., 2023; Ipsen et al., 2019) as well as research into AI and machine learning for automated data collection and predictive modeling (Bibri & Krogstie, 2017; Prasad et al., 2025).

Furthermore, a systematic analysis of existing case studies is essential, as theoretical evaluation approaches vary widely in real-world practice and must be adapted to specific contexts. Existing cases often reveal critical aspects and empirical realities that are overlooked in theoretical models. An example is provided by the evaluation of smart waste management in Södertälje (Brunklaus et al., 2022). This case study demonstrated that the climate impact of the actual IoT solution (sensors and networks) was marginal, while peripheral factors such as trash bags posed the greatest burden. This empirically demonstrates the absolute necessity of a broad, systemic perspective in evaluation.

7 Conclusion

The research makes it clear that the potential of the IoT for smart city development is immense, but a multidimensional, methodologically rigorous evaluation is required. While environmental aspects (via LCA) and economic aspects (via CBA) are often easily quantifiable, the social dimension is crucial for long-term legitimacy. Since no single method is sufficient, hybrid approaches are the most effective way to make trade-offs transparent. Ultimately, the IoT is merely a tool, not an end in itself; its true value depends on needs-based implementation and evidence-based, participatory evaluation.

Acknowledgement

Funding for this research was provided in part by the IoT Sustainability Lab, funded by the Interreg VI Alpenrhein-Bodensee-Hochrhein (ABH) program, which is co-financed by the European Regional Development Fund (ERDF) and the Swiss Confederation.

References

Ahad, M. A., Paiva, S., Tripathi, G., & Feroz, N. (2020). Enabling technologies and sustainable smart cities. Sustainable Cities and Society, 61, 102301. https://doi.org/10.1016/j.scs.2020.102301.

Ahvenniemi, H., Huovila, A., Pinto-Seppä, I., & Airaksinen, M. (2017). What are the differences between sustainable and smart cities? Cities, 60, 234–245. https://doi.org/10.1016/j.cities.2016.09.009.

Akande, A., Cabral, P., Gomes, P., & Casteleyn, S. (2019). The Lisbon ranking for smart sustainable cities in Europe. Sustainable Cities and Society, 44, 475–487. https://doi.org/10.1016/j.scs.2018.10.009.

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. Academy of Management Review, 36(2), 247–271. https://doi.org/10.5465/amr.2009.0188.

Alwar, B., Edgardo, H., & Faustino, E. (2024). Integrating IoT and artificial intelligence for sustainable smart city development: A case study approach. ITEJ (Information Technology Engineering Journals), 9(2). https://doi.org/10.24235/itej.v9i2.137

Baltuttis, D., Häckel, B., Jonas, C. M., Oberländer, A. M., Röglinger, M., & Seyfried, J. (2022). Conceptualizing and Assessing the Value of Internet of Things Solutions. Journal of Business Research, 140, 245–263. https://doi.org/10.1016/J.JBUSRES.2021.10.063

Baumgartner, C., Bohnet, D., Meyer, S., Leber, A., & Ramadane, M. (2025). Assessing the Environmental Impact of IoT Devices - Hotspots and Guidelines for a Better Understanding. 2025 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC), 1–11. https://doi.org/10.1109/ICE/ITMC65658.2025.11106638.

Berville, C., Croitoru, C., & Bode, F. (2025). Life Cycle Analysis in the Context of Smart Cities. E3S Web of Conferences. https://doi.org/10.1051/e3sconf/202560805029.

Bibri, S. E. (2019). On the sustainability of smart and smarter cities in the era of big data: An interdisciplinary and transdisciplinary literature review. Journal of Big Data, 6(1), 1–64. https://doi.org/10.1186/s40537-019-0182-7.

Bibri, S. E., & Krogstie, J. (2017). Smart sustainable cities of the future: An extensive interdisciplinary literature review. Sustainable Cities and Society, 31, 183–212. https://doi.org/10.1016/j.scs.2017.02.016.

Bouramdane, A. A. (2023). Optimal water management strategies: Paving the way for sustainability in smart cities. Smart Cities, 6(5), 2849–2882. https://doi.org/10.3390/smartcities6050128.

Brunklaus, B., Chiew, Y. L., Lundström, A., Nilsson-Lindén, H., & Saarikko, T. (2022). The connected, sustainable, and inclusive society: IoT implementation in a Swedish municipality. E3S Web of Conferences, 349, 11006. https://doi.org/10.1051/e3sconf/202234911006

Cavalieri, A., Reis, J., & Amorim, M. (2024). Socioenvironmental assessment and application process for IOT: A comprehensive approach. *Journal of Cleaner Production*, *436*, 140348. https://doi.org/10.1016/J.JCLEPRO.2023.140348

Duygan, M., Fischer, M., Pärli, R., & Ingold, K. (2022). Where do Smart Cities grow? The spatial and socio-economic configurations of smart city development. Sustainable Cities and Society, 77, 103578. https://doi.org/10.1016/j.scs.2021.103578.

Ferro-Escobar, R., Vacca-González, H., & Gómez-Castillo, H. (2022). Smart and Sustainable Cities in Collaboration with IoT: The Singapore Success Case. Intelligent Systems Reference Library, 121, 213–243. https://doi.org/10.1007/978-3-030-97516-6_12.

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and their associated methodologies. Health Information & Libraries Journal, 26(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x.

Ipsen, K. H., Zimmermann, R. K., Birkved, M., & Ammenberg, J. (2019). Environmental assessment of Smart City Solutions using a coupled urban metabolism - life cycle impact assessment approach. International Journal of Life Cycle Assessment, 24(7), 1239–1253. https://doi.org/10.1007/s11367-018-1453-9.

Kluczek, A., Gladysz, B., & Ejsmont, K. (2021). Application of Lifecycle Measures for an Integrated Method of Environmental Sustainability Assessment of Radio Frequency Identification and

Wireless Sensor Networks. Energies 2021, Vol. 14, 14(10). https://doi.org/10.3390/en14102794.

Kramers, A., Höjer, M., Lövehagen, N., & Wangel, J. (2014). Smart sustainable cities - Exploring ICT solutions for reduced energy use in cities. Environmental Modeling & Software, 56, 52–62. https://doi.org/10.1016/j.envsoft.2013.12.019.

Krúpová, S., Koman, G., Soviar, J., & Holubčík, M. (2025). The Role of Business Models in Smart-City Waste Management: A Framework for Sustainable Decision-Making. Systems 2025, Vol. 13, 13(7). https://doi.org/10.3390/systems13070556.

Lanfranchi, G., Crupi, A., & Cesaroni, F. (2025). Internet of Things (IoT) and Environmental Sustainability: A Literature Review and Recommendations for Future Research. *Corporate Social Responsibility and Environmental Management*, *32*(6), 7648–7670. https://doi.org/10.1002/CSR.70098

Paré, G., Trudel, M. C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. Information & Management, 52(2), 183-199. https://doi.org/10.1016/J.IM.2014.08.008.

Pirson, T. (2023). Environmental perspectives and limits for the Internet of Things in the Anthropocene: Going beyond systematic Life-Cycle Assessment (Doctoral dissertation, UCL - Université Catholique de Louvain). http://hdl.handle.net/2078.1/275373

Prasad, M. V. S., & Subrahmanyam, V. (2025). Intelligent IoT-Enabled Framework for Sustainable Smart Cities. International Journal of Science and Research (IJSR), 406–409. https://doi.org/10.21275/sr25910183656.

Quijano, A., Hernández, J. L., Nouaille, P., Virtanen, M., Sánchez-Sarachu, B., Pardo-Bosch, F., & Knieilng, J. (2022). Towards Sustainable and Smart Cities: Replicable and KPI-Driven Evaluation Framework. Buildings 2022, Vol. 12, 12(2). https://doi.org/10.3390/buildings12020233.

Ragab, A., Osama, A., & Ramzy, A. (2023). Simulation of the Environmental Impact of Industries in Smart Cities. Ain Shams Engineering Journal, 14(6), 102103. https://doi.org/10.1016/j.asej.2022.102103.

Shao, Q. G., Jiang, C. C., Lo, H. W., & Liou, J. J. H. (2023). Establishing a sustainable development assessment framework for a smart city using a hybrid Z-fuzzy-based decision-making approach. Clean Technologies and Environmental Policy 2023 25:9, 25(9), 3027-3044. https://doi.org/10.1007/s10098-023-02547-7.

Shamsuzzoha, A., Niemi, J., Piya, S., & Rutledge, K. (2021). Smart city for sustainable environment: A comparison of participatory strategies from Helsinki, Singapore and London. Cities, 114, 103194. https://doi.org/10.1016/j.cities.2021.103194.

Sharma, P., Bamini, J., Vijayalakshmi, S., Vijayalakshmi, N., Meenambika, A., & Singh, V. P. (2023). Exploring the Implications of IoT Integration in Urban Infrastructures for Sustainable Smart Cities. 2023 International Conference on Data Science, Agents and Artificial Intelligence, ICDSAAI 2023, 1-6. https://doi.org/10.1109/ICDSAAI59313.2023.10452458.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tate, M., Furtmueller, E., Everman, J., & Bandara, W. (2015). Introduction to the Special Issue: The Literature Review in Information Systems. Communications of the Association for Information Systems, 37(1), 5. https://doi.org/10.17705/1CAIS.03705.

Yigitcanlar, T., Kamruzzaman, M., Buys, L., Ioppolo, G., Sabatini-Marques, J., da Costa, E. M., & Yun, J. J. (2018). Understanding 'smart cities': Intertwining development drivers with desired outcomes in a multidimensional framework. Cities, 81, 145-160. https://doi.org/10.1016/j.cities.2018.04.003.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), 118–144. https://doi.org/10.1016/J.JSIS.2019.01.003

Wissenschaftsverbund Vierländerregion (ed.). (n.d.). *IoT Sustainability Lab - Nachhaltige Vierländerregion*. Retrieved April 27, 2026, from https://www.wissenschaftsverbund.org/labs/iot-and-sustainability-lab

Zimmermann, H.-D. (2025). Digitalization Meets Circular Economy. In A. Pucihar, M. Kljajić Borštnar, S. Blatnik, M. Marolt, R. Bons, K. Smits, & M. Glowatz (Eds.), 38th Bled eConference: Empowering Transformation: Shaping Digital Futures for All: Conference Proceedings (pp. 437–452). University of Maribor Press. https://doi.org/10.18690/um.fov.4.2026.27.

A Cross-Layer Integration Framework for the Development of a Platform Business Model, Applications and Services

Vjeran Strahonja, Dijana Oreški
University of Zagreb, Faculty of Organization and Informatics, Varaždin, Croatia
vjeran.strahonjaov@foi.hr, dijana.oreski@foi.hr

DOI
https://doi.org/10.18690/um.fov.4.2026.58

ISBN
978-961-299-152-4

Digital platforms have transformed contemporary economic structures by enabling multi-sided interactions, network effects, and ecosystem-based value creation. However, existing research frequently treats business models, technical architectures and service design as separate domains, resulting in conceptual fragmentation and practical misalignment. This paper proposes a cross-layer integration framework for the development of Digital Platform Business Models (DPBMs), platform applications, and platform services. The framework integrates lifecycle perspectives and construct mapping across business, service, and application layers. It introduces a bidirectional development logic that enables both top-down (business-driven) and bottom-up (technology-driven) innovation. By formalizing the transformation of Business Model Canvas (BMC) constructs into service design artifacts and Unified Modeling Language (UML) models, the framework enhances traceability, consistency, and alignment across layers. The study contributes to digital platform research by offering a structured methodological foundation for cross-layer integration and identifying future directions for tool-supported model transformation.

Keywords:
digital platform business models,
digital platform applications,
digital services,
service design artefacts,
digital platform design

University of Maribor Press

1 Introduction

The motivation for this research stems from the observation that the success of digital platforms (DPs) depends on the alignment of the business model (BM) with technical capabilities and service design. However, these dimensions are often examined separately in the literature. Business models originated as instruments of strategic management and are still most frequently viewed in this context. Nevertheless, contemporary research emphasizes their broader functional applications in entrepreneurship, operational management, financial planning, innovation management, digital transformation, and sustainable development.

Business model research focuses on value creation, delivery, and capture mechanisms, whereas technical architecture research emphasizes modularity, scalability, interoperability, and system performance. The gap between these perspectives may result in platforms that are technically sophisticated but commercially unviable, or strategically sound but technically infeasible

The subject of this research comprises business models based on digital platforms, as well as platform applications and platform services. Digital platforms have fundamentally transformed business models and organizational structures across industries by creating ecosystems that connect producers, consumers, and complementors through technology-mediated interactions. Understanding the lifecycles of these DPs, from both business model and technical architecture perspectives, is essential for practitioners and researchers seeking to design, implement, and sustain successful platform ventures.

The primary research questions are:

1. How can a business model be systematically propagated into the development of platform applications and services that support it?
2. How can a business model be digitally transformed or innovated based on platform applications and services that create new opportunities for value creation?

The ultimate goal of this research is the development of a cross-layer integration framework for the development of a platform business model, applications and services.

This paper emphasizes two aspects of the proposed framework:

1. Lifecycle aspects (Chapter 3);
2. Cross-layer mapping of constructs between business models and application/service architectures (Chapter 4).

2 Problem Positioning and Literature Review

To position the research problem, the literature review synthesizes research across four interconnected domains: (1) platform business models and their representation through the Business Model Canvas (BMC); (2) platform application architectures modeled using Unified Modeling Language (UML); (3) service lifecycle models; (4) integrated approaches that bridge business and technical concerns.

2.1 Digital Platform Business Models and the Business Model Canvas

The origins of business model development can be traced back to early twentieth-century economic theories concerning how companies create and deliver value and generate profit through innovation and market mechanisms. Systematic approaches to modern business model development emerged in the early 2000s, when Osterwalder and Pigneur (2010) introduced the Business Model Canvas (BMC), a widely adopted framework for conceptualizing and communicating business models.

The BMC comprises nine building blocks: customer segments, value propositions, channels, customer relationships, revenue streams, key resources, key activities, key partnerships, and cost structure. In the context of digital platforms, the BMC has been adapted to capture the specific characteristics of multi-sided markets, network effects, and ecosystem orchestration.

Traditionally, business models serve as instruments of strategic management. They define and connect the fundamental structural elements of a business and support business innovation, including the redefinition of value propositions, revenue logic,

cost structures, and customer relationships. Foss and Saebi (2017) conceptualize business model innovation (BMI) as organizational transformation that does not necessarily require changes in the underlying technology or product. BMI is particularly relevant in conditions of disruptive change and digitalization, where competitive advantage increasingly depends on the configuration of activities rather than solely on resource ownership.

Beyond strategic management, the business model concept serves multiple purposes. In entrepreneurship and startup development, the business model functions as a dynamic learning and validation instrument rather than a static document. Lean and agile startup approaches emphasize iterative hypothesis testing and validation (Blank, 2013).

In operational management and organizational alignment, the business model performs an integrative function. Zott, Amit, and Massa (2011) define it as a system of interconnected activities aimed at creating value for customers.

From a financial perspective, the business model describes and operationalizes the logic of revenue generation and cost management. Morris, Schindehutte, and Allen (2005) argue that the business model reflects fundamental economic mechanisms, including revenue structure, profit margins, and market positioning.

In the domain of sustainable development, sustainable business models emphasize environmental and social values, resources, and activities. Bocken et al. (2014) propose archetypes of sustainable business models that support the transition toward circular economy principles and long-term sustainability.

In the context of digital transformation and platformization, the digital platform business model (DPBM) plays a central role in the monetization of innovation. Teece (2010) emphasizes that the business model defines the manner in which a company appropriates value from technological innovation. In platform companies, the business model incorporates mechanisms such as network effects, digital ecosystems, and multi-sided market structures.

Steffen et al. (2023) identified 30 platform success factors and integrated them into the BMC framework, demonstrating its utility for platform design. These factors span corporate value integration, platform value definition, and essential IT considerations, providing a comprehensive view of what differentiates successful platforms from failures.

However, in the context of DP development, the function of BMC is primarily descriptive and configurational, not processual or architectural. Since DPs are based on multilateral interactions, network effects, ecosystem orchestration and digital infrastructure, their elaboration requires methodological approaches that go beyond the static representation of business logic. Therefore, BMC is complemented by approaches such as "lean startup" and agile development, as well as modeling methods that provide a stronger foundation for the design and operationalization of platform services and applications. The Activity System Perspective (Zott and Amit, 2024) is the concept of a business model as an activity system, where the focus is on content, structure and transaction management. From the perspective of DPs, it enables the mapping of interdependent activities of different actors, explains how the configuration of activities generates network effects and analyzes governance mechanisms (control, access rules, algorithmic management).

Unlike traditional linear business models, DMss are based on network effects, ecosystem orchestration and multi-sided markets (Parker, Van Alstyne, 2023; Daradkeh, 2023).

2.2 Digital Platform Architectures and Applications

A digital platform is an Internet-based socio-technical system that leverages digital technologies to connect participants and facilitate interactions, communication, transactions, and value exchange. From a functional perspective, digital platforms represent scalable, user-centric ecosystems rather than single, isolated applications.

Digital platform applications (DPA) require alignment between the business model and the IT architecture. Enterprise modeling approaches, such as ArchiMate, provide layered representations of business, application, and infrastructure components. Compared to the BMC, which operates at the level of business logic,

ArchiMate offers formal links between business processes and software implementation (Lankhorst, 2017).

The Unified Modeling Language (UML) serves as a standardized notation for modeling software-intensive systems. UML diagrams—such as use case diagrams, class diagrams, sequence diagrams, activity diagrams, component diagrams, and deployment diagrams—enable architects to visualize, specify, construct, and document the structure and behavior of platform systems.

Scheer (2023) utilized UML to model platform composition, detailing technical architecture layers that support application composition. The work demonstrates how UML can link development lifecycle phases to technical models, facilitating communication between business stakeholders and technical teams. However, although UML provides strong support for technical modeling, it does not explicitly represent business model constructs such as value propositions or revenue mechanisms.

2.3 Design of Digital Platform Services

Service design is inherently multidisciplinary, integrating insights from service research, design, marketing, operations management, information systems, and interaction design (Prestes Joly et al., 2019). Contemporary service design is grounded in the value co-creation perspective, which challenges traditional provider-centric approaches and aligns with service-dominant logic (Yu et al., 2018). This perspective emphasizes resource integration and collaborative value creation between providers and customers.

The theoretical foundation of contemporary service design rests on the value co-creation perspective, which challenges traditional product-oriented and provider-centric approaches to service development. This perspective aligns with service-dominant logic, which emphasizes resource integration and collaborative value creation between providers and customers (da Fontoura Vieira et. al., 2025).

Recent research highlights the need for systemic and holistic approaches to service design. Cardoso et al. (2024) propose that service design should first be understood holistically, incorporate user-centered practices for structuring development steps,

and employ co-creation mechanisms with partners to make services transparent and meaningful to users.

2.4 Cross-Layer Transformation

Ensuring traceability and consistency between the business model and platform applications and services requires the transformability of concepts and constructs across layers. The Business Model Canvas frequently serves as the primary source framework for such mappings.

Key elements frequently mapped include:

- Value Propositions: Mapped to service specifications, application functions, and service design value propositions (Koyama et al. 2023; Meertens et. al., 2012; Fritscher and Pigneur, 2020)
- Customer Segments: Aligned with personas, user roles in UML, and ArchiMate business actors (Fritscher and Pigneur, 2020; Blaschke et al., 2016)
- Channels: Corresponded to service delivery channels, application interfaces, and customer touchpoints (Allweins et al., 2021), (Koyama et al. 2023)
- Key Activities: Mapped to business processes in ArchiMate, UML activity diagrams, and service blueprint activities (Meertens et al., 2012; Fritscher and Pigneur, 2020)
- Key Resources: Aligned with ArchiMate resources, UML components, and service design resources (Meertens et al., 2012), (Lê, 2016). Platform-specific extensions like the Platform Canvas introduce additional constructs such as platform sides, network effects, and ecosystem governance mechanisms, which require specialized mappings to architectural elements (Allweins et al., 2021).

3 Lifecycle Aspects of the Cross-Layer Development of Business Models, Applications and Services

There is a clear need for integrated lifecycle approaches that align business model evolution with technical architecture and application development, as well as the design of platform services. Existing studies predominantly focus on either business model lifecycles or technical lifecycles, with limited integration. The need for integrated approaches is particularly acute in platform contexts, where business model decisions (e.g., pricing strategies, partnership models) have direct implications for technical architecture (e.g., API design, data governance), and vice versa. The following sections examine lifecycle models in each domain before returning to the integration challenge.

The proposed integration framework includes a lifecycle model and process patterns that define the phases, processes, and activities of platform business model development, as well as the development of the platform itself and its associated applications and services.

The lifecycle approach is bidirectional. Development may proceed top-down, from an existing or newly designed platform business model toward the development of services and applications. Alternatively, development may proceed bottom-up, from a platform idea or prototype toward the formulation or transformation of the business model. Iterative development cycles analogous to software engineering methodologies may be applied.

Figure 1. shows the development of a business model. It starts with a phase of inspiration and idea development. Various approaches and methods of Design Thinking, a discipline that has been developing since the middle of the last century, are applied here. General approaches and methods are specialized in service design. Service Design Thinking (SDT) is a human-centered, co-creative, iterative approach that treats services as end-to-end experiences across touchpoints, frontstage/backstage activities, actors, and supporting systems (Stickdorn, Schneider, 2012). It helps co-design across multiple participant groups (e.g., consumers, providers, partners).

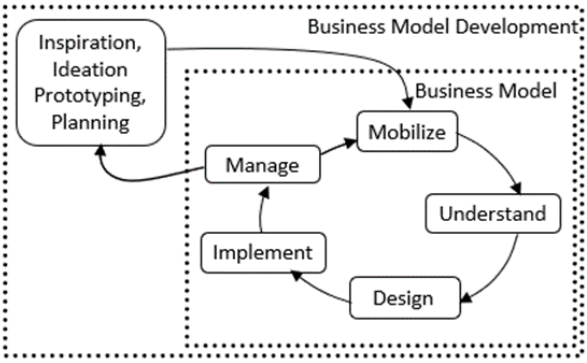

Figure 1: Business Model Development Cycle
Source: Own

According to Alexander Osterwalder and Yves Pigneur (2010) and BMC, the business model development cycle includes mobilization, understanding, design, implementation, and management phases. The transition from the business model layer to the service design layer entails the development of platform applications and customer-centric service models focused on user experience.

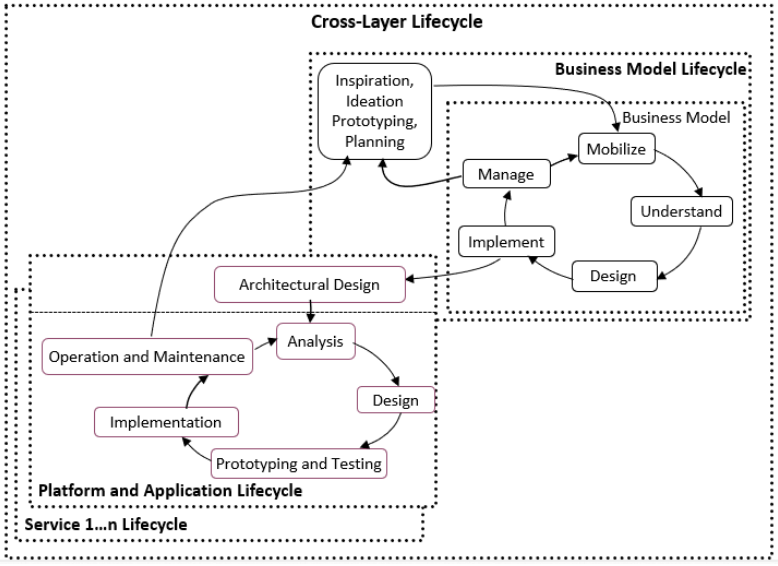

Figure 2: Cross-Layer Lifecycle
Source: Own

Platform applications and services are the implementation of platform business models, as shown on the Cross-Layer Lifecycle (Figure 2).

Methodological approaches of different authors or companies, differ in the number, names and content of the phases, but it is mostly possible to map Lifecycles among different methodological approaches.

Architectural design of digital platform applications encompasses the definition of system components, interfaces (APIs), data flows, infrastructure, security mechanisms, and scalability strategies. Analogously, the architectural design of a platform service encompasses defining the system structure, key components, interfaces (APIs), data flows, security mechanisms, infrastructure, and scalability strategies to ensure reliable, secure, and efficient service delivery.

Platform application and service lifecycles typically include analysis, design, prototyping and testing, implementation, and operation and maintenance. All these phases are carried out in accordance with the chosen methodology of business modeling, and the development of software and digital services.

Development is typically incremental and iterative, in accordance with agile methodologies that combine best practices from systems and software engineering.

The overall cross-cycle development spans three distinct lifecycles, each with its own focus and progression. The development can be conducted in an incremental and iterative manner, consistent with agile development principles.

The figure does not show who initiates development cycles and government mechanisms, as well as the layers of other management processes, such as management of changes, configuration, releases, quality, etc.

Changes in the business model are most often driven by the strategic management of the organization. However, in cyclical development, the impetus for business model development can come from the operation and maintenance phases of the development cycle of platform applications or services, where evaluation can indicate that modifications are needed within the BM. The new development cycle

may also be triggered by technological advancements, the emergence of new tools and platforms, as well as by the redesign or improvement of platform services.

4 Cross-Layer Mapping of Constructs Between Business Models, Application and Service Architectures and Models

Business models, such as those based on the BMC framework, provide high-level representations of value creation, delivery, and capture. Since key platform service design models contain similar constructs, such as value, activity, resource, partner communication, only elaborated in more detail, the conversion of BM constructs into SDM constructs will not be presented in more detail in this article.

However, platform application development requires formal modeling constructs capable of specifying system structure, behavior, and interactions. Unified Modeling Language (UML) offers such formalization through standardized diagrams (e.g., use case, class, sequence, activity, and component diagrams).

This section analyzes how key BMC elements (Value Proposition, Customer Segments, Channels, Customer Relationships, Key Partners, Key Resources, and Key Activities) are transformed into UML artifacts during platform application development (Fritscher and Pigneur, 2020). The objective is to clarify which business-level concepts serve as direct inputs for UML modeling and how they are translated into system-level abstractions.

Table 1 summarizes the mapping of BMC constructs to Platform Architecture layer constructs, ArchiMate constructs and UML concepts with a description of the transformation.

Some constructs are directly transformed from BMC to UML (e.g., Customer Segments → Actors; Key Activities → Activity Diagrams; Key Resources → Domain Classes).

Table 1: Mapping of BMC constructs to Platform Architecture, ArchiMate constructs and UML constructs and concepts

BMC Construct	Platform Architecture Layer Construct	ArchiMate Construct	UML Concept(s)	Transformation into UML (Detailed Description)
Value Proposition	Platform Service / Platform Capability	Business Service, Value, Capability	Use Case, Service Interface, Component	Value propositions are decomposed into platform services. Each service becomes a Use Case. Services exposed via APIs are modeled as UML interfaces. At the structural level, they become Components implementing those interfaces.
Customer Segments	Platform Actor Roles (e.g., Provider, Consumer, Admin)	Business Actor, Business Role	Actor (Use Case Diagram), Actor Generalization	Customer segments are abstracted into system roles. Each role becomes a UML Actor. If multiple subtypes exist (e.g., Premium User), Actor generalization hierarchies are introduced.
Channels	Interaction Layer (Web App, Mobile App, API Gateway)	Application Interface, Application Component	Boundary Class, Interface, Component	Channels are operationalized as interaction interfaces. In UML, they become boundary classes or provided interfaces in component diagrams. They define system entry points.
Customer Relationships	Engagement Mechanisms (Subscription logic, Reputation System)	Business Process, Application Function	State Machine Diagram, Sequence Diagram	Relationship mechanisms (e.g., subscription lifecycle) are transformed into state machines describing user states and transitions. Interaction rules are formalized in sequence diagrams.
Key Activities	Core Platform Processes (Matchmaking, Payment Processing)	Business Process, Application Function	Activity Diagram, Sequence Diagram	Activities become workflows. Each key activity is modeled as an activity diagram. Interactions among components during execution are captured in sequence diagrams.
Key Resources	Platform Assets (Data, Algorithms, Infrastructure)	Data Object, Application Component, Artifact	Class Diagram, Component Diagram	Data resources become domain classes. Algorithms become service components. Infrastructure becomes deployment nodes (in Deployment Diagrams if extended).
Key Partners	External Systems / Ecosystem Actors	Business Actor, Application Collaboration	External Actor, External Component, Dependency	Partners (e.g., Payment Provider) become external actors or components. Dependencies are modeled explicitly in component diagrams and interaction flows in sequence diagrams.

Source: own

Other constructs require derived transformations (e.g., Value Propositions → Use Cases; Customer Relationships → State/Interaction Models; Channels → Interface Definitions).

Revenue streams and cost structures function primarily as contextual inputs for requirements engineering and are not directly mapped to UML structural elements.

A major challenge for future research lies in the development of computer-supported tools enabling systematic cross-layer transformation between BMC components, service design artifacts, and UML architectural models.

5 Conclusion

Digital platform development requires systematic alignment between business logic, service design, and technical architecture. The proposed cross-layer integration framework aligns lifecycle phases and formalizes transformation rules between business model constructs and architectural artifacts. By bridging strategic and technical perspectives, the framework enhances methodological rigor and practical applicability. Future research should focus on meta-model formalization and computer-supported transformation tools to enable automated traceability across layers.

The proposed framework is founded on the cyclical and iterative co-development and consolidation of three interrelated domains: (1) business model development through the application of the Business Model Canvas framework, (2) service development employing selected service design methods, and (3) platform (application) development using Unified Modeling Language. The primary objective of the framework is to reduce the methodological discontinuities among these domains by enabling the systematic integration and cross-utilization of artifacts generated within each domain throughout the development process. In addition, the framework seeks to establish coordinated and synchronized development cycles, thereby facilitating the continuous alignment of business, service, and technological perspectives during the iterative refinement of the overall system.

The proposed framework reduces discontinuities between strategic, operational, and technical modeling domains by enabling cyclical co-development and traceable artifact propagation across business model, service design, and platform engineering activities.

Acknowledgment

The authors would like to acknowledge the support given by the European Commission through the Interreg Danube Region Programme Digital services for Circular Economy - a Toolbox for Regional Developers & SME (DECIDE).

References

Allweins, M. M., Proesch, M., & Ladd, T. (2021). The platform canvas—conceptualization of a design framework for multi-sided platform businesses. *Entrepreneurship Education and Pedagogy*, *4*(3), 455-477.

Blank, S. (2013). Why the lean start-up changes everything. *Harvard Business Review, 91(5)*, 63–72.

Blaschke, M., Cigaina, M., Riss, U. V., & Shoshan, I. (2016). Designing business models for the digital economy. In *Shaping the digital enterprise: Trends and use cases in digital innovation and transformation* (pp. 121-136). Cham: Springer International Publishing

Bocken, N.M.P., Short, S.W., Rana, P., & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. Journal of Cleaner Production, 65, 42–56.

Cardoso, A. M. K., Canciglieri Junior, O., & Benitez, G. B. (2024). Service design: a critical examination and future research directions in servitization literature. *Journal of Business & Industrial Marketing, 39*(9), 1950-1968.

Daradkeh, M. (2023, May). Exploring the boundaries of success: A literature review and research agenda on resource, complementary, and ecological boundaries in digital platform business model innovation. In *Informatics* (Vol. 10, No. 2, p. 41). MDPI.

da Fontoura Vieira, J. F., Echeveste, M. E. S., Tinoco, M. A. C., Marcon, A., Marcon, É., & Lermen, F. H. (2025). Enhancing value cocreation orientation in service innovation: a new service development and service design integrated process. *Humanities and Social Sciences Communications*, *12*(1), 31.

Foss, N.J., & Saebi, T. (2017). Fifteen years of research on business model innovation. Journal of Management, 43(1), 200–227.

Fritscher, B., & Pigneur, Y. (2020). A visual approach to business IT alignment between business model and enterprise architecture. In *Sustainable Business: Concepts, Methodologies, Tools, and Applications* (pp. 543-566). IGI Global.

Koyama, S., Tsuji, Y., Masuda, R., Onoda, H., Ono, H., Natsuyama, M., ... & Yamamoto, S. (2023). A Technique to integrate service business models with archimate. *Procedia Computer Science*, *219*, 479-485.

Lankhorst, M. (2017). Enterprise Architecture at Work: Modeling, Communication and Analysis. Springer.

Lê, L.S., (2016). Enterprise Architecture. Mapping of BMC, DEMOand ArchiMate. *Master Thesis Leiden Institute of Advanced Computer Science (LIACS), theses.liacs.nl/pdf/Le-Minh-non-confidential.pdf*

Meertens, L. O., Iacob, M. E., Nieuwenhuis, L. J., van Sinderen, M. J., Jonkers, H., & Quartel, D. (2012). Mapping the business model canvas to ArchiMate. In *Proceedings of the 27th annual ACM symposium on applied computing* (pp. 1694-1701)

Morris, M., Schindehutte, M., & Allen, J. (2005). The entrepreneur's business model: Toward a unified perspective. Journal of Business Research, 58(6), 726–735.

Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: a handbook for visionaries, game changers, and challengers* (Vol. 1). John Wiley & Sons.

Parker, G., & Van Alstyne, M. (2023). Platforms: Their structure, benefits, and challenges. In *Introduction to Digital Humanism: A Textbook* (pp. 523-542). Cham: Springer Nature Switzerland.

Prestes Joly, M., Teixeira, J. G., Patrício, L., & Sangiorgi, D. (2019). Leveraging service design as a multidisciplinary approach to service innovation. *Journal of service Management*, *30*(6), 681-715.

Scheer, A. W. (2023). Application composition platform architecture. In *Digital Business Engineering* (pp. 47-58). Springer. https://doi.org/10.1007/978-3-658-42483-1_4

Stickdorn, M., & Schneider, J. (2012). *This is service design thinking: Basics, tools, cases*. John Wiley & Sons.

Steffen, M., Hunke, F., Hagen, S., & Schlereth, A. (2023). Success factors of digital platform design. *ITM Web of Conferences*, 51, 05001. https://doi.org/10.1051/itmconf/20235105001

Teece, D. J. (2010). Business models, business strategy and innovation. Long Range Planning, 43(2–3), 172–194.

Yu, E., & Sangiorgi, D. (2018). Service design as an approach to implement the value cocreation perspective in new service development. *Journal of Service Research*, *21*(1), 40-58.

Zott, C., Amit, R., & Massa, L. (2011). The business model: Recent developments and future research. Journal of Management, 37(4), 1019–1042.

Zott, C., & Amit, R. (2024). Business models and lean startups. Journal of Management, 50(8), 3183-3201.

Effects of Personality on Perceived Immersion in Low-Immersive Virtual Reality Stores

Markus Makkonen,[1, 2] Lauri Frank,[2]
Tiina Kemppainen,[3] Matilda Holkkola,[2]
Tiina Kaleton[2]

[1] Tampere University, Faculty of Management and Business, Tampere, Finland
markus.makkonen@tuni.fi
[2] University of Jyvaskyla, Faculty of Information Technology, Jyvaskyla, Finland
lauri.frank@jyu.fi, matilda.i.holkkola@jyu.fi, tiina.e.kaleton@jyu.fi
[3] University of Jyvaskyla, School of Business and Economics, Jyvaskyla, Finland
tiina.j.kemppainen@jyu.fi

Virtual reality (VR) stores can be differentiated into two main types: immersive and low-immersive. In immersive VR stores, the virtual store environment is typically experienced by putting on a head-mounted display, whereas in low-immersive VR stores, it is typically experienced via the flat screen of some commonly used information technology device, such as a computer or a mobile phone. In this study, we examine how the Big Five personality traits of extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience affect perceived immersion in low-immersive VR stores. To do this, we first collect data from 420 US-based consumers via an online survey in which the respondents are requested to visit a real-life low-immersive VR store. We then analyse the collected data with partial least squares structural equation modelling (PLS-SEM). Our findings suggest that especially conscientiousness, but also extraversion and openness to experience, positively affect perceived immersion in low-immersive VR stores.

DOI
https://doi.org/10.18690/um.fov.4.2026.59

ISBN
978-961-299-152-4

Keywords:
personality, Big Five personality traits, perceived immersion, virtual reality, low-immersive virtual reality stores

University of Maribor Press

1 Introduction

Over the recent years, the use of virtual reality (VR) in marketing and retail has captured the attention of both academics (e.g., Alcañiz et al., 2019; Loureiro et al., 2019; Xi & Hamari, 2021; Barrera & Shah, 2023; Yoo et al., 2023; Lu et al., 2025) and practitioners alike, which is not surprising when considering the forecasted growth of the global market of VR in retail from $3.33 billion in 2023 to $26.28 billion by 2032 (SNS Insider, 2025). However, when talking about the use of VR in marketing and retail, or in any other application area for that matter, it is important to note that there exists not only one kind of VR but numerous different kinds of VR, of which two notable examples are *immersive VR* and *non-immersive VR* (e.g., Mills & Noyes, 1999; Suh & Lee, 2005; Suh & Prophet, 2018). Of these, the former typically necessitates users putting on a head-mounted display (HMD) to experience the virtual environment in which they are fully enclosed. In contrast, the latter typically does not necessitate users putting on an HMD, but the virtual environment can be experienced via the flat screen of some commonly used information technology (IT) device, such as a desktop computer, a laptop computer, a tablet computer, or a mobile phone. In technological terms, non-immersive VR can be seen as somewhat less sophisticated than immersive VR. However, in practical terms, it is still able to offer users many similar benefits as immersive VR, such as the experience of telepresence (Peukert et al., 2019). This means that non-immersive VR can also be seen as immersive to some extent and should, therefore, more accurately be called *low-immersive VR* (Peukert et al., 2019). Because of this, also in this paper, we will be using the term low-immersive VR instead of non-immersive VR.

A similar differentiation can be made between *immersive VR stores* and *low-immersive VR stores* (e.g., Lombart et al., 2020; Zhang et al., 2025), of which the former typically necessitate putting on an HMD to experience the virtual store environment, whereas the latter do not typically necessitate putting on an HMD, but consumers can experience the virtual store environment via the flat screen of their commonly used IT device, such as a desktop computer, a laptop computer, a tablet computer, or a mobile phone. Of these, the low-immersive VR stores, which are also commonly called *360-degree virtual stores* or more simply *360-virtual stores* (Ruusunen et al., 2023), can be considered an extremely tempting option for many retailers to begin experimenting with VR and the metaverse because although they not necessitate

consumers owning and putting on an expensive HMD to experience the virtual store environment, they are still able to offer consumers immersive experiences to some extent and, therefore, also many similar benefits as immersive VR stores (e.g., Peukert et al., 2019; Lombart et al., 2020; Zhang et al., 2025). However, few prior studies have examined how immersive the experiences offered by low-immersive VR stores to consumers actually are and what kind of factors affect the immersion perceptions of consumers when they are visiting such stores. One critical factor is likely to be the personal characteristics of consumers, and especially the personality of consumers.

Although personality has been found to affect almost all aspects of human behaviour (e.g., Roccas et al., 2002), few prior studies have examined its effects on perceived immersion. In this study, our objective is to address this gap in the context of low-immersive VR stores by examining *how the Big Five personality traits of extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience affect perceived immersion in low-immersive VR stores*. To do this, we first collect data from 420 US-based consumers via an online survey in which the respondents are requested to visit a real-life low-immersive VR store. We then analyse this collected data with partial least squares structural equation modelling (PLS-SEM). Our findings suggest several interesting effects of the Big Five personality traits on perceived immersion, which have important implications for both theory and practice.

2 Theoretical Background

The two main theoretical concepts of this paper are personality and perceived immersion. Today, the Big Five personality traits of extraversion, agreeableness, conscientiousness, emotional stability (or neuroticism), and openness to experience, which are also commonly referred to as the five-factor model (FFM) of personality, are the dominant paradigm for conceptualising human personality (e.g., Roccas et al., 2002). These five personality traits have been found to successfully capture human personality in a wide variety of contexts, ranging from clinical and counselling to industrial and organisational contexts (e.g., Roccas et al., 2002), which is why we have chosen to apply them also in this study. Often, the Big Five personality traits are defined by describing how individuals who score either high or low in each of the traits tend to be (e.g., Roccas et al., 2002). First, individuals who are extraverted have the tendency to be sociable, talkative, assertive, and active,

whereas individuals who are introverted have the tendency to be retiring, reserved, and cautious. Second, individuals who are agreeable have the tendency to be good-natured, compliant, modest, gentle, and cooperative, whereas individuals who are disagreeable have the tendency to be irritable, ruthless, suspicious, and inflexible. Third, individuals who are conscientious have the tendency to be careful, thorough, responsible, organised, and scrupulous, whereas individuals who are unconscientious have the tendency to be irresponsible, disorganised, and unscrupulous. Fourth, individuals who are emotionally stable (or not neurotic) have the tendency to be calm, poised, and emotionally stable, whereas individuals who are emotionally unstable (or neurotic) have the tendency to be anxious, depressed, angry, and insecure. And finally, fifth, individuals who are open to experience have the tendency to be intellectual, imaginative, sensitive, and open-minded, whereas individuals who are not open to experience have the tendency to be down-to-earth, insensitive, and conventional.

In turn, perceived immersion refers to a psychological state experienced by individuals in which they may become so involved, absorbed, engaged, or engrossed in an environment or experience that they lose awareness of their surroundings in terms of place and time (Agrawal et al., 2020). As summarised in the recent literature review by Colombini et al. (2026), few prior studies have examined the effects of the Big Five personality traits on perceived immersion. And those that have, have made inconsistent findings. For example, the prior study by Jasper et al. (2023) found extraversion and agreeableness – but not conscientiousness, neuroticism (or emotional stability), and openness to experience – to have a positive effect on perceived immersion. In contrast, the prior study by Weibel et al. (2010) found neuroticism (or emotional stability), openness to experience, and extraversion – but not agreeableness and conscientiousness – to have a positive effect on immersive tendency. Thus, although many of the Big Five personality traits have been found to have mainly positive (or negative in the case of emotional stability) effects on perceived immersion, the exact effects seem to vary considerably between different studies conducted in different contexts, thus calling for more context-specific examinations. In this study, we conduct such a more context-specific examination in the specific context of low-immersive VR stores, which has so far been completely overlooked in prior research. More particularly, we propose and test the following five hypotheses concerning the effects of the Big Five personality traits on perceived immersion:

H1: Extraversion has a positive effect on perceived immersion.

H2: Agreeableness has a positive effect on perceived immersion.

H3: Conscientiousness has a positive effect on perceived immersion.

H4: Emotional stability has a negative effect on perceived immersion.

H5: Openness to experience has a positive effect on perceived immersion.

3 Methodology

The data for this study was collected identically to a prior study by Makkonen (2026) with an online survey conducted by using the LimeSurvey service in September 2025. It consisted of three parts. We first asked the respondents questions about their basic background information, such as demographics and personality. We then requested the respondents to visit a real-life low-immersive VR store for about five minutes. Finally, we asked the respondents questions about their visit to the store and their perceptions and intention of using these kinds of stores in the future. Each respondent was randomly assigned to visit one of two low-immersive VR stores belonging to Amazon's Beyond program: the Logitech store or the Star Wars store. This was to promote the generalisability of our findings by covering not only one but a wider variety of stores in terms of their product offerings, with the products of the Logitech store being more utilitarian (e.g., mice, keyboards, headphones, loudspeakers, and web cameras) and the products of the Star Wars store being more hedonic (e.g., clothing, collectables, and toys). For illustrations of the two stores, please see Appendix B. In both stores, the respondents were able to freely move in and explore the virtual store environment enriched with both audio and video content, view the offered products embedded in the environment, and click on the offered products to see more information about them or purchase them. Based on their visit, the respondents were also requested to respond to a question concerning the price of one particular product in the store to prove that they had really conducted the visit.

In the survey, the Big Five personality traits were measured by using the ten-item personality inventory (TIPI) by Gosling et al. (2003), whereas perceived immersion was measured by using a four-item scale adapted from tom Dieck et al. (2023). These scales were selected due to their wide use, brevity, and clarity. In the wordings of the items that measured perceived immersion, low-immersive VR stores were referred to simply as "virtual stores", which were defined in the survey as "online stores that enable you to move and interact in three-dimensional 360-degree environments via your screen without a virtual reality headset or any other extra equipment". For the precise wordings of all the items, please see Appendix A. The standard seven-point Likert scale, ranging from strong disagreement (1) to strong agreement (7), was used to measure all the items. To avoid forced responses, the measurement scale also allowed for missing values. Before its launch, four researchers who had experience in the topic pilot tested the survey, resulting in some small revisions. The average completion time of the survey was about 22.5 minutes.

We recruited the respondents by using the Prolific crowdsourcing service. In general, crowdsourcing services have been found to be a reliable and valid data collection method in information systems (IS) research (Lowry et al., 2016). More specifically, the Prolific service has been found to provide a more heterogeneous participant population and at least equal data quality in comparison to its competitors, such as the Amazon Mechanical Turk (MTurk) service (Peer et al., 2017, 2022). To promote data quality, we adhered to the guidelines for using crowdsourcing services in IS research by Jia et al. (2017) by adding two attention trap items in the survey and recruiting only respondents who had at least 20 but no more than 10,000 prior submissions and whose approval rate for their prior submissions was at least 98%. To homogenise the visiting experience as much as possible, we also recruited only respondents who were fluent in English, resided in the US, and were willing to complete the survey by using either a desktop computer or a laptop computer. All the respondents who completed the survey received a minor monetary reward of a few US dollars, on average exceeding the US federal minimum wage of $7.25 per hour.

When analysing the collected data, the Big Five personality traits constructs were inserted into the estimated model as composite variables formed by first unreversing the coding of the reverse-coded items and then averaging the scores of the two items that measured each personality trait. Model estimation was conducted with partial

least squares structural equation modelling (PLS-SEM) by using the SmartPLS 4.1.1.8 software (Ringle et al., 2026) and adhering to the guidelines by Hair et al. (2022). For example, bootstrapping with 10,000 subsamples was used to test the statistical significance of the model estimates, whereas the statistical significance of potential differences in the model estimates between the Logitech and Star Wars subsamples was tested by using bootstrapped multi-group analysis with 10,000 subsamples, in which the Logitech and Star Wars subsamples constituted the two compared groups. In both, $p < 0.05$ was used as the threshold for statistical significance. Mean replacement was used to handle the potential missing values. When estimating the model, we also used gender, age, educational attainment, and prior experience in visiting low-immersive VR stores as control variables for perceived immersion.

4 Results

4.1 Sample Statistics

The survey was completed by 455 respondents, of whom 35 (7.7%) had to be discarded due to incorrect responses to the two attention trap items or to the question concerning product price. Thus, 420 respondents formed the final sample. Table 1 reports the descriptive statistics of this sample in terms of the gender, age, educational attainment, and prior experience in visiting low-immersive VR stores of the respondents. The sample had a very balanced gender distribution between men and women (49.3% vs. 49.8%). The age of the respondents ranged from 19 to 81 years, with a mean of 42.9 years and a standard deviation of 13.0 years. In terms of education, most of the respondents (66.0%) had attained an undergraduate, graduate, or postgraduate degree, diploma, or certificate. And in terms of prior experience in visiting low-immersive VR stores, most of the respondents (66.7%) had never previously visited such a store. In addition, Table 1 reports the visited low-immersive VR store during the survey, showing that 207 respondents (49.3%) visited the Logitech store and 213 respondents (50.7%) visited the Star Wars store.

Table 1: Sample statistics (N = 420)

	N	%
Gender		
Man	207	49.3
Woman	209	49.8
Other	4	1.0
Age		
18–29 years	64	15.2
30–39 years	133	31.7
40–49 years	88	21.0
50–59 years	90	21.4
60 years or over	45	10.7
Educational attainment		
Less than secondary school or high school	2	0.5
Secondary school or high school	65	15.5
Some post-secondary studies but no degree, diploma, or certificate	76	18.1
Undergraduate degree, diploma, or certificate	191	45.5
Graduate or postgraduate degree, diploma, or certificate	86	20.5
Prior experience in vising low-immersive VR stores		
Has never visited	280	66.7
Has visited 1–2 times	75	17.9
Has visited 3–5 times	29	6.9
Has visited 6–10 times	8	1.9
Has visited 10 times or more	15	3.6
No response	13	3.1
Visited low-immersive VR store		
Logitech	207	49.3
Star Wars	213	50.7

4.2 Measurement Reliability and Validity

The reliability of the perceived immersion and the Big Five personality traits constructs was evaluated by using their Cronbach's alpha (CA), which are reported in Table 2. For all constructs except for agreeableness and openness to experience, these were found to be greater than or equal to 0.7, thus suggesting acceptable reliability (Nunnally & Bernstein, 1994). However, we also consider the lower CAs of agreeableness (0.508) and openness to experience (0.624) as acceptable because they were found to be in the typical range from 0.40 to 0.73 of the TIPI scale by Gosling et al. (2003). In addition, as instructed by Fornell and Larcker (1981), the reliability of the reflectively measured perceived immersion construct was evaluated by using its composite reliability (CR), which is reported in Table 2. This was found to be greater than or equal to 0.7, thus suggesting acceptable reliability (Nunnally & Bernstein, 1994). In turn, the validity of the reflectively measured perceived

immersion construct was evaluated by using the two criteria proposed by Fornell and Larcker (1981), which enable its evaluation from the aspects of both convergent and discriminant validity. They both are based on the average variance extracted (AVE) of a construct, which is the average proportion of variance explained by a construct in its indicators. Of these, the first criterion expects a construct to have an AVE of greater than or equal to 0.5 to suggest acceptable convergent validity. In turn, the second criterion expects a construct to have a square root of AVE greater than or equal to its absolute correlations with the other model constructs to suggest acceptable discriminant validity. The AVE of the perceived immersion construct, its square root (the on-diagonal cells of the correlation matrix), and the correlations of the perceived immersion construct with the other model constructs (the off-diagonal cells of the correlation matrix) are reported in Table 2, showing that the perceived immersion construct met both these criteria.

Table 2: Construct statistics (N = 420)

	CA	CR	AVE	1	2	3	4	5	6
1 Perceived immersion	0.880	0.916	0.732	0.856					
2 Extraversion	0.733	–	–	0.246	–				
3 Agreeableness	0.508	–	–	0.249	0.248	–			
4 Conscientiousness	0.720	–	–	0.340	0.250	0.480	–		
5 Emotional stability	0.809	–	–	0.226	0.353	0.411	0.519	–	
6 Openness to experience	0.624	–	–	0.254	0.435	0.255	0.231	0.223	–
7 Gender	–	–	–	0.137	0.039	0.160	-0.004	-0.074	0.114
8 Age	–	–	–	0.035	0.111	0.223	0.112	0.182	-0.012
9 Education	–	–	–	0.103	0.122	0.102	0.186	0.181	0.136
10 Prior experience	–	–	–	0.059	0.098	0.036	0.006	0.045	0.167

Finally, the reliability and validity of the indicators that were reflectively measuring the perceived immersion construct were evaluated by using their standardised loadings, which are reported in Table 3. In the typical case of each indicator loading on only one construct, these are typically expected to be statistically significant and greater than or equal to 0.707 (Fornell & Larcker, 1981) in order to suggest acceptable reliability and validity. As shown, this criterion was met by all indicators. In addition, Table 3 reports the means (M) and standard deviations (SD) of and the

percentages of missing values in the indicator scores for all the indicators measuring the perceived immersion and the Big Five personality traits constructs.

Table 3: Item statistics (N = 420, *** = p < 0.001)

Item	M	SD	Missing	Loading	Item	M	SD	Missing	Loading
PI1	5.401	1.564	0.2%	0.881***	**AGR2**	5.307	1.709	0.0%	–
PI2	4.538	1.743	0.0%	0.843***	**CON1**	5.784	1.312	1.0%	–
PI3	5.064	1.618	0.0%	0.886***	**CON2**	5.557	1.613	0.0%	–
PI4	4.340	1.793	0.0%	0.811***	**ES1**	5.459	1.462	0.5%	–
EXT1	3.809	1.968	0.2%	–	**ES2**	4.810	1.839	0.0%	–
EXT2	3.077	1.820	0.7%	–	**OTE1**	5.513	1.307	0.7%	–
AGR1	5.792	1.278	1.4%	–	**OTE2**	5.307	1.696	0.0%	–

4.3 Estimation Results

Table 4: Estimation results (N = 420, *** = p < 0.001, ** = p < 0.01, * = p < 0.05)

	All	Logi-tech	Star Wars	Diffe-rence
Hypothesised effects				
H1: Extraversion → Perceived immersion	0.110*	0.129	0.086	-0.043
H2: Agreeableness → Perceived immersion	0.051	0.032	0.056	0.024
H3: Conscientiousness → Perceived immersion	0.252***	0.308**	0.213**	-0.095
H4: Emotional stability → Perceived immersion	0.025	0.025	0.021	-0.004
H5: Openness to experience → Perceived immersion	0.106	-0.015	0.234**	0.249*
Effects of the control variables				
Gender → Perceived immersion	0.129*	0.113	0.132	0.019
Age → Perceived immersion	-0.050	-0.086	-0.023	0.063
Education → Perceived immersion	0.025	-0.003	0.015	0.018
Prior experience → Perceived immersion	0.030	0.015	0.047	0.032
R^2				
Perceived immersion	0.178***	0.170**	0.211***	0.041

Table 4 reports the estimation results in terms of the standardised path coefficients of the hypothesised effects and the effects of the control variables, first for the whole sample (N = 420) and then for the Logitech (N = 207) and Star Wars (N = 213) subsamples. As can be seen, in the whole sample, only conscientiousness (0.252***), extraversion (0.110*), and gender (0.129*) were found to have statistically significant and positive effects on perceived immersion, suggesting that women and individuals who are more conscientious and extraverted perceive low-immersive VR stores as more immersive. In the Logitech subsample, only conscientiousness (0.308**) was found to have statistically significant and positive effect on perceived immersion,

whereas in the Star Wars subsample, both conscientiousness (0.213**) and openness to experience (0.234**) were found to have statistically significant and positive effects on perceived immersion. The effect of openness to experience on perceived experience was also the only effect in which a statistically significant difference (0.249*) was found between the Logitech and Star Wars subsamples. In terms of the proportion of explained variance (R^2) in the perceived immersion construct, the Big Five personality traits constructs and the control variables were together found to explain 17.8% of the variance in the whole sample, 17.0% of the variance in the Logitech subsample, and 21.1% of the variance in the Star Wars subsample.

5 Discussion and Conclusion

In this study, we examined the effects of the Big Five personality traits on perceived immersion in low-immersive VR stores. Of the five hypotheses concerning the effects of extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience on perceived immersion, we found either full (supported in the whole sample and in both of the two subsamples) or partial (supported in at least one of the three samples or subsamples) support for three hypotheses (i.e., H1, H3, and H5), whereas no support was found for two hypotheses (i.e., H2 and H4).

First, we found conscientiousness to have the most stable effect on perceived immersion in terms of being positive both in the whole sample and in the Logitech and Star Wars subsamples. This finding is best explained by the fact that individuals who score higher in conscientiousness tend to be more organised by nature, thus likely being able to better concentrate on the virtual store environments of low-immersive VR stores when visiting them and thus also perceiving these environments as more immersive. Second, we also found extraversion to have a positive effect on perceived immersion in the whole sample, although this effect was found to be weaker than the effect of conscientiousness on perceived immersion, so much so that no statistically significant effect could be found when examined separately in the Logitech and Star Wars subsamples. This finding is best explained by the fact that individuals who score high in extraversion tend to be more sociable by nature, thus being more ready to engage in social interactions that may be encountered in the virtual store environments of low-immersive VR stores and thus also perceiving these environments as more immersive. But because such social interactions may not be encountered in the virtual store environments of all low-

immersive VR stores, the effect of extraversion on perceived immersion remains weaker than the effect of conscientiousness on perceived immersion. Third, we found openness to experience to have the most interesting effect on perceived immersion because whereas this effect was found to be positive in the Star Wars subsample, no such effect was found in the Logitech subsample or in the whole sample. This finding is best explained by the fact that individuals who score higher in openness to experience tend to be more open-minded by nature, thus being able to better imagine themselves in different environments. This is likely to make it easier for them to imagine themselves especially in the virtual store environments of low-immersive VR stores that are not mere replicas of physical store environments (as was the case of the Logitech store) but more fantasy-like environments (as was the case of the Star Wars store), thus helping them to perceive especially these latter environments as more immersive.

We see that our study makes important theoretical and practical contributions. From a theoretical perspective, our study promotes the current understanding of the effects of the Big Five personality traits on perceived immersion in low-immersive VR stores by both supporting and conflicting with some of the findings of prior research. For example, our study supports the positive effects of extraversion and openness to experience on perceived immersion, which were also found by Weibel et al. (2010), but conflicts with the studies by both Jasper et al. (2023) and Weibel et al. (2010) by finding also conscientiousness to have a positive effect on perceived immersion. From a practical perspective, the retailers managing low-immersive VR stores can use this understanding to further promote the adoption and diffusion of their stores by better targeting them to consumers with different personality traits. An example of this would be using low-immersive VR stores with more fantasy-like virtual store environments to target consumers who are more open to experience.

6 Limitations and Future Research

We see that our study has three main limitations, each of which calls for future research to appropriately address them. First, our study focused on consumers who resided only in the US, thus calling for future studies focusing on consumers residing also in other countries, especially in countries with different cultures. Second, although our study was based on a survey in which the respondents were requested to visit a real-life low-immersive VR store, this store was visited only for about five

minutes, thus limiting the visiting experience and raising a call for future studies that are less cross-sectional and more longitudinal. Third, although our study examined the effects of personality traits on perceived immersion in the case of not only one but two real-life low-immersive VR stores, of which one offered more utilitarian products and the other offered more hedonic products – these two stores are only two examples of low-immersive VR stores, thus calling for future studies focusing on a wider variety of stores in terms of the offered products and otherwise.

Acknowledgment

This study was funded by the Foundation of Finnish Business College.

References

Agrawal, S., Simon, A., Bech, S., Bæntsen, K., & Forchhammer, S. (2020). Defining immersion: Literature review and implications for research on audiovisual experiences. *Journal of the Audio Engineering Society*, *68*(6), 404–417. doi:10.17743/jaes.2020.0039

Alcañiz, M., Bigné, E., & Guixeres, J. (2019). Virtual reality in marketing: A framework, review, and research agenda. *Frontiers in Psychology*, *10*, 1530. doi:10.3389/fpsyg.2019.01530

Barrera, K. G., & Shah, D. (2023). Marketing in the Metaverse: Conceptual understanding, framework, and research agenda. *Journal of Business Research*, *155 A*, 113420. doi:10.1016/j.jbusres.2022.113420

Colombini, G., Frigati, L., Gori, N., Duradoni, M., & Guazzini, A. (2026). What makes us feel (tele)present? A PRISMA systematic review of psychological factors. *International Journal of Human–Computer Interaction*, 1–42. doi:10.1080/10447318.2026.2631665

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. doi:10.1177/002224378101800104

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. doi:10.1016/S0092-6566(03)00046-1

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Thousand Oaks, CA: Sage.

Jasper, A., Sepich, N. C., Gilbert, S. B., Kelly, J. W., & Dorneich, M. C. (2023). Predicting cybersickness using individual and task characteristics. *Computers in Human Behavior*, *146*, 107800. doi:10.1016/j.chb.2023.107800

Jia, R., Steelman, Z. R., & Reich, B. H. (2017). Using Mechanical Turk data in IS research: Risks, rewards, and recommendations. *Communications of the Association for Information Systems*, *41*(14), 301–318. doi:10.17705/1CAIS.04114

Lombart, C., Millan, E., Normand, J.-M., Verhulst, A., Labbé-Pinlon, B., & Moreau, G. (2020). Effects of physical, non-immersive virtual, and immersive virtual store environments on consumers' perceptions and purchase behavior. *Computers in Human Behavior*, *110*, 106374. doi:10.1016/j.chb.2020.106374

Loureiro, S. M. C., Guerreiro, J., Eloy, S., Langaro, D., & Panchapakesan, P. (2019). Understanding the use of virtual reality in marketing: A text mining-based review. *Journal of Business Research*, *100*, 514–530. doi:10.1016/j.jbusres.2018.10.055

Lowry, P. B., D'Arcy, J., Hammer, B., & Moody, G. D. (2016). "Cargo Cult" science in traditional organization and information systems survey research: A case for using nontraditional

methods of data collection, including Mechanical Turk and online panels. *Journal of Strategic Information Systems*, *25*(3), 232–240. doi:10.1016/j.jsis.2016.06.002

Lu, S.-H., Tran, H. T. T., Ngo, T.-S., & Huang, C.-H. (2025). Exploring the evolution and research directions of virtual reality in marketing: An information systems perspective. *Online Information Review*, *49*(3), 534–551. doi:10.1108/OIR-03-2024-0165

Makkonen, M. (2026). Explaining the use intention of low-immersive virtual reality stores: An extension of UTAUT2 with perceived novelty and personal innovativeness. In J. L. Reis, E. Bigne, Z. Bogdanović, & J. P. M. dos Santos (Eds.), *Smart Innovation, Systems and Technologies: Vol. 473. Marketing and Smart Technologies. Proceedings of the International Conference on Marketing and Technologies 2025: Vol. 1* (pp. 789–806). Cham, Switzerland: Springer. doi: 10.1007/978-3-032-16284-7_50

Mills, S., & Noyes, J. (1999). Virtual reality: An overview of user-related design issues. Revised paper for special issue on "Virtual reality: User issues" in Interacting with Computers, May 1998. *Interacting with Computers*, *11*(4), 375–386. doi:10.1016/S0953-5438(98)00057-5

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. doi:10.1016/j.jesp.2017.01.006

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. doi:10.3758/s13428-021-01694-3

Peukert, C., Pfeiffer, J., Meißner, M., Pfeiffer, T., & Weinhardt, C. (2019). Shopping in virtual reality stores: The influence of immersion on system adoption. *Journal of Management Information Systems*, *36*(3), 755–788. doi:10.1080/07421222.2019.1628889

Ringle, C. M., Wende, S., & Becker, J.-M. (2026). *SmartPLS 4*. https://www.smartpls.com

Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The big five personality factors and personal values. *Personality and Social Psychology Bulletin*, *28*(6), 789–801. doi:10.1177/0146167202289008

Ruusunen, N., Hallikainen, H., & Laukkanen, T. (2023). Does imagination compensate for the need for touch in 360-virtual shopping? *International Journal of Information Management*, *70*, 102622. doi:10.1016/j.ijinfomgt.2023.102622

SNS Insider (2025). *Virtual reality in retail market*. https://www.snsinsider.com/reports/virtual-reality-in-retail-market-6007

Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, *86*, 77–90. doi:10.1016/j.chb.2018.04.019

Suh, K. S., & Lee, Y. E. (2005). The effects of virtual reality on consumer learning: An empirical investigation. *MIS Quarterly*, 29(4), 673–697. doi:10.2307/25148705

tom Dieck, M. C., Cranmer, E., Prim, A. L., & Bamford, D. (2023). The effects of augmented reality shopping experiences: Immersion, presence and satisfaction. *Journal of Research in Interactive Marketing*, *17*(6), 940–958. doi:10.1108/JRIM-09-2022-0268

Weibel, D., Wissmath, B., & Mast, F. W. (2010). Immersion in mediated environments: The role of personality traits. *Cyberpsychology, Behavior, and Social Networking*, *13*(3), 251–256. doi:10.1089/cyber.2009.0171

Xi, N., & Hamari, J. (2021). Shopping in virtual reality: A literature review and future agenda. *Journal of Business Research*, *134*, 37–58. doi:10.1016/j.jbusres.2021.04.075

Yoo, K., Welden, R., Hewett, K., & Haenlein, M. (2023). The merchants of meta: A research agenda to understand the future of retailing in the metaverse. *Journal of Retailing*, *99*(2), 173–192. doi:10.1016/j.jretai.2023.02.002

Zhang, Y., Shao, W., Quach, S., Thaichon, P., & Li, Q. (2025). VR retailing: When and why immersion pays off. *Journal of Retailing*, 101(3), 348–365. doi:10.1016/j.jretai.2025.04.004

Appendix A: Item Wordings

Item	Wording
Perceived immersion (PI)	
PI1	I was completely immersed in the virtual store.
PI2	I lost track of time while visiting the virtual store.
PI3	I became very involved in my visit to the virtual store, forgetting about other things.
PI4	I became unaware of my surroundings while visiting the virtual store.
Extraversion (EXT)	
EXT1	I see myself as extroverted, enthusiastic.
EXT2*	I see myself as reserved, quiet.
Agreeableness (AGR)	
AGR1	I see myself as sympathetic, warm.
AGR2*	I see myself as critical, quarrelsome.
Conscientiousness (CON)	
CON1	I see myself as dependable, self-disciplined.
CON2*	I see myself as disorganized, careless.
Emotional stability (ES)	
ES1	I see myself as calm, emotionally stable.
ES2*	I see myself as anxious, easily upset.
Openness to experience (OTE)	
OTE1	I see myself as open to new experiences, complex.
OTE2*	I see myself as conventional, uncreative.

* = reverse-coded item

Appendix B: Logitech and Star Wars Stores

Source: Own

Source: Own

DOI
https://doi.org/10.18690/um.fov.4.2026.60

ISBN
978-961-299-152-4

Expert-Driven Criteria for Assessing SME Readiness for Open Data: Insights from an AI-Supported Brainstorming Workshop

Andrej Škraba,[1] Andreja Pucihar,[1] Staša Blatnik,[1] Nejc Čelik,[1] Franc Lavrič,[1] Radovan Stojanović,[2] Mirjana Kljajić Borštnar[1]

[1] University of Maribor, Faculty of Organizational Sciences, Cybernetics & Decision Support Systems Laboratory, Kranj, Slovenia
andrej.skraba@um.si, andreja.pucihar@um.si, stasa.blatnik1@um.si, nejc.celik1@um.si, franc.lavric@student.um.si, mirjana.kljajic@um.si
[2] University of Montenegro, Faculty of Electrical Engineering, Podgorica, Montenegro
stox@ucg.ac.me

The growing importance of open data ecosystems highlights the need for SMEs to assess their maturity for open data use. This study presents findings from an expert workshop aimed at identifying the factors that drive or hinder SME open data maturity. The workshop, conducted at the 37th Bled eConference, involved 17 experts in a structured brainstorming session supported by the Kresilnik digital tool. A total of 260 ideas were generated through a two-phase process combining traditional ideation and AI-supported suggestion generation. The results indicate that AI assistance can mitigate the typical decline in brainstorming productivity by shifting participant roles toward evaluation and refinement of generated ideas. In addition to methodological insights into human–AI collaborative ideation, the study proposes a structured set of expert-derived criteria for assessing SME maturity for open data. The findings demonstrate the potential of hybrid brainstorming approaches for rapid knowledge generation in emerging digital transformation domains.

Keywords:
open data,
brainstorming,
AI,
LLM,
kresilnik

University of Maribor Press

1 Introduction

Open data (OD) is increasingly recognized as a key enabler of transparency, innovation, and data-driven decision-making across public and private sector ecosystems (Janssen et al., 2012; Page et al., 2023; Zuiderwijk et al., 2014). The concept of OD is commonly defined as data that can be freely used, reused and redistributed by anyone, subject to no proprietary restrictions and with at most requirements for attribution and share-alike (OPSI, 2016). Open data policies have become an important instrument for fostering transparent, accountable, and participatory governance while supporting innovation and economic development (European Commission, 2020; Publications Office of the European Union, 2024; Zuiderwijk et al., 2014). From an economic perspective OD create new opportunities for enterprises by lowering barriers to accessing valuable databases, that would otherwise be costly or unavailable (Jetzek et al., 2019; Magalhaes & Roseira, 2020). Studies estimate that the economic value of OD in the EU reaches billions of euros annually, driven by innovation, efficiency gains and the creation of new products and services (European Commission, 2020). This is particularly important for SMEs, which frequently face resource, capability, and data access constraints, positioning open data as a potentially valuable strategic resource (Janssen et al., 2012).

However, despite its potential the adoption of OD in SMEs remains underutilized. Prior research highlights several barriers among which include problems with OD standardization, quality and usability, security, availability, knowledge management, cooperation with stakeholders, and regulatory (Çaldağ & Gökalp, 2023; Crusoe & Melin, 2018; Ferencek & Kljajić Borštnar, 2025; Huber et al., 2020; Janssen et al., 2012; Wang, 2020; Wieczorkowski, 2019). Jetzek et al. (2019) further explore organizational barriers such as management support and data culture that influences the enterprises' ability to effectively use OD. As a result, understanding and assessing SME maturity to use OD is a critical step toward enabling broader adoption and value creation.

The concept of OD maturity refers to the extent to which organizations possess the capabilities to effectively discover, access, integrate, and exploit OD for value creation. Existing sources suggest that OD maturity is a multidimensional construct encompassing technological, organizational and strategic dimensions (Page et al.,

2023; Zuiderwijk et al., 2015). Maturity includes the technological context, such as availability of appropriate IT infrastructure, data management tools and mechanisms that enable interoperability and integration of external datasets (OECD, 2019). Organizational context, on the other hand, relates to internal processes, employee skills and knowledge management, and its culture. This includes the ability to interpret data, identify relevant use cases and translate insights from data into business value. Management support and strategic alignment are also important factors, as they influence resource allocation and prioritization of OD initiatives (D'Addario et al., 2025; Jetzek et al., 2014; OECD, 2019; Zuiderwijk et al., 2014). Companies that integrate OD into their core strategies are more likely to create new products and services, improve operational efficiency and enhance customer relationships (Magalhaes & Roseira, 2020).

Despite these insights, there is still a lack of standardized and empirically validated methods for assessing OD maturity in SMEs. Most existing models are designed for public sector organizations, which differ significantly from SMEs in terms of resources, capabilities and strategic priorities (European Commission, 2020). This gap highlights the importance of developing context-specific criteria that reflect the unique characteristics and needs of SMEs. To address this gap an expert group was convened to systematically identify and refine relevant criteria for assessing OD maturity in SMEs through a structured brainstorming session supported by the "Kresilnik" tool (Lavrič & Škraba, 2023).

The integration of artificial intelligence (AI) into collaborative ideation processes is redefining traditional brainstorming methodologies. This study investigates the impact of AI-assisted brainstorming within the context of open data maturity for SMEs, as explored during a workshop "Open Data Maturity – Readiness of SMEs" at the 37th Bled eConference in June 2024. The "Kresilnik" tool (Lavrič & Škraba, 2023a) employed at the workshop is a digital platform designed to facilitate structured brainstorming sessions (Lavrič & Škraba, 2023b; Škraba et al., 2025; Škraba et al, 2022).

Recent research increasingly examines how collaboration between humans and AI reshapes teamwork, productivity, and collective problem solving. Field experiments show that integrating AI agents into collaborative work environments can significantly improve team performance and productivity by complementing human

capabilities and facilitating information processing (Ju & Aral, 2025). At the same time, studies of knowledge workers indicate that the use of generative AI may influence cognitive processes during task performance, including reductions in perceived cognitive effort and changes in confidence when evaluating results (Lee et al., 2025). These findings highlight the importance of carefully designing human–AI collaboration so that AI systems augment rather than replace human reasoning and creativity. In group-based problem-solving contexts such as brainstorming, the structure of interaction among participants also plays a critical role, as decentralized collaboration and collective performance have been shown to enhance learning and idea development in multidisciplinary environments (Meluso & Hébert-Dufresne, 2023). Together, these insights provide an important theoretical background (Škraba et al., 2003) for examining hybrid brainstorming processes where human participants interact with AI systems to generate, refine, and evaluate ideas.

This paper analyses idea-generation dynamics in hybrid brainstorming sessions by examining the interaction between human creativity and AI support in the identification of OD maturity criteria. By comparing AI-assisted and traditional brainstorming approaches, the study explores how AI can enhance collaborative ideation for assessing SME open data readiness.

2 Methodology

The overall research goal was to design a multi-criteria decision model for evaluating SME open data maturity. To support the systematic development of the model, we adopted the Design Science Research approach focusing specifically on the exploratory criteria identification phase (Hevner et al., 2004). In the design phase, we applied the DEX (Decision EXpert) methodology, which is suitable for modelling complex multi-criteria decision models using qualitative scales (Bohanec & Rajkovič, 1990). The model development process includes problem definition, criteria identification, hierarchical structuring of criteria, definition of domain values and utility functions, as well as model verification and refinement. Since criteria identification represents a key phase of the modelling process, we first conducted a systematic literature review to identify factors that drive or hinder open data use in SMEs. We then organized an expert workshop to independently identify and validate the criteria against the factors identified in the literature. This paper focuses on that workshop phase.

Eight participants were from Slovenia and nine from other countries. The group included representatives from governmental institutions (2), companies (3), and academic institutions, all with professional or research interests in open data. An additional three participants assisted in conducting the workshop. A structured brainstorming session was conducted in 2024 using the custom-developed "Kresilnik" digital brainstorming tool, which lasted 30 minutes. Participants initially engaged in conventional idea generation. After 1080 seconds, AI support was introduced through an AI-assisted interface that enabled participants to generate, edit, and submit AI-generated ideas they considered valuable. During the 30-minute session, 17 participants collectively generated 260 ideas. Figure 1 presents the initial brainstorming question together with the prompt used to initiate the ideation process and its tokenized representation generated by the OpenAI tokenizer. Each coloured segment represents an individual token identified by the tokenizer, illustrating how the natural language prompt is decomposed into smaller computational units that are processed by the AI model. This visualization highlights the underlying mechanism through which the AI system interprets the brainstorming prompt before generating suggestions or supporting idea generation during the session.

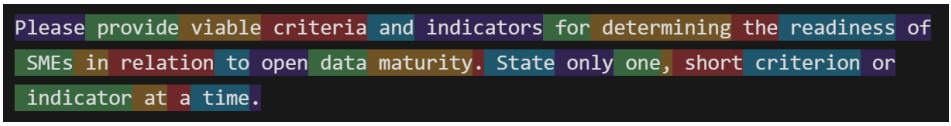

Figure 1: Tokenized representation of the initial brainstorming prompt
Source: Created by the authors

The brainstorming session was divided into two distinct phases: Phase 1) A traditional brainstorming format and Phase 2) An enhanced mode incorporating real-time support from artificial intelligence. Within the session the following activities were conducted: a) Topic presentation, b) Collecting of ideas (in the first part of the session, the session followed a classical format, while in the second part, it was enhanced with the application of artificial intelligence), c) Categorization, d) Voting about the importance of the categories, and e) Voting about the importance of ideas in the first category.

Figure 2 presents an extended hybrid brainstorming workflow that integrates AI into the classical brainstorming process and expands the concept proposed by Lavrič and Škraba (2023a). After the initial call for ideas, the participants start with ideation process. If human ideas are available, they proceed directly into the traditional group brainstorming flow. When no ideas are produced, the participants can activate an AI-assisted pathway that generates ideas using OpenAI. Important technical addition is a database of previously generated AI ideas which reduces the redundancy of generated ideas by AI. The generated ideas are then reviewed by participants, who may check, modify, or correct them before evaluating whether they are unique and appropriate. Approved ideas are subsequently introduced into the group brainstorming process, while non-unique or inappropriate ideas trigger a new AI generation cycle. The key extension compared to the earlier model is the introduction of the AI-generated ideas database, which enables reuse, accumulation, and refinement of ideas across brainstorming sessions.

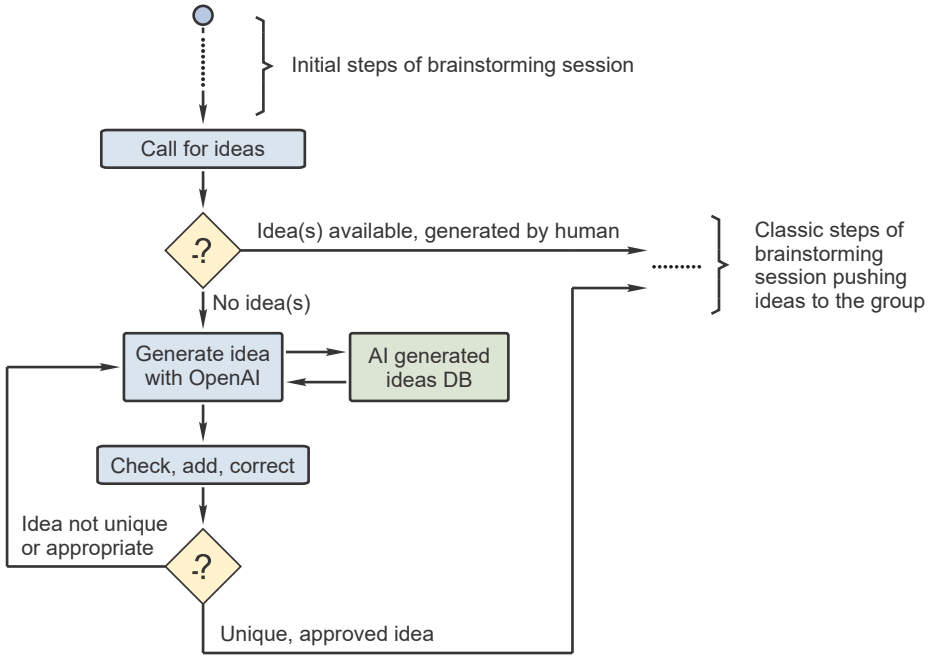

Figure 2: Extended hybrid brainstorming process integrating AI-generated ideas and an AI idea database

Source: Created by the authors

3 Results

A total of 30,759 interaction logs were recorded in a millisecond-resolution JSON database, capturing user input (keystrokes etc.) and interface interactions. These data form the foundation for analysing the dynamics and intensity of participation throughout the session.

Notably, among the 260 generated ideas, the one ranked third—'Use of open data to create innovative products or services addressing specific market needs'—was generated by the AI system, demonstrating that AI-supported idea generation can produce highly relevant contributions that are evaluated by expert groups as comparable to, and competitive with, human-generated ideas in the context of assessing SME open data readiness.

Figure 3 shows a word cloud generated from the collected ideas, highlighting the most frequent terms and concepts identified by the experts. Here the following stopwords were used at the generation of the word cloud: **`open data|SME|SMEs|within|s`**.

Figure 3: Word cloud of the Bled2024 brainstorming session
Source: Created by the authors

To gain a deeper understanding of participant engagement during the brainstorming session, we generated a visualization based on precisely timestamped keyboard interactions. Each interaction-such as a keystroke or idea entry-was recorded with millisecond accuracy and represented as a vertical line on a time axis. This approach captures the rhythm and intensity of the creative process in a way that is both intuitive and revealing.

The resulting Figure 4 can be seen as a spectral fingerprint of the brainstorming session-an abstract, high-resolution imprint of cognitive and collaborative energy unfolding over time. Time is mapped along the horizontal axis, covering the full 30-minute session (1800 seconds), while the vertical axis is used purely for plotting purposes, with each black line extending from 0 to 1. Every vertical line marks a precise moment of user interaction, and the collective density of these lines reflects the dynamic flow of ideation and engagement.

The visualization distinctly reveals two behavioural phases. In the first phase-spanning approximately the first 1080 seconds- Figure 4 displays a dense cluster of vertical lines, indicative of frequent and sustained keyboard input. This suggests a period of heightened cognitive effort and active idea generation, as participants were fully immersed in the creative process. The second phase begins after the 1080-second mark, coinciding with the introduction of AI assistance. From this point on, a visible reduction in line density occurs, signalling a decline in manual input. This shift may reflect participants transitioning from active ideators to evaluators, curators, or facilitators of AI-generated suggestions. It could also suggest the onset of cognitive fatigue, changes in group dynamics, or increased reliance on the AI's output.

Overall, this visual encoding of timestamped activity serves as an elegant and powerful method for uncovering the hidden temporal patterns of group brainstorming. It not only highlights key transitions in engagement but also provides a compact, interpretable snapshot of the human-machine creative interplay as it evolved during the session.

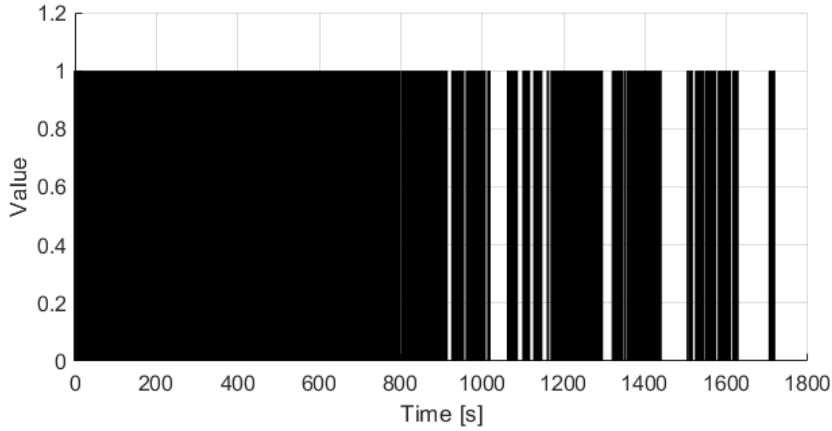

Figure 4: Spectral fingerprint of the brainstorming session
Source: Created by the authors

These results provide an indicator, that the AI support introduced at 1080s fundamentally changed the group's ideation dynamics, not just in level but in direction of growth.

Table 1 presents the final thematic categories derived from the expert brainstorming workshop together with their ranking scores assigned by participants. The categories were obtained through clustering and consolidation of the 260 generated ideas into broader conceptual dimensions representing SME readiness for open data. The ranking results indicate that experts considered business-related and data-related aspects to be the most critical determinants of SME readiness, while legal and regulatory aspects received comparatively lower, though still relevant, importance scores. The relatively low coefficient of variation (CV) values across most categories indicate a moderate to high level of agreement among participants regarding the importance of the identified dimensions.

Table 1: Ranked categories for assessing SME readiness for open data

Rank	Category	Example Criteria	Mean Score	Std. Dev.	CV
1	Business Potential	Value, innovation, competitiveness	8.41	1.68	0.2
2	Data Quality	Governance, metadata, consistency	8.12	1.41	0.17
3	Capabilities	Analytics, expertise, visualization	7.59	2.12	0.28
4	Innovation	Culture, agility, prototyping	7.35	2.08	0.28
5	IT Infrastructure	Platforms, dashboards, cybersecurity	6.82	2.15	0.32
6	Internal Strategy	Leadership, strategy, alignment	6.82	1.89	0.28
7	Business Environment	Ecosystems, funding, collaboration	6.59	2	0.3
8	Data Integration	Repositories, accessibility, interoperability	6.59	2	0.3
9	Training and Education	Literacy, workshops, learning	6.29	2.44	0.39
10	Regulations and Legal Compliance	GDPR, compliance, privacy	5.88	1.91	0.32

The results suggest that experts perceive SME readiness for open data primarily as a strategic and business-oriented capability rather than solely a technological issue. The highest-ranked category, Business Potential, emphasizes that SMEs must recognize and operationalize the business value of open data through its integration into products, services, and decision-making processes. The detailed ranking within this category further reinforces this interpretation, as the two highest-ranked ideas, "see value in open data" and "reliability of open data sources", both achieved the highest average score (8.50), highlighting that perceived business value and trust in data quality are considered fundamental prerequisites for successful open data adoption. Particularly noteworthy is the third-ranked idea, "Use of open data to create innovative products or services that address specific market needs", which was generated by the AI-supported brainstorming system and achieved an average score of 8.42. Its high ranking demonstrates that AI-generated contributions can be evaluated by experts as highly relevant and comparable to human-generated ideas in the context of open data readiness assessment. At the same time, the slightly higher coefficient of variation for this idea suggests somewhat greater disagreement among

participants regarding the role of innovation-driven value creation as an immediate readiness factor.

The second-ranked category, Data Quality, further highlights the importance of governance, metadata quality, consistency, accessibility, and reliability of open data sources as prerequisites for effective data utilization within SMEs. Similarly, the highly ranked Capabilities category emphasizes the importance of analytical competencies, data literacy, visualization skills, and the ability to derive actionable knowledge from data. The categories Innovation and Internal Strategy additionally underline the importance of organizational culture, agility, leadership support, and strategic alignment in fostering open data adoption. Interestingly, more purely technical dimensions such as IT Infrastructure and Data Integration received lower rankings than business and strategic factors, indicating that experts perceive organizational readiness and value creation as more important determinants of open data maturity than technology alone. Overall, the findings support the view that SME readiness for open data is a multidimensional construct combining technological, organizational, strategic, cultural, and governance-related dimensions.

Figure 5 shows the evolution of average idea length over time during the 30-minute brainstorming session, with time divided into equal bins (60 seconds) along the horizontal axis and average idea length displayed on the vertical axis. The dashed red vertical line at 1080 seconds marks the point at which AI assistance was introduced, separating the session into a pre-AI phase (blue bars) and an AI-supported phase (orange bars). In the initial phase, idea length remains relatively moderate and variable, generally ranging between about 25 and 70 characters with occasional peaks, reflecting concise, human-driven idea generation. After the introduction of AI, a clear shift is observed, as average idea length increases substantially and stabilizes at higher values, typically exceeding 100 characters and reaching up to around 140–145 characters. This indicates that ideas in the AI-supported phase are more elaborated and detailed, suggesting a transition from rapid idea generation toward more reflective refinement and expansion of ideas, likely influenced by AI-generated suggestions.

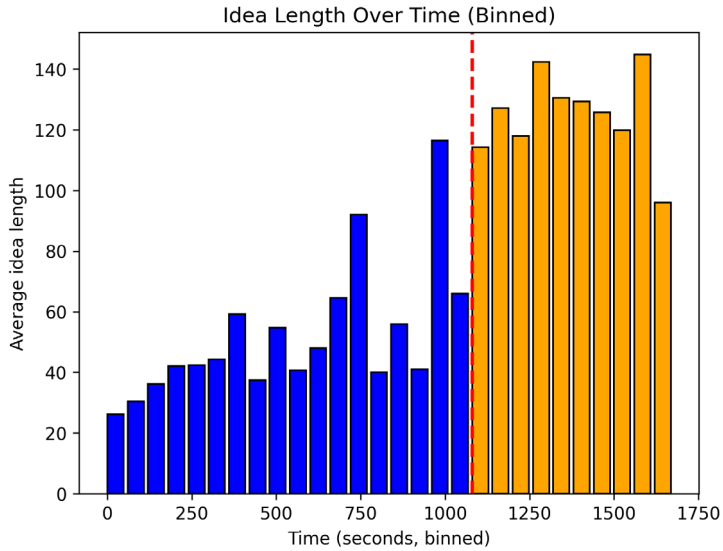

Figure 5: Average idea length (characters) before and after introduction of AI

Table 2 presents the ten most semantically unique ideas generated after the introduction of AI support, identified based on their lowest cosine similarity scores to any idea produced in the pre-AI phase. Idea similarity was computed using sentence embeddings generated with the SentenceTransformer model (Hugging Face Community, 2026), where lower similarity scores indicate greater semantic distance and therefore possible higher novelty. For each idea, the table lists its closest matching counterpart from the pre-AI phase along with the corresponding similarity score. The similarity scores indicate varying degrees of novelty among ideas generated in the AI-supported phase. A small number of ideas exhibit very low similarity scores (below 0.3), representing strong semantic divergence from previously generated ideas; however, these correspond to outliers or contextually irrelevant contributions by participants (e.g., "good coffee"), suggesting that low similarity alone does not guarantee conceptual relevance. More substantively, several ideas with similarity scores between 0.3 and 0.45 introduce genuinely new thematic dimensions, such as risk assessment and compliance, indicating meaningful expansion of the problem space; based on their style, phrasing, and level of contextual grounding, these ideas were most likely contributed by human participants rather than generated by the AI system. Ideas generated with support of AI fall within the range with similarity score of 0.5 and above, reflecting moderate

novelty mostly in the form of elaborations or refinements of existing concepts rather than entirely new directions. Overall, the results suggest that the AI-supported phase contributed to both the emergence of new conceptual areas and the extension of previously identified themes, highlighting a combination of radical and incremental innovation. Based on stylistic characteristics, it is likely that most, if not all, human-generated ideas produced after the introduction of AI are represented among the top ten most unique contributions, indicating that human participants remained the primary source of conceptual novelty in the later phase of the session.

Table 2: Top 10 most unique ideas after AI introduction by similarity score to any idea from before AI introduction

Idea	Most similar	Similarity score
;-- SELECT * FROM users WHERE 1;	Easy to find data	0.230109721
Good coffee	Strategy	0.240747333
Doing a proper risk assessment	Use of the data to aid decision making	0.362663835
Understanding the challenge of compliance	Understanding the regulation governing open data	0.454871476
Culture - employees coming from different multicultural background - how do they accept changes?	Organizational culture	0.509587228
Clear communication channels and collaborative platforms are in place to facilitate seamless data sharing and collaboration across teams. --> I agree with this AI :-)	Taking part in data sharing ecosystems	0.513242185
Digital capability (such as technology use - especially data analytics, investment in IT department, what data (or digital) strategy they have as an organisation)	Technological capabilities	0.535972476
Evaluation of the SME's capability to leverage machine learning algorithms for extracting actionable insights from large open data sets.	Does the model/ prediction generated from open data have impact on real life business	0.537490427
Engagement in continuous learning and professional development programs focused on the latest advancements in open data technology and methodologies.	Open data being sufficient and insightful to the development process	0.541098237
Development of a disaster recovery plan to ensure continuity of open data initiatives in the event of a system failure or data breach.	Reliability of open data sources	0.554412842

5 Limitations

A limitation of present study is that the findings are based on a single expert workshop conducted within the context of the Bled 2024 eConference, which limits the generalizability of the results. Although the workshop involved participants from different countries and professional backgrounds, the sample size remained relatively small and may not fully represent the broader population of SMEs, policymakers, or open data practitioners. The study should therefore be considered exploratory in nature, primarily intended to identify and structure potentially relevant criteria for assessing SME readiness for open data. In addition, the brainstorming outcomes may have been influenced by the specific expertise, experiences, and perspectives of the participating experts, introducing a degree of expert-selection bias. While the integration of AI support contributed to increased ideation and idea elaboration, the quality and relevance of AI-generated suggestions were not independently validated outside the workshop context. Furthermore, the proposed criteria structure has not yet undergone longitudinal or large-scale empirical validation in real organizational settings. Future research should therefore include broader and more diverse participant groups, repeated workshops, and empirical testing of the resulting DEX-based assessment model across different SME environments and industries.

6 Discussion and Conclusion

The results of this study offer compelling evidence that the integration of artificial intelligence (AI) into group brainstorming sessions can significantly reshape both the trajectory and intensity of ideation processes. Through timestamped interaction data, sliding-window activity analysis, and linear trend comparisons, a nuanced and paradoxical behavioural pattern emerges—one that challenges traditional assumptions about group creativity over time.

Before the AI intervention, the group followed a classical brainstorming dynamic: high initial engagement, followed by a gradual decline in both interaction intensity and idea generation, reflecting expected cognitive fatigue and diminishing returns.

The integration of AI into group ideation appears to reverse the typical pattern of creative decline. It compensates for physical fatigue, sustains and even amplifies ideational output, and reconfigures participants' roles from individual contributors to collaborative co-curators. The findings suggest that AI has the potential to reshape collaborative brainstorming processes by supporting ideation continuity and refinement.

This study addresses the research gap related to the lack of standardized and empirically validated approaches for assessing OD maturity in SMEs by focusing on the identification of relevant evaluation criteria. Through an expert-driven brainstorming workshop using the "Kresilnik" tool, a comprehensive set of criteria reflecting technological, organizational and strategic dimensions of OD maturity was generated. It highlights that SME maturity to use open data extends beyond technical infrastructure and includes critical organizational factors such as data culture, skills, management support, and the ability to translate data into business value, which corroborates the findings of the systematic literature review (Blatnik et al., 2024). A smaller group of experts used the generated criteria for consequent modelling phases (Kljajić Borštnar et al., 2026). The findings demonstrate that the integration of AI into the ideation process significantly alters brainstorming dynamics, shifting participants from primary idea generators to evaluators and curators, while sustaining productivity and enhancing idea elaboration. The hybrid human–AI approach proved effective in rapidly generating diverse and structured knowledge in a complex and emerging domain. Overall, this study contributes both methodologically, by demonstrating the potential of AI-supported collaborative ideation, and substantively, by providing a foundation for the development of SME-specific OD maturity assessment model.

Acknowledgement

The authors acknowledge the financial support from the Slovenian Research and Innovation Agency (research core funding No. P5-0018) and the Ministry of Higher Education, Science, and Innovation of the Republic of Slovenia as part of the Next Generation EU National Recovery and Resilience Plan (Grants No. 3330-22-3515; NOO No: C3330-22-953012).

References

Blatnik, S., Pucihar, A., & Borštnar, M. K. (2024). Spodbude in ovire uporabe odprtih podatkov v podjetjih—Pregled literature. Uporabna informatika, 32(4). https://doi.org/10.31449/upinf.238

Bohanec, M., & Rajkovič, V. (1990). DEX: An Expert System Shell for Decision Support. Sistemica, 1(1), 145–157.

Çaldağ, M. T., & Gökalp, E. (2023). Understanding barriers affecting the adoption and usage of open access data in the context of organizations. Data and Information Management, 100049. https://doi.org/10.1016/j.dim.2023.100049

Crusoe, J., & Melin, U. (2018). Investigating Open Government Data Barriers: A Literature Review and Conceptualization. In P. Parycek, O. Glassey, M. Janssen, H. J. Scholl, E. Tambouris, E. Kalampokis, & S. Virkar (Eds), Electronic Government (Vol. 11020, pp. 169–183). Springer International Publishing. https://doi.org/10.1007/978-3-319-98690-6_15

D'Addario, J., Tarrant, D., Temple, D., Thuermer, G., Corby, T., Newman, A., Lee, T., Evans, E., de Marco, O., Lynch, L., Dinnage, D., & Pieroni, T. (2025). Open Data Maturity Model Guide (2.0). Open Data Institute.

e2303568120.OECD. (2019). Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies. OECD Publishing. https://doi.org/10.1787/276aaca8-en

European Commission. (2020). The European data market monitoring tool: Key facts & figures, first policy conclusions, data landscape and quantified stories : d2.9 final study report Executive summary. Directorate General for Communications Networks, Content and Technology. https://data.europa.eu/doi/10.2759/837980

Ferencek, A., & Kljajić Borštnar, M. (2025). Open Government Data Topic Modeling and Taxonomy Development. Systems, 13(4). https://doi.org/10.3390/systems13040242

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. Management Information Systems Quarterly, 28(1), 75–106. https://doi.org/10.2307/25148625

Huber, F., Wainwright, T., & Rentocchini, F. (2020). Open data for open innovation: Managing absorptive capacity in SMEs. R&D Management, 50(1), 31–46. https://doi.org/10.1111/radm.12347

Hugging Face Community. (2026, March 20). Sentence-transformers. All-MiniLM-L6-V2. https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

Ju, H., & Aral, S. (2025). Collaborating with ai agents: Field experiments on teamwork, productivity, and performance. arXiv preprint arXiv:2503.18238

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management, 2012(29), 258–268. https://doi.org/10.1080/10580530.2012.716740

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2014). Data-Driven Innovation through Open Government Data. Journal of Theoretical and Applied Electronic Commerce Research, 9(2), 15–16. https://doi.org/10.4067/S0718-18762014000200008

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2019). The Sustainable Value of Open Government Data. Journal of the Association for Information Systems, Copenhagen Business School, 702–734. https://doi.org/10.17705/1jais.00549

Kljajić Borštnar, M. K., Rajkovič, U., Blatnik, S., Marolt, M., Lenart, G., & Pucihar, A. (2026). Metodologija za merjenje zrelosti za uporabo odprtih podatkov v podjetjih. In Univerzitetna založba Univerze v Mariboru. Univerzitetna založba Univerze v Mariboru. https://doi.org/10.18690/um.fov.2.2026

Magalhaes, G., & Roseira, C. (2020). Open government data and the private sector: An empirical view on business models and value creation. Government Information Quarterly, 37(3), 101248. https://doi.org/10.1016/j.giq.2017.08.004

Lavrič, F., & Škraba, A. (2023a). Brainstorming will never be the same again—a human group supported by artificial intelligence. Machine Learning and Knowledge Extraction, 5(4), 1282-1301.

Lavrič, F., & Škraba, A. (2023b, March). Group brainstorming support by chatgpt & ayoa at the design of regional development plan. In Proceedings of the 42th International Conference on Organizational Science Development, Portorož, Slovenia (pp. 22-24).

Lavrič, F., Semenkin, E., Stanovov, V., & Škraba, A. (2022, March). Review of tools and methodology for efficient group idea generation and decision making. In Proceedings of the 41st International Conference on Organizational Science Development (pp. 489-501). Maribor, Slovenia: University of Maribor Press.

Lee, H. P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025, April). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In Proceedings of the 2025 CHI conference on human factors in computing systems (pp. 1-22).

Meluso, J., & Hébert-Dufresne, L. (2023). Multidisciplinary learning through collective performance favors decentralization. Proceedings of the National Academy of Sciences, 120(34).

OPSI. (2016). OPSI [Online post]. Kaj so Odprti Podatki? https://podatki.gov.si/posredovanje-podatkov/kaj-so-odprti-podatki

Page, M., Hajduk, E., Lincklaen Arriens, E., Cecconi, G., & Brinkhuis, S. (2023). Open data maturity report 2023 (Vol. 2023). Publications Office of the EU. https://data.europa.eu/doi/10.2830/384422

Publications Office of the European Union. (2024). Results of the European Data Market study 2021-2023 | Shaping Europe's digital future. Https://Digital-Strategy.Ec.Europa.Eu/En/Library/Results-European-Data-Market-Study-2021-2023. https://doi.org/10.2759/632809

Škraba, A., Kljajić, M., & Leskovar, R. (2003). Group exploration of system dynamics models—Is there a place for a feedback loop in the decision process? System Dynamics Review, 19(3), 243–263. https://doi.org/10.1002/sdr.274

Wang, H. J. (2020). Adoption of Open government data: Perspectives of user innovators. Information Research, 2020(25 (1)). https://web.archive.org/web/20200224193543/http://informationr.net/ir/25-1/paper849.html

Wieczorkowski, J. (2019). Open Data as a Source of Product and Organizational Innovations. Proceedings of the 14th European Conference on Innovation and Entrepreneurship ECIE 2019, 2019, 1118–1127. https://doi.org/10.34190/ECIE.19.190

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2015). Socio-technical Impediments of Open Data. Electronic Journal of E-Government, 10(2).

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. Information Polity, 19(1,2), 17–33. https://doi.org/10.3233/IP-140329

Appendix

Figure 6 shows the mechanism used for dynamic prompt construction in iterative AI-supported brainstorming. During the first API call, the prompt consists only of the initial brainstorming question (`initialQuestion`) and additional instructions (`promptExtension`) that guide the model to generate a single concise idea.

For subsequent API calls, the prompt is extended by appending a list of previously generated ideas (`previouslyGeneratedIdeas`). These ideas are introduced in the prompt with the instruction "Please note that the following ideas have already been suggested:". This mechanism provides contextual information to the model and encourages the generation of novel, non-redundant ideas in later iterations. As a result, the system supports a structured and iterative AI-assisted brainstorming process.

```
// Prompt construction for iterative AI-supported brainstorming
let promptString = "";

if (numberOfChatGPTCalls === 1) {
    promptString =
        initialQuestion +
        promptExtension;
} else {
    promptString =
        initialQuestion +
        promptExtension +
        " Please note that the following ideas have already been suggested: " +
        previouslyGeneratedIdeas;
}
```

Figure 6: Dynamic prompt construction for iterative AI-supported brainstorming, where previously generated ideas are appended to the prompt to encourage the generation of novel and non-redundant ideas

RESEARCH IN PROGRESS

DESIGN AND EVALUATION OF A DIGITAL TRANSFORMATION CANVAS FOR SME

LUKAS NEIDLINGER, JÜRGEN JUNG

Frankfurt University of Applied Sciences, Faculty of Computer Science and Engineering, Hesse, Germany
lukas.neidlinger@stud.fra-uas.de, juergen.jung@fra-uas.de

Due to the rapid evolution of digital technologies and associated business benefits, companies are forced to adapt their business models and operational processes. However, especially small and medium-sized companies perceive a significant risk imposed by a digital transformation as they lack digital capabilities and necessary resources. Overwhelmed by heavy-weight frameworks and methods, such companies rather need a simple approach like a business model canvas for digital transformation. The objective of the research presented in the paper at hand is to present reesarch in progress on the design and validation of a digital transformation canvas for small and medium-sized enterprises. It has been developed together with the Chamber of Industry and Commerce in Offenbach (Germany) as well as two member companies.

DOI
https://doi.org/10.18690/um.fov.4.2026.61

ISBN
978-961-299-152-4

Keywords:
digital transformation, canvas, SME, digitalization, business model

University of Maribor Press

1 Motivation and Research Objective

Digital Transformation (DT) is widely perceived as a business transformation enabled by modern digital technology (Morakanyane et al., 2017). Digital technology is an enabler for business model innovation (Ancillai et al., 2023) or the adaption of existing business models with digital services (Tian et al., 2022). Typical drivers for a DT are expectations of digital customers (Verhoef et al., 2021), innovative (or disruptive) digital technologies (Ancillai et al., 2023) or digital ecosystems (Ancillai et al., 2023). Furthermore, DT can be used for improving existing business processes or even achieve operational excellence (Draganic, 2023).

However, due to its holistic impact on an organisation, DT still remains a challenging task even though dedicated methods and frameworks are available (Ciruskabiri & Varnaseri, 2023; Torres et al., 2024). Especially Small and Medium-sized Enterprises (SME) struggle with such an endeavour, as digitalization is not their core competency and they lack of necessary resources and digital capabilities (Pelletier & Cloutier, 2019) and existing canvases are only partially suitable (Neidlinger, 2025). In order to provide a tangible starting point for a DT in SME, the objective of the paper at hand is to present the design and evaluation of a corresponding canvas. A Design Science Research (DSR) approach (Hevner et al., 2004) has been chosen as the research aims to develop an artefact that can be applied in a business context. A first draft canvas was derived based on results from a Structured Literature Review (SLR) on existing canvases, digital transformation as well as peculiarities of SME. This draft has been evaluated in an interview with an expert of the Industrie- und Handelskammer (IHK)[1] in the city of Offenbach (Germany) and then validated by conducting two scenario workshops with selected SME.

2 Overview on Existing Research

A couple of candidates for DT canvases for SME have been identified based on the SLR as documented in (Neidlinger, 2025). The databases of Springer, Elsevier and IEEE have been searched resulting in 140 publications. These have been filtered by excluding very specific topics or other contexts and including those with a holistic

[1] Chamber of Industry and Commerce in Offenbach, Germany

perspective for DT. Candidate fields have been identified and related to SME requirements based on the final 25 papers.

The Business Model Canvas (BMC) by Osterwalder and Pigneur (Osterwalder & Pigneur, 2013) aims to support shaping a (new) business model (Osterwalder & Pigneur, 2013). Even though it does not directly address a DT (Bachmann et al., 2025), it offers a validated base to build upon. Dedicated DT canvases in particular focus on employees (e.g. Moencks et al., 2022) and the usage of data as a driver for DT (e.g. Elia et al., 2024). Both points are relevant for SME having a limited number of employees and difficulties in recognizing potential for successfull usage of data in their business processes. Nevertheless, these canvases are not fully suited as comprehensive DT canvases, as they focus on their specific subject areas while other essential aspects of DT are not included. Security and safety aspects are emphasised as relevant topics for SME (e.g. Azinheira et al., 2023; Jäntti, 2020) but not adequatly recognised in existing DT canvases except for one: (Elia et al., 2024). A considerable amount of canvases follow a business focused approach, aiming to display most aspects of a business model on a high level of abstraction (e.g. Elia et al., 2024). Additionally, there are human-centered canvases (e.g. Moencks et al., 2022), moving the employee into the focus to display its relevance in DT, while others are technology-focused (e.g. Andre et al., 2018; May et al., 2024) or guiding the introduction of key drivers for DT.

3 DT Canvas for SME

The structure of the canvas, as shown in Figure 1, is based on Osterwalder and Pigneur's BMC (Osterwalder & Pigneur, 2013). *Employee* is positioned to the left of *Value Creation* to emphasize its role as the core of value creation in the company. New elements *Data* and *Digital Technologies* are inserted into the center of the canvas, to reflect their relevance for processes of *Value Creation* as well as *Delivery*. *Security & Regulations* is placed at the top to emphasize its impact on other elements. The canvas contains two types of arrows to indicate dependencies. Large arrows follow the flow of value creation from within the company on the left towards the customer on the right. Small arrows (with labels) describe the interconnectedness between canvas fields. The following explains the fields and their relations to each other in depth.

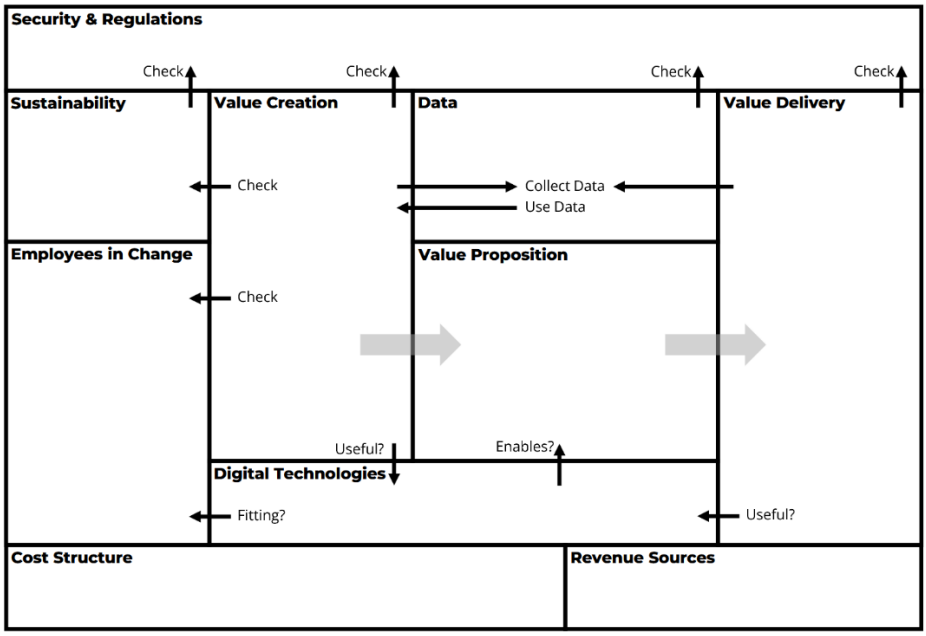

Figure 1: DT canvas overview
Source: (Neidlinger, 2025), p. 55

The field *Data* encourages to think about what data can be collected from *Value Creation* as well as *Value Delivery* (arrow labeled *collect data*). For *Value Creation*, this may encompass data from past activities or machines (Gualtieri et al., 2020), while on the *Value Delivery* side, data about consumer behavior or demographics could be gathered (Enshassi et al., 2025). Encouraged by the *use data* arrow, users should further reflect on how data can be used for value creation, like for example predictive maintenance (Khan et al., 2022), or gather more customer data for better tailored marketing.

Security and Regulations is located above *Data* to challenge data security and check for relevant regulations. Attention has also been paid to implementing measures to prevent or mitigate security risks, like introducing regular IT-security trainings (Le et al., 2024). The field should also address physical security hazards that can occur during a digitalization process, for example through the introduction of robots in manufacturing (Cagno et al., 2024; Romero et al., 2019).

Digital Technologies serves as a collection of technologies currently in use with the objective to make two initial considerations: Does technology help with the current *Value Creation* or Value *Delivery*? Or does it enable new ways of *Value Creation* and potentially a new *Value Proposition*? SMEs should further consider the fit of the technology to the employees, because in the end they are the ones using it (Fischer et al., 2019). This is emphasized by the proximity of both fields.

Employees are defined by Osterwalder and Pigneur to be part of a company's *Value Creation*. However, the field *Employees in Change* does not refer to employees as individual resources, but depicts how employees, being a key driver of value creation in SMEs, are affected by change. This field shows how an employee's work environment is changed through digitalization, how they react to change and what measures can be taken to prepare them or inspire them for DT.

Sustainability addresses both ecological and economic sustainability of the business model. It makes the user consider whether planned changes may strengthen or weaken the organization's ecological footprint. Further, it encourages reflection on how the adapted business model can remain competitive in the long term.

The DT Canvas combines the results of the SLR and SME requirements and, therefore, does not focus on one specific level of abstraction but perceived usefulness of SME. It enables them to ideate business models at a strategic level while still addressing crucial operational aspects, such as changes to employees' work environments. This mix of abstraction is the result from SME requirements identified in the SLR as well as the expert interview and scenario workshops. However, future research is necessary to confirm its usefulness.

The canvas is supposed to be filled in by business stakeholders together with a facilitator who supports adding contents and keeps it consistent. A questionnaire has been developed to guide such a facilitator through the process of filling in the DT Canvas. This is accomplished by simple questions for the user of the canvas to be answered in order to complete it. Details on the questionnaire and its application can be found in (Neidlinger, 2025).

4 DT Canvas Evaluation

An online interview has been conducted with an expert of the Chamber of Industry and Commerce of Offenbach on July 8, 2025 using Microsoft Teams. Even though the feedback was generally encouraging, the expert raised concerns with respect to the usability and comprehensiveness of the canvas for SMEs in particular. As a consequence, the facilitator's guide and the guiding arrows as explained in the previous section have been introduced to improve user understanding.

The DT Canvas was then evaluated in two consecutive workshops with SME. There were no explicit sampling criteria as only two companies were willing to support the evaluation. They were chosen to get early feedback as recommended in (Yin, 2007). Results from the workshop were noted ex post by the researcher based on observations of the behavior, issues and feedback. Topics have been identified with respect to usability (e.g. canvas field size), user understanding (e.g. field names) and flexibility (e.g. different usage scenarios).

We describe the workshop to provide an example on how it can be applied with a facilitator in a SME: During the first workshop, a blank A0 print of the canvas was used at the company's premises. After a short overview of the canvas, post-its where used to map the "AS-IS" business model of the company by addressing each field according to the order defined in the guiding questionnaire. As an exception, the field *Employees in Change* was left empty as no specific change occurred, yet. The resulting canvas was then used as a base to discuss possible scenarios for digitalization or DT. With the help of the guiding arrows, opportunities for leveraging data from the side of the customer or the value creation were explored, as well as the potential use of new technologies and the implications of these changes for security and employees. These scenarios were added to the canvas by using sticky notes having different colors. This workshop has also shown that clearer naming of canvas elements was needed as the practitioner experienced some confusion. Thus, the initial fields *Employees* and *Technologies* where renamed to *Employees in Change* and *Digital Technologies* respectively. Furthermore, the version of the canvas used in this workshop had an additional field labelled *Competitors*. The field was removed to allow more focus on core elements and speed up the process of filling out the canvas, as this took a considerable amount of time. Too much focus on competitors may also

lead to an inadequate business model, for example by copying a business model or technology without sufficient resources or capabilities (C. Matt et al., 2015).

The second workshop was conducted online. While no direct changes to the canvas were made, the current structure and layout of the canvas was confirmed with respect to understanding, usability and flexibility. However, the observation was made that consultation of additional employees during the process of filling in the DT Canvas may lead to better results, especially since employees are a core topic in the canvas.

5 Summary and Outlook

The paper at hand described the design and evaluation of a canvas addressing a DT in SME. A first version of the canvas has been derived from results of a SLR and then evaluated by an expert from a Chamber representing industry and commerce firms in the Offenbach area (Germany). This interview already implied changes to the canvas and the practicability check continued with two scenario workshops conducted with selected SMEs. Experiences from the evaluations lead to changes in the canvas' structure as well as the creation of a guiding document explaining the application of the canvas through dedicated questions.

Future research will have to be done in order to improve the canvas. As already discussed in section 4, further workshops should be conducted with additional employees as it was not in the scope of the evaluation. Interviewing additional experts will lead to more diverse insights and better validity of results. Especially since the second workshop was conducted online, another on-site test could lead to additional insights in regard to usability of the printed-out canvas. To further assess user acceptance of the DT canvas, future research will incorporate the Technology Acceptance Model (TAM) as an evaluation framework (Davis & Granić, 2024). Another limitation, which is common in Design Science Research, is that the researcher also acted as the evaluator of the developed artifact (Hevner et al., 2004). This may introduce some bias into the evaluation process, even though care was taken to interpret evaluation as objectively as possible. Thus, future evaluation by several researchers is recommended.

References

Ancillai, C., Sabatini, A., Gatti, M., & Perna, A. (2023). Digital technology and business model innovation: A systematic literature review and future research agenda. *Technological Forecasting & Social Change*, *188*.

Andre, L., Michael, B., Daniel, R., & Christian, K. (2018). Framework for the identification and demand-orientated classification of digital technologies. *2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)*, 31–36. https://doi.org/10.1109/ITMC.2018.8691135

Azinheira, B., Antunes, M., Maximiano, M., & Gomes, R. (2023). A methodology for mapping cybersecurity standards into governance guidelines for SME in Portugal. *Procedia Computer Science*, *219*, 121–128. https://doi.org/10.1016/j.procs.2023.01.272

Bachmann, N., Thienemann, A.-K., Tüzün, A., Brunner, M., Tripathi, S., Pöchtrager, S., & Jodlbauer, H. (2025). The Evolution of the Business Model Canvas for Digital Sustainability. *Procedia Computer Science*, *253*, 1012–1023. https://doi.org/10.1016/j.procs.2025.01.163

Cagno, E., Accordini, D., Neri, A., Negri, E., & Macchi, M. (2024). Digital solutions for workplace safety: An empirical study on their adoption in Italian metalworking SMEs. *Safety Science*, *177*, 106598. https://doi.org/10.1016/j.ssci.2024.106598

Ciruskabiri, S., & Varnaseri, A. (2023). Digital Transformation Planning and Frameworks. *International Journal of Digital Content Management (IJDCM)*, *4*(7), 397–433.

Davis, F. D., & Granić, A. (2024). *The Technology Acceptance Model: 30 Years of TAM*. Springer International Publishing. https://doi.org/10.1007/978-3-030-45274-2

Draganic, M. (2023). Managing the Company's Digital Capability: A Case for Operational Excellence. *Journal of Information and Organizational Sciences*, *47*(1), 223–243.

Elia, G., Solazzo, G., Lerro, A., Pigni, F., & Tucci, C. L. (2024). The digital transformation canvas: A conceptual framework for leading the digital transformation process. *Business Horizons*, *67*(4), 381–398. https://doi.org/10.1016/j.bushor.2024.03.007

Enshassi, M., Nathan, R. J., Soekmawati, & Ismail, H. (2025). Unveiling barriers and drivers of AI adoption for digital marketing in Malaysian SMEs. *Journal of Open Innovation: Technology, Market, and Complexity*, *11*(2), 100519. https://doi.org/10.1016/j.joitmc.2025.100519

Fischer, H., Rittmeier, F., Strothmann, T., & Schwenniger, N. (2019). Partizipation von Beschäftigten in der Gestaltung einer digitalisierten Arbeitswelt 4.0 mittels einer Canvas-Methode. In C. K. Bosse & K. J. Zink (Eds), *Arbeit 4.0 im Mittelstand: Chancen und Herausforderungen des digitalen Wandels für KMU* (pp. 177–195). Springer. https://doi.org/10.1007/978-3-662-59474-2_11

Gualtieri, L., Rojas, R. A., Ruiz Garcia, M. A., Rauch, E., & Vidoni, R. (2020). Implementation of a Laboratory Case Study for Intuitive Collaboration Between Man and Machine in SME Assembly. In D. T. Matt, V. Modrák, & H. Zsifkovits (Eds), *Industry 4.0 for SMEs: Challenges, Opportunities and Requirements* (pp. 335–382). Springer International Publishing. https://doi.org/10.1007/978-3-030-25425-4_12

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.

Jäntti, M. (2020). Studying Data Privacy Management in Small and Medium-Sized IT Companies. *2020 14th International Conference on Innovations in Information Technology (IIT)*, 57–62. https://doi.org/10.1109/IIT50501.2020.9299050

Khan, M., Ahmad, A., Sobieczky, F., Pichler, M., Moser, B. A., & Bukovsky, I. (2022). A Systematic Mapping Study of Predictive Maintenance in SMEs. *IEEE Access*, *10*, 88738–88749. https://doi.org/10.1109/ACCESS.2022.3200694

Le, T. D., Le-Dinh, T., & Uwizeyemungu, S. (2024). Search engine optimization poisoning: A cybersecurity threat analysis and mitigation strategies for small and medium-sized enterprises. *Technology in Society*, *76*, 102470. https://doi.org/10.1016/j.techsoc.2024.102470

Matt, C., Hess, T., & Benlian, A. (2015). Digital Transformation Strategies. *Business & Information Systems Engineering*, *57*(5), 339–343. https://doi.org/10.1007/s12599-015-0401-5

May, M. C., Glatter, D., Arnold, D., Pfeffer, D., & Lanza, G. (2024). IIoT System Canvas—From architecture patterns towards an IIoT development framework. *Journal of Manufacturing Systems*, *72*, 437–459. https://doi.org/10.1016/j.jmsy.2023.12.001

Moencks, M., Roth, E., Bohné, T., Romero, D., & Stahre, J. (2022). Augmented Workforce Canvas: A management tool for guiding human-centric, value-driven human-technology integration in industry. *Computers & Industrial Engineering*, *163*, 107803. https://doi.org/10.1016/j.cie.2021.107803

Morakanyane, R., Grace, A. A., & O'Reilly, P. (2017). Conceptualising Digital Transformation in Business Organisations: A Systematic Review of Literature. In *BLED 2017 Proceedings* (pp. 427–444).

Neidlinger, L. (2025). *Design and Evaluation of a Digital Transformation Canvas for SME* [Master Thesis, Faculty of Computer Science and Engineering, Frankfurt University of Applied Sciences]. https://eamlab.frankfurt-university.de/wp-content/uploads/2026/02/Masterthesis_LukasNeidlinger_1533200_orig-1.pdf

Osterwalder, A., & Pigneur, Y. (2013). *Business model generation: A handbook for visionaries, game changers, and challengers*. Wiley&Sons.

Pelletier, C., & Cloutier, L. M. (2019). Challenges of Digital Transformation in SMEs: Exploration of IT-Related Perceptions in a Service Ecosystem. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 4967–4976.

Romero, D., Gaiardelli, P., Powell, D., Wuest, T., & Thürer, M. (2019). Rethinking Jidoka Systems under Automation & Learning Perspectives in the Digital Lean Manufacturing World. *IFAC-PapersOnLine*, *52*(13), 899–903. https://doi.org/10.1016/j.ifacol.2019.11.309

Tian, J., Coreynen, W., Matthyssens, P., & Shen, L. (2022). Platform-based servitization and business model adaptation by established manufacturers. *Technovation*, *118*, 102222.

Torres, A. A., Collazos, C. A., & Mon, A. (2024). A Review of the Literature on models or frameworks for Implementation of Digital Transformation Processes. *Journal of Computer Science & Technology*, *24*(2), 153–167.

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research*, *122*, 889–901.

Yin, R. K. (2007). *Case study research: Design and methods* (3. ed., [Nachdr.]). Sage.

RESEARCH IN PROGRESS

TRAINING ORAL COMMUNICATION-SKILLS USING AN AVATAR

NANDA VAN DER STAP,[1] EMANUEL VAN DONGEN,[2]
STAN VAN GINKEL[1]

[1] University of Applied Sciences Utrecht, Utrecht, the Netherlands
nanda.vanderstap@hu.nl, stan.vanginkel@hu.nl
[2] Utrecht University, Utrecht, the Netherlands
e.g.d.vandongen@uu.nl

Training participants to give quality feedback requires them to receive feedback themselves on the feedback they have given to someone else. Such training is time-consuming and anxiety-inducing. Nevertheless, the reception of feedback is primarily dependent on its perceived quality, therefore training is essential. In the present study, an AI-driven Avatar was used as the recipient of feedback, generating a feedback report concerning the feedback given. 200 participants will train themselves using the Avatar, in which their purported findings on perceived value/usefulness of the training, their competence and confidence will be measured (pre and post) using questionnaires based on IMI scales. Additionally, Learning Analytics will be exploited to measure the number of times students logged in, and the duration of their sessions. The overall results should show the effects of using an Avatar on students' perceptions and confidence on feedback-giving. A future study may show how this affects students' professional practice.

DOI
https://doi.org/10.18690/um.fov.4.2026.62

ISBN
978-961-299-152-4

Keywords:
avatar,
oral feedback training,
human values,
digital ethics,
aritificial intelligence

University of Maribor Press

1 Introduction

Professionals need to be competent in oral communication, such as providing quality oral feedback. Repeatedly practising oral feedback-giving, one-on-one with a critical trainer, thereby receiving feedback on the feedback given, is necessary to improve its quality (Carless & Boud, 2018), but impractical and highly time-consuming (Chow, 2024), hence barely undertaken in higher (vocational) education. Additionally, training feedback-giving skills with a human agent may induce anxiety (Sichterman et al. 2025), hence inhibiting the learning process.

An AI-driven Avatar could help by taking the role of a one-on-one critical trainer that could be adaptable to different scenarios in which oral communication is necessary. Accordingly, anxiety would be reduced as practising can take place in privacy whilst remaining anonimous. Moreover, since such a training can be done in the user's own home, the practicial implication of time-consumption would become obsolete. To this extent, an Avatar would interact with the provider using verbal and visual responses, followed by a feedback report on the method of feedback given.

Furthermore, said design can be used for training oral feedback-giving for professionals as most higher vocational training programmes focus on peer-feedback only (Van der Stap, 2025). Focussing solely on peer-feedback is insufficient for training feedback-giving as this requires the presence of a one-on-one recipient, not as a peer, but as a critical trainer (Chow, 2024). Whilst some studies have implemented AI-technology on feedback training, these are related to peer-learning processes (Noroozi et al. 2023; Ouyang et al. 2021; Darvishi et al. 2024; Zheng at al. 2024), not to differing feedback-giving scenarios using the Avatar as a critical trainer. Consequently, there are no studies on training and practising oral feedback-giving, applicable to different situations, using an Avatar as critical trainer. This results in a scientific gap.

A major concern of using an AI-driven Avatar is users' wariness as AI-technology is becoming increasingly autonomous (Zhai et al. 2024; Skulmowski, 2023) resulting in concerns of how this may influence experiences and viewpoints (Van der Stappen & Van Steenbergen, 2020). These concerns affect users' confidence in AI-technology (Glikson & Wooley, 2020), inhibiting adoption of the Avatar (Soulami et al. 2024). Accordingly, users need to be confident that the interaction with the

Avatar does not affect their privacy and autonomy (Zhai et al. 2024; Golgeci et al. 2025), hence the Avatar design needs to be conducted using a structured approach (Nguyen et al. 2023; Markauskaite et al. 2022). One such structured approach is the Value-Sensitive-Design (VSD) approach. This approach in particular structurally collects human values, potential harms and benefits, relevant for developing a skill through the use of AI (Van der Molen et al. 2025). This should leverage users' adoption of the Avatar as it increases the confidence that such interaction does not affect autonomy and privacy (Zhai et al. 2024; Golgeci et al. 2025).

In the present study, the Avatar was designed through the implementation of the VSD approach. As a result, the Avatar's features include the systematic capture of the perspectives of different stakeholders on the potential harms and benefits of AI-use, making the Avatar itself sensitive in its responses to avoid potential harms whilst maintaining the benefits. The present study is a follow-up from said study, and examines the effects of the AI-driven Avatar that incorporates given human values and design principles.

1.1 Aim

Training oral communication skills is important for professionals, but impractical for higher vocational training programmes, leaving professionals ill-prepared or risking early resignations, for example in education (Golubtchik, 2024), resulting in a practical gap. There are studies on using an Avatar in feedback-giving, however, these focus on peer-feedback only, leaving a scientific gap. The aim of this study therefore is to fill said gaps by using the Value-Sensitive-Design method to structurally collect professionals' values and concerns into the Avatar-design as a one-on-one critical trainer. The research question therefore is:

What is the effect of AI feedback training on participants' perceived competence, and confidence to give quality feedback, and to what extent do participants find the training valuable and useful.

2 Methodology

In this research, we are interested in the impact that practising giving feedback-giving has in an Avatar feedback practice system, in particular on perceived usefulness of feedback-giving training, perceived competence in feedback-giving, and confidence

in feedback-giving. In short, in order to provide powerful, quality feedback as outlined above (Hattie & Timperley, 2007; Van Ginkel, 2019), being confident in giving feedback (Mukhia & Tenzin, 2025), perceiving competence in feedback-giving (Ryan & Deci, 2000), as well as seeing the usefulness of training in feedback-giving (Boud & Dawson, 2021), it is essential that training is provided in an anxiety-reduced environment such as AI (Sichterman, et al. 2025). Training is an essential element of providing quality feedback, hence experiencing the usefulness of such training may increase perceived competence and confidence. As explained by Wong et al. (2023), perceiving an AI training as useful increases engagement and learning effectiveness of the training.

The Avatar environment incorporates the criteria of quality feedback as outlined above, and provides feedback on feedback given to an Avatar, ensuring the minimalisation of anxiety. Participants can practise giving feedback as often and as long as they like and receive feedback on the feedback provided by them based on aforementioned criteria, in order to improve their perception of the three aspects as given. Our main research question is:

Does AI training in feedback-giving contribute positively to the perceived usefulness, perceived competence, and confidence of giving quality feedback?

In order to seek an answer to this question, the following sub-questions are posed:

1. How useful do participants find the training?
2. How does the training effect participants perceived competence?
3. How does the training effect participants confidence?

2.1 Research Setting

Whilst feedback is an essential element of professional growth in any work-related setting as an essential element of learning, most of the research setting will take place in an educational environment, namely a graduate teacher-trainer programme of HU University of Applied Sciences (HU) in which giving oral feedback is a critical component of professional training. Likewise, the students at the teaching-trainer programme at the Utrecht University (UU) will need to be competent in giving oral feedback as part of their professional practice training.

2.2 Participants and data collection

The participants at the HU consist of graduate students (n=50) in said teacher-trainer programme, participants at the UU comprise of graduate students in a teacher-training programme (n=15). In both groups a convenience sampling method was used, recruiting participants in professional development courses as these groups were larger and most likely to apply oral feedback giving in their professional practice. As the participants are from two different universities, sampling bias is limited. Nevertheless, the sample may not be fully representative of the broader population, limiting generalisability. Social desirability bias is lifted as all the surveys are fully anonymous. Prior to the main part of the empirical study, a pilot study was conducted to validate the questionnaire and method. This pilot took place at a revalidation institute and consisted of managers and doctors (n=17), followed up by retrospective interviews (n=3) to ascertain whether the answers given were in fact the answers they had intended to give. The pilot study served to assess the proposed data analysis techniques, and to uncover potential problems (Holloway, 1997). Whilst it was not the aim of the pilot to measure the Avatar's effectiveness, participants voluntarily commentend on the usefulness of the Avatar, ranging from »*this should be standard in any professional educational programme*«, or »*great way of practising communication skills without anyone looking over my shoulder*«, to »*very helpful, I wish I had known of this earlier*«, and »*this should form part of training in coaching as it improves communication skills without judgment*«.

2.3 Questionnaires

All participants will receive two sets of identical questionnaires, focussing on the elements of perceived competence and usefulness, both constructed from the Intrinsic Motivation Inventory (Ryan, 1982), and the element confidence, taken from the Course Interest Survey (Keller, 1987). One questionnaire will be distributed prior to the AI training, the other will be conducted after the training. The questions consist of 21 items, based on existing scales, and contextualised to the context of the present study (providing quality feedback). A 5-points Likert scale will be used ranging from 'rarely or never true' (1) to 'always or almost always true' (5). The elements will be measured by multiple items distributed throughout the questionnaire (see Appendix A). To avoid social desirability bias (Nederhof, 1985), the questionnaires will be conducted anonymously.

The questionnaires include several elements of which the first element, Value/Usefulness, has seven items, e.g.: I find practising giving feedback using an Avatar an important activity, the second element, Perceived Competence, also has seven items, e.g.: I think I'm good at giving feedback. The third element, Confidence, has seven items as well, e.g.: As I take this training, I think I can successfully complete (the final session of) giving feedback. The questionnaire, made by Van Dongen (2020), had been validated twice prior this present study and used in a multi-measured study among law students in the years 2020-2021 and 2023 (Van Dongen, 2024). Said questionnaire has subsequently been adapted by Van der Stap (2025) and validated in a pilot case study, in order to suit the present study aimed at students in a teacher-trainer college during the years 2025-2026.

2.4 Data Analysis

After all the questionnaires are received, the results will be analysed using SPSS (descriptive analysis, correlations, repeated measures ANOVA). Additionally, the AI programme has a Learning Analytics (LA) construct that measures the number of times students logged into the programme, and how long they stayed in the AI environment. The LA results, i.e. the amount of times logged in and the duration of each practise session will be compared to student reactions in surveys, which will be coded to guarantee anonymity. To this extent, a full picture will emerge through which the usefulness and effectiveness of the avatar training can be assessed.

References

Boud, D., & Dawson, P. (2021). What feedback literate teachers do: an empirically-derived competency framework. *Assessment & Evaluation in Higher Education*, *48*(2), 158–171.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315-1325.

Chow, M. F. (2024). Understanding student perspectives on peer feedback: Written versus video versus face-to-face dialog. *Active Learning in Higher Education*.

Darvishi, A., Khosravi, H., Sadiq, S., Gasevic, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, 210, 104967. https://doi.org/10.1016/j.compedu.2023.104967.

Glikson, E. & Woolley, A.W. (2020). *Academy of Management Annals* 14(2), 627–660. https://doi.org/10.5465/annals.2018.0057.

Golgeci, I., Ritala, P., Arslan, A., McKenna, B. & Ali, I. (2025). Confronting and alleviating AI resistance in the workplace: An integrative review and a process framework. Human Resource Management Review. 35(2). https://doi.org/10.1016/j.hrmr.2024.101075.

Golubtchik, L. (2024). Increasing Teacher Retention by Improving Self-Efficacy and Classroom Management Skills in Pre-Service Teachers. Journal of Education and Learning, 13(4).

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81-112.

Holloway, I. (1997). *Basic Concepts for Qualitative Research.* Oxford: Blackwell Science.

Keller, J. (1987). Development and Use of the ARCS Model of Instructional Design', *Journal of Instructional Development* 13), 2-10.

Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Buckingham Shum, S., Gašević, D., & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers and Education: Artificial Intelligence*, 3, 100056. https://doi.org/10.1016/j.caeai.2022.100056.

Mukhia, S., Tenzin, K. Training academic staff for effective feedback in workplace-based assessment: a study in Bhutan. *BMC Med Educ* **25**, 748 (2025).

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology, 15*(3), 263–280.

Nguyen, T., Walker, C., & Bainbridge, D. (2023). Ethical considerations in AI-enhanced digital learning environments: Balancing innovation and responsibility. *AI & Society*, 38(1), 67–84. https://doi.org/10.1007/s00146-022-01345-9.

Noroozi, O., & de Wever, B. (2023). *The power of peer learning: Fostering students' learning processes and outcomes*. Springer.

Noroozi, O., Banihashem, S. K., Biemans, H. J., Smits, M., Vervoort, M. T., & Verbaan, C. L. (2023). Design, implementation, and evaluation of an online supported peer feedback module to enhance students' argumentative essay quality. *Education and Information Technologies*, 1–28. https://doi.org/10.1007/s10639-023-11683-y

Ouyang, F., Chen, Z., Cheng, M., Tang, Z., & Su, C. Y. (2021). Exploring the effect of three scaffoldings
on the collaborative problem-solving processes in China's higher education. *International Journal of Educational Technology in Higher Education*, 18(35). https://doi.org/10.1186/s41239-021-00273-y

Ryan , R. M. 1982) . Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory . *Journal of Personality and Social Psychology* , 43 : 450 – 461.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation,
social development, and well-being. *American Psychologist, 55*(1), 68–78.

Sichterman, B., Banihashem, S. K., Verstappen, M., Schipper, M., Waisvisz, P., Noroozi, O., & van Ginkel, S. (2025). Can AI feedback stand alone for fostering students' presentation performance? A comparison of AI-only versus teacher-supported AI feedback. *Interactive Learning Environments*, 1–18.

Skulmowski, A. (2023). Ethical implications of immersive technology in education. *Educational Technology & Society*, 26(2), 1–12.

Soulami, M., Benchekroun, S. & Galiulina, A. (2024). Frontiers in Artificial Intelligence, 7. https://doi.org/10.3389/frai.2024.1473872

Van Dongen, E. (2024). Pleading in the Virtual Courtroom: Exploring Experiential Learning in Law through Virtual-Reality-Based Exercises and Student Feedback. *European Journal of Legal Education*, 5(1), 57–190.

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2015). Towards a set of design principles for developing oral presentation competence: A synthesis of research in higher education. *Educational Research Review, 14,* 62-80.

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2017a). The impact of the feedback

source on developing oral presentation competence. *Studies in Higher Education, 42*(9), 1671-1685.

Van Ginkel, S., Gullikers, J., Biemands, H., Noroozi, O., Roozen, M., Bos, T., Van Tilborg, R., Van Halteren, M., Mulder, M. (2019). Fostering oral presentation competence through a virtual reality-based task for delivering feedback. Computers & Education, 134, 78-97.

Van der Stap, N. (2025). Internal Report: a study of higher education programmes (*n*=17) throughout The Netherlands

Van der Stap, N., van den Bogaart, T., van Ginkel, S., Rahimi, E., & Versendaal, J. (2024b). Towards teaching strategies addressing online learning in blended learning courses. Journal of Computers and Education. https://doi.org/10.1016/j.compedu.2024.105103.

Van der Stappen, E., & Van Steenbergen, N. (2020). Ethics by design: Practical applications in educational technology. *Journal of Educational Innovation*, 2(1), 45–62.

Vanhoof, S., & Speltincx, G. (2021). *Feedback in de klas. Verborgen leerkansen.* LannooCampus.

Van der Molen, O., Van Dongen, E., & Van Ginkel, S. (2025). Developing a Value Sensitive VR Learning Environment Fostering Professionals' Communication Skills. Presented at BLED Conference, Slovenia.

Wong, E.Yc., Hui, R.Ty. & Kong, H. (2023). Perceived usefulness of, engagement with, and effectiveness of virtual reality environments in learning industrial operations: the moderating role of openness to experience. *Virtual Reality,* **27**, 2149–2165.

Zhai, X., Qin, N., Zhang, J., & Song, G. (2024). AI and education: Risks of cognitive outsourcing and the need for critical thinking reinforcement. *Educational Review*, 76(1), 95–112. https://doi.org/10.1080/00131911.2023.1934567.

Zheng, L., Fan, Y., Chen, B., Huang, Z., Gao, L., & Long, M. (2024). An AI-enabled feedback-feedforward approach to promoting online collaborative learning. *Education and Information Technologies*, 29(9), 11385–11406. https://doi.org/10.1007/s10639-023-12292-5.

About the authors

Dr. **Nanda van der Stap** is Curriculum Coordinator for the Master Teaching English at the HU University of Applied Sciences, and researcher at the Research Centre Digital Ethics. Additionally, she is involved in several projects that bring together developing competences, education, and AI-driven technology.

Dr. **Emanuel G.D. van Dongen** is Associate Professor at the Molengraaff Institute, Utrecht University, and researcher at the Utrecht Centre for Accountability and Liability Law. Alongside his teaching and research, he is actively involved in (technology-driven) innovation projects, connecting insights from Educational Sciences with Technological Developments.

Dr. **Stan van Ginkel** is Professor of Applied Sciences at the Research Group Digital Ethics at the HU University of Applied Sciences. Further, he supervises PhD students and other researchers in projects on digital literacy, digital didactics and digital ethics.

RESEARCH IN PROGRESS

DESIGNING EVIDENCE-BASED MANAGERIAL DECISION SUPPORT TOOLS FOR OLDER ADULT MANAGERS

KHALED HASSANEIN, MILENA HEAD
McMaster University, DeGroote School of Business, Hamilton, Canada
hassank@mcmaster.ca, headm@mcmaster.ca

DOI
https://doi.org/10.18690/um.fov.4.2026.63

ISBN
978-961-299-152-4

Evidence-based managerial decision making (EBMDM) is increasingly critical in digital and data-rich organizational environments. At the same time, the managerial workforce is aging, with many senior and experienced managers continuing to make high-stakes decisions well into later adulthood. Prior research suggests that normal age-related cognitive changes can affect how older adults search for, integrate, and synthesize information, particularly in technology-mediated settings. Yet, decision support systems and managerial tools are rarely designed with age-related cognitive variability in mind. This work-in-progress paper presents an ongoing research program that leverages neurophysiological and human–computer interaction methods to understand evidence gathering and decision making among older adult managers at a deeper level and to inform the design of customizable EBMDM tools. By integrating EEG, eye-tracking, physiological and behvioural measures with traditional techniques to observe older adult managers engaged in realistic managerial decision tasks, this research aims to contribute to decsion making theory and develop adaptive decision aids that reduce cognitive burden, mitigate bias, and support high-quality decision making across the adult lifespan.

Keywords:
decision support, older adults, evidence-based, technology design, neurophysiology

University of Maribor Press

1 Introduction

Organizations increasingly expect managers to make complex, high-stakes decisions based on large volumes of digital evidence drawn from internal databases, analytics platforms, and AI-enabled decision support tools that assist with data aggregation, analysis, and visualization. In today's data-rich and data-driven business environment, such technologies can enhance decision speed and consistency, positively influencing competitive advantage. At the same time, as managerial decision contexts become more data-intensive and technologically mediated, they place greater cognitive demands on decision makers, particularly with respect to attention, reasoning, and the integration of multiple sources of evidence. As a result, decision makers often resort to strategies such as making decisions based on intuition or satisficing, where one settles for suboptimal decisions to reduce one's cognitive burden (Misuraca & Teuscher 2013).

Against this backdrop, demographic shifts are reshaping the managerial workforce (Heyden et al. 2025). Older adults are remaining in managerial and executive roles longer, bringing deep domain expertise and institutional knowledge, but also experiencing normal age-related changes in cognition. Research in cognitive aging shows that while crystallized intelligence and experience often remain stable or improve, fluid cognitive abilities (such as reasoning, pattern recognition, and problem solving) tend to decline with age due to neurocognitive, structural, and functional changes in the brain (Harada et al. 2013). Age-related declines in attention and memory can further affect learning, information synthesis, and interaction with digital technologies, including contemporary AI-supported decision tools (de Bruin et al. 2012). Despite these well-documented changes, most managerial decision support tools implicitly assume cognitively uniform users. This mismatch risks increasing cognitive overload, reinforcing reliance on intuition or satisficing, and ultimately degrading decision quality, particularly for older adult managers operating in digital environments (Sin et al. 2021).

This paper reports on a work-in-progress research program that seeks to understand how aging-related cognitive changes affect evidence-based managerial decision making (EBMDM) and to translate these insights into the design of adaptive decision support tools tailored to older adult managers. Specifically, this research is guided by two interrelated objectives:

Objective 1: To understand the real-time cognitive and neurophysiological mechanisms underlying evidence gathering and decision making among older adult managers in realistic EBMDM tasks.
Objective 2: To design and evaluate customizable EBMDM decision aids that adapt to age-related cognitive characteristics, reducing cognitive burden while maintaining or improving decision quality.

Importantly, we position aging not as a deficit to be corrected, but as a source of cognitive variability that should inform inclusive and adaptive DSS design. By explicitly accounting for age-related changes in attention, memory, and reasoning, EBMDM tools can better support sustained managerial performance across the lifespan.

2 Background

2.1 Evidence-Based Managerial Decision Making (EBMDM)

EBMDM can be viewed as a dynamic process through which managers, supported by information technology, deliberately search for, evaluate, and integrate multiple sources of evidence when making decisions. While EBMDM has been shown to improve decision quality and help mitigate against decision biases (Arnott 2006), it is cognitively demanding, requiring sustained attention, working memory, and executive control. In data-rich digital environments, these demands are amplified by information overload and heterogeneous data formats. When cognitive resources are strained, decision makers often resort to heuristics or satisficing strategies, settling for suboptimal decisions to reduce cognitive effort. For older adult managers, whose cognitive profiles or capacities may differ from their younger counterparts, these pressures may be especially pronounced (Wei et al. 2011).

2.2 Cognitive Aging and Managerial Decision Making

Normal aging is associated with gradual changes in brain structure and function that affect fluid cognitive abilities. Declines in processing speed, working memory, and attentional control can make it more difficult to compare alternatives, detect patterns, and integrate complex information streams (Juras et al. 2025). Importantly, these changes do not imply diminished managerial competence; rather, they suggest

that older adults may benefit from decision environments that better align with their cognitive strengths and limitations (Duffy et al 2025).

In technology-mediated contexts, age-related changes can influence how information is visually scanned, how long attention is sustained on particular evidence sources, and how confidently decisions are made under uncertainty (Frank & Seaman 2023). Yet, empirical research directly examining these processes in realistic managerial tasks remains limited.

3 Methodology

3.1 Experimental Setting and Decision Tasks

The empirical component of this research will be conducted in a state-of-the-art research facility designed to support the synchronized collection of neurophysiological and behavioural data in realistic decision-making contexts. Observation rooms are configured to resemble typical managerial workspaces, including individual offices and conference-room settings, allowing participants to engage with digital decision support tools in contexts that closely approximate real-world managerial environments.

A key focus of the study is on older adult managers, broadly defined as individuals in later adulthood who continue to hold decision-making authority within organizations. Participants who fit our inclusion criteria will engage in a series of realistic EBMDM tasks that reflect common managerial decision domains in digitally mediated environments. These tasks include, but are not limited to, hiring and promotion decisions, supply-chain and vendor selection scenarios, and performance evaluation tasks based on data dashboards and digital reports. Each task requires participants to search for, evaluate, and integrate multiple sources of evidence presented through a computer-based interface.

The experimental design manipulates key characteristics of information presentation, such as data density, visual layout, sequencing, and framing, to examine how these factors interact with age-related cognitive changes. Tasks are structured to elicit varying levels of cognitive demand, enabling comparison of decision strategies and cognitive load across conditions. Both individual and group-based

decision scenarios are incorporated, reflecting the increasing prevalence of collaborative and virtual decision making in contemporary organizations. This design allows the research team to observe how older adult managers adapt their evidence-gathering strategies under different informational and social conditions.

3.2 Neurophysiological and Behavioural Measures

Traditional decision-making research relies heavily on self-report measures and behavioural outcomes, which can obscure real-time cognitive processes and are susceptible to bias. For example, self-reported data are susceptible to subjectivity bias, social desirability bias, and demand effects (Dimoka et al. 2012). Neurophysiological methods (such as EEG, eye-tracking, and physiological arousal measures) offer a complementary approach by capturing continuous, real-time indicators of cognitive load, attention, and engagement (Aimone et al 2016; Huk et al. 2018). These measures provide insight into how older adult managers search for evidence, allocate attention across competing data sources, and prioritize information during decision making. Behavioural performance indicators (e.g., decision accuracy, consistency, and response time) and post-task questionnaires assessing perceived effort and confidence will complement the neurophysiological data, enabling triangulation across multiple levels of analysis. Combining neurophysiological data with HCI methods enables a more holistic understanding of decision processes, particularly in complex, dynamic, and group-based managerial contexts.

3.3 Design and Evaluation of the EBMDM Decision Aids

Insights generated from the empirical analyses will be used to inform the iterative design of customizable EBMDM decision aids. These tools will adapt interface features (such as information chunking, visual emphasis, and decision pacing) to the cognitive needs and preferences of older adult managers. Prototype decision aids will be evaluated through follow-up experiments using the same multimodal measurement framework. Effectiveness will be assessed based on improvements in decision quality, reductions in cognitive workload and stress, and enhanced user experience. Rather than adopting a one-size-fits-all approach, the design philosophy emphasizes flexibility and personalization, allowing decision aids to accommodate individual differences in cognitive capacity and decision style.

4 Expected Contributions

This research is expected to make several important contributions to the literature on decision support systems and managerial decision making by explicitly foregrounding aging as a central design consideration rather than a peripheral demographic variable. First, at a theoretical level, the study advances EBMDM research by integrating insights from cognitive aging and neuroscience into models of managerial decision making. While prior EBMDM and DSS research has largely assumed cognitively homogeneous users, this work highlights how normal age-related changes in fluid cognitive abilities, attention, and memory systematically shape evidence gathering, information synthesis, and decision strategies. In doing so, the research reframes aging not as a deficit but as a source of meaningful cognitive variability that must be accounted for in digitally mediated managerial environments.

Second, the research contributes methodologically by demonstrating the value of multimodal neurophysiological approaches for studying managerial decision making in realistic, technology-enabled contexts. By combining EEG, eye tracking, physiological measures, and traditional behavioural and self-report data, the study provides a more granular understanding of real-time cognitive processes than is possible with conventional methods alone. This approach is particularly valuable for examining subtle age-related differences in cognitive load and attentional allocation that may not be consciously accessible to decision makers themselves. As such, the methodology offers a replicable framework for future researchers interested in studying decision making under conditions of complexity, information overload, and cognitive change.

Finally, the study is expected to generate actionable design knowledge with direct practical implications for organizations facing an aging managerial workforce. The findings will inform concrete design principles for EBMDM tools that reduce unnecessary cognitive burden, improve information presentation, and support sustained attention among older adult managers. By translating insights about cognitive aging into prescriptive guidance for DSS design, the research supports more inclusive decision support technologies that enhance decision quality, reduce reliance on intuition or satisficing, and accommodate cognitive diversity across the managerial lifespan.

References

Aimone, Jason & Ball, Sheryl & King-Casas, Brooks. (2016). It's not what you see but how you see it: Using eye-tracking to study the risky decision-making process. Journal of Neuroscience, Psychology, and Economics. 9. 137-144. 10.1037/npe0000061.

Arnott, D., Cognitive biases and decision support systems development: a design science approach. Information Systems Journal, 2006. 16(1): p. 55-78.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2012). Explaining adult age differences in decision-making competence. Journal of Behavioral Decision Making, 25(4), 352-360.

Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P.H., Pavlou, P.A., Müller-Putz, G., Riedl, R., Brocke, J.V., and Weber, B. (2012). On the Use of Neurophysiological Tools in IS Research: Developing a Research Agenda for Neurois, MIS Quarterly, 36(3), 679- 702.

Duffy, J. S., Bellgrove, M. A., Murphy, P. R., & O'Connell, R. G. (2025). Disentangling sources of variability in decision-making. Nature Reviews Neuroscience, 1-16.

Frank, C. C., & Seaman, K. L. (2023). Aging, uncertainty, and decision making: A review. Cognitive, Affective, & Behavioral Neuroscience, 23(3), 773-787.

Harada, C. N., Love, M. C. N., & Triebel, K. L. (2013). Normal cognitive aging. Clinics in geriatric medicine, 29(4), 737-752.

Heyden, M. L., Wechtler, H. M., & van Doorn, S. (2025). Wrinkle of change? The reproduction of executive age profiles across CEO succession episodes. Human Relations, 00187267251327955.

Huk, A. et al. "Beyond Trial-Based Paradigms: Continuous Behavior, Ongoing Neural Activity, and Natural Stimuli." The Journal of Neuroscience 38 (2018): 7551 - 7558.

Juras, L., Martincevic, M., & Vranic, A. (2025). A tale of two ages: fluid reasoning as a predictor of working memory training efficacy in middle-aged and older adults. Aging, Neuropsychology, and Cognition, 1-23.

Misuraca, R., & Teuscher, U. (2013). Time flies when you maximize—Maximizers and satisficers perceive time differently when making decisions. Acta psychologica, 143(2), 176-180.

Sin, J., L. Franz, R., Munteanu, C., & Barbosa Neves, B. (2021). Digital design marginalization: New perspectives on designing inclusive interfaces. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1-11.

Wei, K. K., Teo, H. H., Chan, H. C., & Tan, B. C. (2011). Conceptualizing and testing a social cognitive model of the digital divide. Information Systems Research, 22(1), 170-187.

RESEARCH IN PROGRESS

TEACHING A SYSTEMS APPROACH TO MANAGE SMART SUSTAINABLE CITIES

QUAN ZHU,[1] STEVEN HAVEMAN[2]

[1] HU University of Applied Sciences Utrecht, Research Group Process Innovation & Information Systems, Utrecht, the Netherlands
quan.zhu@hu.nl
[2] HU University of Applied Sciences Utrecht, Research Group Organisations in Digital Transition, Utrecht, the Netherlands
steven.haveman@hu.nl

Cities are pushing their boundaries to transform into smart and sustainable urban spaces, where factors such as ecological conservation, economic development, and democratic equity should be considered at the same time. Systems thinking is therefore among the most needed skills for practitioners managing smart sustainable cities. The purpose of this paper is to initiate the development of a curriculum for a systems approach to manage smart sustainable cities, leveraging theoretical concepts of systems thinking. Our framework will enable learners to view societal challenges as outcomes of social-ecological system behaviour over time, so that key leverage points for interventions can be identified.

DOI
https://doi.org/10.18690/um.fov.4.2026.64

ISBN
978-961-299-152-4

Keywords:
systems thinking, smart sustainable cities, skills development, social-ecological system, system behaviour, leverage points

University of Maribor Press

1 Introduction

"Smart city" has transpired to address challenges and possibilities in urban planning and city development for a vast majority of the population across several disciplines (Kutty et al., 2020). Simply speaking, cities that strive to make themselves "smarter" are termed as smart cities (Chourabi et al., 2012). Meanwhile, modern societies face many sustainability challenges from the environment (e.g., air quality, water quality, and waste handling), the economy (e.g., unemployment, wealth gap, and economic inequality), and the society (e.g., food security, water scarcity, and poverty) (Zander & Mosterman, 2017). Cities are therefore pushing their boundaries to transform into smart and sustainable urban spaces, where factors such as ecological conservation, economic development, and democratic equity should be considered at the same time (Kutty et al., 2020). This is in line with the finding of *Smart Communities Skills Development in Europe* (SMARCO, 2025) that systems thinking is among the most needed skills for practitioners managing smart sustainable cities.

Teaching a system approach (to manage smart sustainable cities) encourages learners to engage in nonlinear, synthetic and adaptive thinking founded upon systems thinking (Capra, 1996). Systems thinking is defined as "a discipline for seeing wholes. ... a framework for seeing interrelationships rather than things, for seeing patterns of change rather than static 'snapshots'" (Senge, 2006, p. 68). It calls for managerial decision-making based on insight into the complex dynamics of social-ecological systems to broaden horizons for understanding and addressing societal challenges in new and different ways (Sterman, 1994).

The purpose of this paper is to initiate the development of a curriculum for a systems approach to manage smart sustainable cities, leveraging theoretical concepts of systems thinking. With reference to Kennedy et al. (2026), our framework comprises four conceptual moves: *developing a complex adaptive systems (CAS) mindset, identifying structure and form of systems, understanding behavioural dynamics of systems over time,* and *intervening in systems*. In addition, we discuss one illustrative example (Breukelman et al., 2022) to make the approach more tangible and domain-specific.

2 The Learning Framework

Below, we offer guidance through a learning framework that integrates problem-solving approaches of systems thinking (Ackoff, 1974; Forrester, 1994; Meadows, 2008; Senge, 2006; Sterman, 2001).

2.1 Developing a CAS Mindset

Societal challenges are dynamic, changing through a process of coevolution involving multiple entities (such as population, governance quality, budget, solid waste management performance, etc.) within complex entanglements (Kennedy et al., 2026). A CAS mindset therefore enables learners to embrace complexity and understand different entities as inherently situated and interdependent within broader contexts (Shrivastava & Kennelly, 2013). Below, we will use the example about Solid Waste Management (SWM) in cities (Breukelman et al., 2022) to illustrate such interdependency.

SWM research has mostly focused on symptoms (such as lack of resources, infrastructure, awareness, and institutional strength), while underlying reasons have attracted little attention (Breukelman et al., 2019). In fact, the performance of waste collection can be considered as the product of a complex urban system in which processes on demographics, economics, and physical (infra)structures interact with social and political processes (Breukelman et al., 2022). For instance, the combined effect of economic growth and urbanisation drives an increase of build-up area, population density, traffic, and consumption (Mangi et al., 2020). In many cases, such growth overtakes the capacity-building of administrative organisations to adapt themselves and their cities (Mason et al., 2020). Where this happens in an urban context with low institutional quality (Beard et al., 2016), fragmented governance (Ahrend et al., 2017), discontinuities in management (Ezeah & Roberts, 2014), and absence of stable leadership (Cornea et al., 2017), tax incomes and urban investments start to shrink (Rice & Rice, 2012). In turn, this leads to diseconomies of agglomeration (Venables, 2017), initiating a vicious circle (Beard et al., 2016).

2.2 Identifying Structure and Form of Systems

A systems approach teaches learners to exercise caution before intervening in societal challenges (Kennedy et al., 2026). Learners are taught to explore underlying structures and causalities of social challenges and the values and beliefs that shape this structuring (Kennedy et al., 2026). For this purpose, understanding the concepts of feedback and delays is the starting point.

A feedback loop is a "closed chain of causal connections" (Meadows, 2008, p. 27) typically categorized as those that amplify an effect (reinforcing or positive feedback loops) and those that dampen an effect (balancing or negative feedback loops). Feedback loops help learners develop closed-loop thinking about why societal challenges persist and how entities are part of these causal connections and act to reinforce or break feedback loops. Delays refer to the time lag between cause and effect as materials and information flow through the system (Sterman, 2001). Recognizing delays helps learners understand long-term consequences and reduce the risk of instability or breakdowns by taking corrective action when unnecessary (Senge, 2006). Figure 1 shows a simplified model for the example from Breukelman et al. (2022), with "R" denoting a reinforcing loop, "B" denoting a balancing loop, and "‖" denoting a delay.

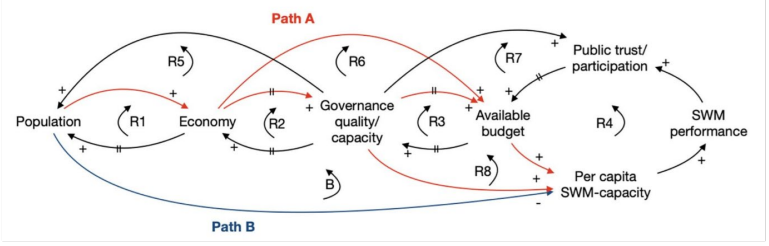

Figure 1: A simplified SWM model
Source: Breukelman et al. (2022)

2.3 Understanding Behavioural Dynamics of Systems Over Time

Building on an understanding of the structure of a system, learners can seek to gain an appreciation for the state of societal challenges and their patterns of change over time (Kennedy et al., 2026). Rather than viewing single events of societal challenges

in isolation, learners understand them as behavioural patterns in systems influenced by fast-paced effects of temporally proximate actions combined with slow-paced effects of those from the distant past (Meadows, 2008). Learners gain an appreciation for nonlinear dynamics of change and the crossing of potential threshold points whereby underlying dynamics driving behaviour of societal challenges may fundamentally transform (Kennedy et al., 2026).

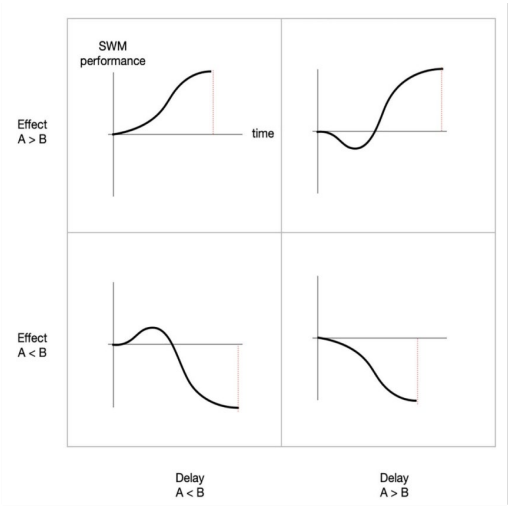

Figure 2: Projections of SWM performance under different regimes of effects and delays in the dominant paths
Source: Breukelman et al. (2022)

In the example from Breukelman et al. (2022), the expected behaviour may be either continued growth or decline in SWM performance, which may be strongly influenced by the pace of population growth. Figure 1 shows two important paths starting from "Population": Path A going through "Economy", "Governance quality/capacity", and "Available budget" towards "Per capita SWM capacity" (which are all reinforcing loops), while Path B going directly to "Per capita SWM capacity" (which is a balancing loop). The effects and delays in both paths dominate the evolution of SWM performance over time, which can be further simplified as the interaction between the effect of a reinforcing loop (i.e., Path A) and the effect of a balancing loop (i.e., Path B) with delays respectively. Such a structure is a typical Limits to Growth system archetype (Senge, 2006). The plots in Figure 2 demonstrate such system behaviour over time (under different regimes of effects and delays in

Paths A and B). A huge effect and short delay in Path A may be the most beneficial scenario for SWM performance. Longer delay in Path A will lead to an initial reduction in SWM performance, but will gradually be overcome in the long run. A small effect and long delay in Path A do the opposite.

2.4 Intervening in Systems

The purpose of this conceptual move is to understand how systems may be transformed by appreciating what is governing stability and what forces enable change. Learners are invited to search for leverage points for change (Meadows, 2008) and to consider broad ranges of impacts and unintended consequences (Sterman, 2002).

For example, when identifying leverage points to improve SWM performance (Breukelman et al., 2022), it is important to consider: (1) the population and economy part (on the left-hand side of Figure 1) is not useful for interventions, as these processes are autonomous, difficult to manage, and any intervention may lead to counter effects (i.e., via both Path A and Path B). For instance, lowering the inflow of new inhabitants may seem attractive, but it will also hurt the economy and reduce available budgets. (2) The SWM part (on the right-hand of Figure 1) may be a good point for interventions, but it is not. For example, any investment in capacity (i.e., shorten the delay of Path A) may lead to an improved performance in the short run, but if this is not followed by improvements in governance and yearly budgets (i.e., improve the effect of Path A), the upswing will not last long and may even cause the decrease of public trust and a reluctancy on further improvements. (3) The key to success is in the hands of governance. Political and administrative processes should focus on budget availability, which should keep pace with population and city growth. The most promising way to do so is to implement and enforce taxation schemes and, at the same time, safeguarding service availability and quality to prevent eroding public trust (i.e., the upper left corner of Figure 2).

3 Discussion and Future Work

Our framework may enable learners to view societal challenges as outcomes of social-ecological system behaviour over time. Early pilot evidence—such as classroom tests, practitioner reactions, or preliminary learning outcomes—would

make our contribution more concrete. In addition, Breukelman et al.'s (2022) example is mainly about sustainability, while the "smart" side is not fully reflected. A further inclusion of illustrative examples, such as the one in Kutty et al. (2020), is in our plan to link between information systems (IS) and socio-technical system design, with clearer implications on IS education and digital decision support tools.

Acknowledgment

Special thanks to Dr. Mohamed Eledeisy (HU University of Applied Sciences Utrecht), for his efforts on proofreading.

References

Ackoff, R. (1974). The systems revolution. *Long Range Planning*, *7*(6), 2–20.

Ahrend, R., Farchy, E., Kaplanis, I., et al. (2017). What makes cities more productive? Evidence from five OECD countries on the role of urban governance. *Journal of Regional Science*, *57*, 385–41.

Beard, V., Mahendra, A., & Westphal, M. (2016). Towards a more equal city: Framing the challenges and opportunities. *World Resources Institute Report*. Washington, DC: World Resources Institute.

Breukelman, H., Krikke, H. & Löhr, A. (2019). Failing services on urban waste management in developing countries: A review on symptoms, diagnoses, and interventions. *Sustainability*, *11*, 6977.

Breukelman, H., Krikke, H., & Löhr, A. (2022). Root causes of underperforming urban waste services in developing countries: Designing a diagnostic tool, based on literature review and qualitative system dynamics. *Waste Management & Research: The Journal for a Sustainable Circular Economy*, *40*(9):1337-1355. doi:10.1177/0734242X221074189

Capra, F. (1996). *The web of life: A new synthesis of mind and matter*. Harper Collins.

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., Scholl, H. J. (2012). Understanding Smart Cities: An Integrative Framework. Paper presented at: *45th Hawaii International Conference on System Sciences*, 2289–2297.

Cornea, N., Véron, R., & Zimmer, A. (2017). Clean city politics: An urban political ecology of solid waste in West Bengal, India. *Environment and Planning A*, *49*, 728–744.

Ezeah, C. & Roberts, C. L. (2014). Waste governance agenda in Nigerian cities: A comparative analysis. *Habitat International*, *41*, 121–128.

Forrester, J. W. (1994). System dynamics, systems thinking, and soft OR. *System Dynamics Review*, *10*(2–3), 245–256.

Kennedy, S., Grewatsch, S., Liboni, L., & Cezarino, L. O. (2026). Teaching a Systems Approach to Address the Sustainability Management Disconnect. *Business & Society*, *65*(3), 736-770.

Kutty, A. A., Abdella, G. M., Kucukvar, M., Onat, N. C., Bulu, M. (2020). A system thinking approach for harmonizing smart and sustainable city initiatives with United Nations sustainable development goals. *Sustainable Development*, *28*, 1347–1365. https://doi.org/10.1002/sd.2088

Mangi, M. Y., Yue, Z., Kalwar, S., et al. (2020). Comparative analysis of urban development trends of Beijing and Karachi metropolitan areas. *Sustainability*, *12*, 451.

Mason, N., Oyaya, C., & Boulenouar, J. (2020). Reforming urban sanitation under decentralization: Cross-country learning for Kenya and beyond. *Development Policy Review*, *38*, 42–63.

Meadows, D. H. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.

Rice, J., & Rice, J. S. (2012). Debt and the built urban environment: Examining the growth of urban slums in the less developed countries, 1990-2010. *Sociological Spectrum*, *32*, 114–137.

Senge, P. M. (2006). *The fifth discipline: The art and practice of the learning organization*. Random House.

Shrivastava, P., & Kennelly, J. J. (2013). Sustainability and place-based enterprise. *Organization and Environment*, *26*(1), 83–101.

SMARCO (2025). *D4.1. Competence Matrix for Short-Term Trainings*. Retrieved from: https://smarco.eu/project-resources/

Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review*, *10*(2–3), 291–330.

Sterman, J. D. (2001). System dynamics modeling: tools for learning in a complex world. *California Management Review*, *43*(4), 8–25.

Sterman, J. D. (2002). All models are wrong: Reflections on becoming a systems scientist. *System Dynamics Review*, *18*(4), 501–531.

Venables, A. J. (2017). Breaking into tradables: Urban form and urban function in a developing city. *Journal of Urban Economics*, *98*, 88–97.

Zander, J., & Mosterman, P. (2017). *Computation for humanity: Information technology to advance society (1st ed.)*. Boca Raton, FL: CRC Press.

RESEARCH IN PROGRESS

THE IMPACT OF DATA SHARING IN PRODUCT RECALLS: A COST-BASED APPROACH

MAARTEN DE MILDT, SOFIE VERBRUGGE, DIDIER COLLE
IDLab, Department of Information Technology at Ghent University - imec, Ghent, Belgium
maarten.demildt@ugent.be, sofie.verbrugge@ugent.be, didier.colle@ugent.be

Product recalls often incur substantial operational, and reputational costs across supply chains due to precautionary "overkill" practices. While data sharing is increasingly promoted to enable more precise recalls, evidence on its actual impact remains limited. This work-in-progress paper investigates how data sharing can impact recall costs by conducting a multiple-case study of four recall incidents across the food and automotive sectors. For each case, we map tangible and intangible recall costs, as well as the risks associated with sharing data. We estimate the potential reduction in recalled products and preliminarily suggest conditions under which such potential could be reached. In the next step, we aim to support our findings with quantitative proof on the costs and risks of data sharing in product recalls and perform structured mapping of recall cases to uncover patterns in potential for data sharing.

DOI
https://doi.org/10.18690/um.fov.4.2026.65

ISBN
978-961-299-152-4

Keywords:
data sharing,
data spaces,
product recalls,
cost modelling,
case study

University of Maribor Press

1 Introduction

In supply chains, product recalls are typically the result of faulty products considered dangerous to public safety and health (Ketchen et al., 2014). In a best case scenario, referred to as precise recalls, recalls would be efficient, meaning the faulty product is recalled and non-faulty ones can remain untouched (Ketchen et al., 2014). However, out of safety precautions, recalls occur frequently, and often at large scale as an "overkill" measure (Ketchen et al., 2014). This has a major impact on costs, interorganizational trust in supply chains, consumer trust, reputation, and company valuation to name a few (Fang et al., 2024).

Efforts to reduce the impact of recalls have focused on improving data sharing in supply chains. Solutions range from data platforms like the RASFF platform to track food contamination (Djekic et al., 2017), to more advanced data space solutions like Catena-X to track automotive supply chains (Ammann and Hess, 2025; Kernstock et al., 2025). Catena-X allegedly reduced 1.4 million vehicle recalls to just 14, which is considered a major success story (AIAG, 2025). However, little evidence is provided for this reduction in recalled vehicles.

Additionally, related work challenges if the benefits of sharing data in supply chains can be achieved. Agahari et al. (2022) show that data sharing generates value in supply chains but is hindered by concerns over data control, trust, and perceived risks, which makes firms cautious to share sensitive data. Confirming these views in supply chain context, Ströher et al. (2025) emphasize that traceability benefits for supply chain collaboration are clear but depend on high-quality inter-organizational data sharing. Brechtel (2025) further notes that many firms perceive inter-firm data sharing as a competitive threat if the value proposition is not clear.

Nonetheless, the traceability benefits that are put forward are relevant for product recalls. Similarly, Heeß et al. (2024) and Fani et al. (2025) argue that digital product passports and data-sharing architectures like blockchain can enhance transparency while protecting confidentiality. This could reduce fraud, strengthen information integrity, and improve trust in supply chains where product authenticity and safety are critical, which is essential for data sharing in recall cases. These studies show that inter-organizational data sharing is recognized to unlock potential benefits in identifying faulty products and managing recalls. However, they also emphasize that

adoption remains constrained by trust, governance, cost, and control concerns. Therefore, we want to pose the following research question: What is the impact data sharing can have on product recall costs?

2 Methods & Data

To provide an answer to the question if data sharing has an impact on recall cases, we are employing a multiple-case study (Yin et al., 2018) of four recall cases in two domains: 1) a candy manufacturer cannabis contamination case (Clapson, 2025) (candy CS1), 2) a supermarket blueberry contamination case (FAVV, 2025) (blueberry CS2), 3) a car manufacturer fire hazard case (Vanoverbeke, 2025) (motor CS3), and 4) an airbag manufacturer corrosion hazard case (Bernard, 2015) (airbag CS4).

Gathering data on each recall case we are do by a collection of news articles, official company communication, official supply chain partner communication, official government communication channels, data platforms (like the RASFF), online price catalogues, and academic studies performed on related recall cases. When dealing with multiple price points of a product in online catalogues, or varying ranges of for example impact on missed sales by reputation damage, averages are assumed. Such uncertainties in the data, are dealt with also by assuming they can happen, and as such some are nuanced in the discussion section to describe different scenarios.
In the four case studies, we are distilling costs, under the assumption that each recall case incurs similar types of costs (but with different values and nuances). From this we are building a cost breakdown shown in Table 1.

3 Preliminary results

The preliminary cost breakdown from the case studies is shown in Table 1. It includes the tangible costs of recall cases, the intangible costs of recall cases, and finally the costs specific to data sharing. The perspective taken for these costs is that the company to which the faulty product is attributed at first, as they are typically held responsible for first-line handling recalls in the supply chain.

Table 1: Cost breakdown of studied recall cases, including costs of the recall itself and costs that come with sharing data in recall cases

Tangibles	**Reimbursing customer**	reimbursement of product
		shipment cost
		payment processing
	Logistics of return	handling customer returns
		moving products
		destroying products
	Testing	test cost (assuming independent lab)
	Reputational damage	missed sales
		court case settlements
		patching reputation
Intangibles	**Reputational damage**	stock price plummets
		trustworthiness drops
		market share loss
	Employee wellbeing	employee morale
		employee retention
Cost of data sharing	**Revealing information**	product design secrets lost
		process optimization secrets lost
		supplier relationships getting poached
		market insights revealed
		suppliers adjusting prices based on info

Next to mapping the costs of the studied cases, the potential that could be reached with sharing data must be determined. The preliminary results on potential, shown in Table 2 presents a best-case scenario (precise recall) of what the minimum number of products that need to be recalled are. For each scenario, the characteristics of this case, allowing for potential improvement are discussed.

Table 2: Number of products that need to be recalled in the worst (overkill) and best (precise) case scenario

	candy CS1	blueberry CS2	motor CS3	airbag CS4
Overkill recall	≈ 50.000	≈ 101.300	≈ 332.500	≈ 100.000.000
Precise recall	≈ 3	≈ 10–90	≈ 3	≈ 1000

In CS1, specific batches were recalled of a candy product which we assumed to be of typical food industry batch size for a product of this weight and size. After independent investigation, the company producing the product found no contamination in their manufacturing process. There is potential for more precise recalls in this case as the only bags that needed recalling were thus the ones were

people reported feeling ill, which were only three. It remains unclear as to why the contamination occurred.

A similar logic was applied to CS2, as it was also a food industry recall, however actual contamination with Hepatitis A was found in several tens of cases in the Netherlands. Therefore, the number of products a precise recall would reduce is likely more limited. Nonetheless, assuming a positive scenario were only those tens of products turned out to be contaminated; there is still large potential. For example, as a safety precaution, the same product that was processed in the same timeframe of the contaminated ones in Belgium, were also recalled, while not finding any contamination. Therefore, we assume somewhere in the supply chain the contamination reached a split-off point or was physically constrained, which thus allows for potential improvement of the number of recalled products.

CS3 and 4 in turn being automotive cases, concerned faulty products from specific vendors. For both cases, we assume the best possible scenario is that mentioned in the Catena-X success story, being a reduction by factor of 100.000. For CS3, in Germany and the US combined, about 332.500 vehicles were recalled due to corrosion in the starter motor of the engine causing a fire hazard. No cases are reported of any engine failing, and as such this recall occurred out of precaution. The corrosion can occur due to faulty sealing or damaged protective coating during manufacturing, therefore, the best-case scenario will be a challenging one to achieve as likely each car with possibly fauly installtion needs to be checked by a mechanic.

In CS4 more than 100.000.000 vehicles are estimated to be recalled during several large recalls that are linked to the faulty airbags. Again, the potential can be challenging to achieve as each car that has an airbag of the bankrupted company, might get it replaced out of precaution anyways.

4 Discussion & preliminary conclusions

While a potential reduction of recalls seems possible, a discussion on preliminary conclusions if data sharing can offer a solution here, is appropriate.

To reach the potential in candy CS1 and blueberry CS2, for example, the source of the contamination, and its trace needs to be known. Data sharing can have potential impact here if it can show where contamination began and was halted. This would require all supply chain parties to test and share test results of handling of ingredients, packaging, and the final product. This would require substantial interorganizational trust among many parties as it requires sharing details on processes among competing suppliers. This again, potentially raises the costs of the risks of sharing data. Next to this, an additional issue could be the quality of testing, and consequently the quality of the data shared. Faulty testing could trigger false positives and negatives, with the first leading to inefficiency, and the later leading to possible health hazards, possibly diminishing the data sharing potential.

However, there are also conditions under which the risks of data sharing might be reduced. In CS2 for example, no data had to be shared with competing supermarket chains, and product design (being fruit, there is none) is not a crucial asset. As such the value of the data shared like test results among the supply chain and production processes is likely of less value to those involved in the recall, like farmers and distributors of fruit due to the non-competitive nature of the case. If the assumption is made that data can be shared in an obfuscated manner, the costs of data sharing risks are possibly reduced further. On the other hand, in motor CS3 and airbag CS4, product design is likely of high value, given they deal with car parts. Therefore, costs of sharing data might be significant if shared with competitors, and there might be reluctance to share product design info, even in obfuscated manner. The nature of CS3 and CS4, where it is unclear if a faulty installation occurred, would thus require extensive trace of every car part's assembly, raw materials and handling to reach its potential reduction of recalls.

We can preliminarily conclude that data sharing presents an interesting case for recall cases that have a non-competitive nature. Next to this, studying the impact of data sharing technology that preserves confidentially presents a research opportunity in recall cases with a competitive nature. However, further research is needed to determine whether the potential of data sharing, particularly within emerging initiatives such as data spaces, will materialize in practice, and thus lead to the associated public funding to indeed contribute to a more sustainable digital future.

5 Next steps

As our preliminary conclusions on the impact data sharing can have shows potential, but also challenges, we aim to further develop our cost model and structurally map recall cases on the model. This will allow both finetuning of the model and provide quantitative support for our arguments. We aim to quantify the costs of the different recall cases, as well as the risks of data sharing, so that both scenarios of overkill and precise recalls can be compared. The goal is to, by discovering patterns in recall cases, be able to distinguish recall cases with likely potential for data sharing, from those with high risks for data sharing. This must include mapping technology impact in a more structured manner to the recall cases than the short discussion currently provided on what data needs to be shared, and how this impacts the costs and risks of the recall cases.

Acknowledgment

This work was partially funded by UGent under BOF.BAF.2024.0953.01.

References

Agahari, W., Ofe, H., & de Reuver, M. (2022). It is not (only) about privacy: How multi-party computation redefines control, trust, and risk in data sharing. Electronic Markets, 32, 1577–1602. https://doi.org/10.1007/s12525-022-00572-w

Ammann, J., & Hess, T. (2025). To sell, to donate, or to barter? Value creation and capture through business model types in decentralized data ecosystems. Electronic Markets, 35, Article 36. https://doi.org/10.1007/s12525-025-00775-x

Automotive Industry Action Group. (2025, June 5). Quality recall reduced from 1.4 million to 14 vehicles. https://www.aiag.org/expertise-areas/data/aiag-catena-x-na-hub/real-world-use-case/blog-detail-catena-x/catena-x-real-world-use-case/2025/06/05/quality-recall-reduced-from-1.4-million-to-14-vehicles

Bernard, T. S. (2015). Navigating the Takata airbag recall. International New York Times.

Brechtel, M. (2025). From hesitation to participation in industrial data ecosystems: Analysis of motives and incentives in the automotive industry. Electronic Markets, 35, Article 73. https://doi.org/10.1007/s12525-025-00818-3

Clapson, C. (2025, June 2). Haribo recalls bags of 'Happy Cola F!ZZ' after cannabis found in Dutch confectionery. VRT NWS. https://www.vrt.be/vrtnws/en/2025/06/02/haribo-recalls-bags-of-happy-cola-f-zz-after-cannabis-found-in-d/

Djekic, I., Jankovic, D., & Rajkovic, A. (2017). Analysis of foreign bodies present in European food using data from Rapid Alert System for Food and Feed (RASFF). Food Control, 79, 143–149.

Fani, V., Ciccullo, F., Bandinelli, R., et al. (2025). Cultivating trust: An empirical exploration of blockchain's adoption within the Italian wine supply chain. Electronic Markets, 35, Article 35. https://doi.org/10.1007/s12525-025-00782-y

Fang, X., Wang, X., Shao, Y., & Banerjee, P. (2024). Examining the effect of a firm's product recall on financial values of its competitors. Journal of Business Research, 176, 114586.

Federaal Agentschap voor de Veiligheid van de Voedselketen. (2025, January 14). Terugroeping van Albert Heijn: Zakje met blauwe bessen – diepvries (AH). https://favv-afsca.be/nl/producten/terugroeping-van-albert-heijn-16

Heeß, P., Rockstuhl, J., Körner, M. F., et al. (2024). Enhancing trust in global supply chains: Conceptualizing Digital Product Passports for a low-carbon hydrogen market. Electronic Markets, 34, Article 10. https://doi.org/10.1007/s12525-024-00690-7

Kernstock, P., Harms, C., Hein, A., et al. (2025). Establishing and governing data ecosystems at the crossroads of centralization and decentralization. Electronic Markets, 35, Article 71. https://doi.org/10.1007/s12525-025-00810-x

Ketchen, D. J., Jr., Wowak, K. D., & Craighead, C. W. (2014). Resource gaps and resource orchestration shortfalls in supply chain management: The case of product recalls. Journal of Supply Chain Management, 50(3), 6–15.

Ströher, T., Körner, M. F., Paetzold, F., et al. (2025). Bridging carbon data's organizational boundaries: Toward automated data sharing in sustainable supply chains. Electronic Markets, 35, Article 33. https://doi.org/10.1007/s12525-025-00779-7

Vanoverbeke, L. (2025, September 26). BMW roept wereldwijd honderdduizenden auto's terug wegens brandgevaar. VRT NWS. https://www.vrt.be/vrtnws/nl/2025/09/26/bmw-roept-wereldwijd-honderdduizenden-auto-s-terug-wegens-brandg/

RESEARCH IN PROGRESS

CONSTRUCTING A COST–BENEFIT ANALYSIS FRAMEWORK FOR DIGITAL TWIN–ENABLED SPATIAL PLANNING: ARTICULATING (HIDDEN) COSTS AND (IMPLICIT) BENEFITS IN GOVERNMENTAL ORGANIZATIONS

SAM LEEWIS, KOEN SMIT, JOHN VAN MEERTEN, RONALD BUIJSSE

HU University of Applied Science Utrecht, Utrecht, the Netherlands
sam.leewis@hu.nl, koen.smit@hu.nl, john.vanmeerten@hu.nl, ronald.buijsse@hu.nl

Digital Twins (DTs) are increasingly used in urban planning and public-sector spatial development to support evidence-based decision-making and scenario analysis. However, many DT initiatives struggle to progress beyond pilot stages (Hiller et al., 2026), partly due to the absence of a clear and validated cost–benefit analysis (CBA) framework. Existing evaluations often underestimate lifecycle costs while failing to adequately articulate indirect and intangible benefits such as improved decision quality, stakeholder engagement, and transparency. This research-in-progress study tries to address this gap by asking how hidden costs and implicit benefits of DTs in public-area development can be identified and structured into a first version of a CBA framework. Using a Design Science approach, we combine a scoping review further supplemented with exploratory expert interviews in a Dutch spatial planning context. Preliminary findings highlight lifecycle and governance-related costs and indicate that DT value primarily emerges through improved decision quality and reduced downstream failure costs, though further detailing and validation is required in future research.

DOI
https://doi.org/10.18690/um.fov.4.2026.66

ISBN
978-961-299-152-4

Keywords:
cost-benefit analysis, spatial planning, digital twin technology, Bled eConference, research in progress

University of Maribor Press

1 Introduction

Digital twins (DTs) are increasingly adopted in urban planning and public-sector spatial development as instruments for evidence-based decision-making, simulation, and scenario analysis (Asian Development Bank, 2025). By creating dynamic, data-driven virtual representations of physical environments, DTs promise to improve infrastructure planning processes, decrease conflicts during realization, optimize spatial intervention results, and support the evaluation of complex policy trade-offs. Governments and municipalities across Europe, including in the Netherlands, are experimenting with DT applications to address societal challenges such as climate adaptation, mobility transitions, energy infrastructure, and integrated area development (Calzati, 2023), see e.g., Figure 1. Despite these ambitions, many DT initiatives struggle to move beyond pilot or prototype stages. A recurring explanation in both academic and practitioner literature is the absence of a clear and validated cost–benefit analysis (CBA) framework to support long-term investment decisions. While DTs are often framed as transformative technologies, public organizations face substantial uncertainty regarding both their full cost structures and their realized benefits (Hiller et al., 2026).

Figure 1: 3D Utrecht DT Platform (Gemeente Utrecht, 2026)

The literature consistently points to two interrelated challenges. First, many DT-related costs remain "hidden" or underestimated. Beyond technical development, organizations must account for data governance arrangements, cybersecurity, legal

compliance, training, change management, and long-term system evolution. Second, many of the anticipated benefits are implicit or difficult to quantify. These include improved interdepartmental collaboration, enhanced transparency in policy processes, increased stakeholder engagement, better risk awareness and risk prevention, less realization conflicts, better decision results and contributions to environmental resilience and social inclusion. These benefits also reflect underlying human and public values, such as fairness, transparency, inclusivity, and long-term societal well-being, that increasingly shape how digital infrastructures in the public domain are evaluated. Abbas et al. (2024), for instance, observe that most DT studies do not present quantifiable findings and call explicitly for comprehensive evaluation approaches, including formal cost–benefit analyses and clearer taxonomies of DT-specific cost and benefit categories. Without such structure, building a convincing business case remains difficult. In the Dutch context, similar patterns emerge. Municipalities and provinces are actively experimenting with DT solutions, yet these efforts often remain isolated and project-based. Studies of DT maturity in the Netherlands highlight the need for operationalized dimensions that extend beyond technical performance to include financial, operational, legal, policy, ethical, and societal considerations. Ecosystem analyses further describe a landscape characterized by strong technical innovation but limited integration of lifecycle value assessment in public-area development projects. As a result, DT initiatives frequently encounter difficulties in securing sustained political and financial support for scaling and institutionalization.

Taken together, these observations reveal a knowledge gap at the intersection of DT research and information systems evaluation. From an information systems perspective, DTs in governmental spatial planning should be assessed as socio-technical systems whose value depends not only on technical performance, but also on organizational embedding, user acceptance, governance arrangements, and broader decision-support and coordination benefits. Against this background, we address the following research question: How can both visible and hidden costs and both visible and often implicit benefits of DTs in public-area development be identified, articulated, and structured into a validated cost-benefit analysis framework? This Research-in-Progress paper adopts a design science perspective to lay the conceptual foundation for such a framework. By synthesizing existing research on DTs in urban and spatial planning and drawing on emerging Dutch practices, we aim to outline the core dimensions of a structured CBA framework

that integrates both technical and organizational factors, as well as tangible and intangible value components.

2 Background and Related Work

Hidden and Lifecycle Costs: economic assessments show that DT investments extend far beyond initial model development. Thomas (2024) argues that DT viability must be evaluated across the full lifecycle, including data acquisition, validation, cybersecurity, governance, and long-term system use and maintenance. Similarly, Green et al. (2025), though from McKinsey, demonstrate that while DTs may improve infrastructure ROI, realising such gains requires sustained investments in interoperability, data pipelines, and organisational transformation. Urban DT reviews identify persistent integration and standardisation challenges, including data fragmentation and cross-domain coordination (Mazzetto, 2024; Syed Abdul Rahman et al., 2024). In practice, governance embedding generates additional costs: aligning DT outputs with policy cycles and monitoring frameworks requires new indicators, coordination mechanisms, and capacity building (De Lathouwer et al., 2024).

Implicit and Public-Value Benefits: While costs are often underestimated, benefits are frequently under-articulated. International frameworks highlight DT potential for multistakeholder coordination, risk anticipation, and resilience planning (World Economic Forum, 2022). However, such gains materialise indirectly through improved decision quality and reduced uncertainty rather than immediate financial returns. Huang and Tsou (2025) conceptualise DTs as socio-technical systems enabling cross-domain optimisation, potentially generating systemic sustainability benefits. In the Dutch context, Smit et al. (2023) identify quality dimensions such as legal robustness, ethical safeguards, and public-value orientation, indicating that DT value extends beyond efficiency to include trust and accountability. Case research confirms improved transparency and stakeholder understanding, yet difficulties in translating these into formal investment logic (Wolters, 2023). The literature demonstrates that DTs are applied across diverse domains, including industrial manufacturing, asset management (e.g., building information modeling), and healthcare (Hasan & Crawford, 2025; Tudor et al., 2025). However, the underlying assumptions about value differ across these streams. Some studies adopt an efficiency-oriented perspective and treat DT value primarily in terms of return on investment, time savings, or operational optimization, whereas other contributions

emphasize lifecycle viability, governance complexity, and broader public-value effects such as transparency, accountability, and stakeholder inclusion. This creates a conceptual tension: what counts as a "benefit" in one framework may appear as a cost, constraint, or non-financial value dimension in another. In public spatial planning, this tension becomes particularly relevant because governmental DT applications are shaped not only by technical performance and financial considerations, but also by policy processes, interorganizational coordination, and societal legitimacy. The absence of a context-specific and validated CBA framework tailored to spatial planning by governmental organizations therefore motivates the present design science effort to systematically structure hidden costs and implicit benefits.

3 Research method & Preliminary Results

This research-in-progress study follows a design science approach (Hevner et al., 2004) to develop a CBA framework for DT-enabled spatial planning. We conducted a narrative review (Paré et al., 2015) to derive candidate cost and benefit categories. Within this review, we used Google Scholar as our search engine, utilizing "Cost AND Benefit AND Digital AND Twin" as our search query. Empirically, we currently conducted five exploratory semi-structured interviews with six experts (one interview had two experts present) from Dutch public and private organizations involved in DT initiatives in spatial planning. These experts were selected based on convinience sampling, but had multiple years of experience with building and implementing DTs in practice. We used direct coding in one round by one researcher focusing on identifying either candidate costs or benefits from both data sources. These were then reviewed by two reseachers, which led to the first iteration of the CBA presented in the next section.

Preliminary results

Current DT applications are primarily exploratory, utilizing platforms such as Tygron and Cesium, for environmental simulations (e.g., water and heat stress) during early-stage planning. A significant barrier to further formalization of a CBA is the scarcety of empirical, quantitative datasets, as most initiatives are funded through non-recurrent innovation budgets rather than operational cycles that operate long-term. Consequently, the value proposition of DTs currently shifts from

direct operational efficiency toward enhanced decision quality and integrated spatial insights. We present the first (conceptual) iteration of the CBA in Table 1 below.

Table 1: Preliminary results of the cost-benefit analysis regarding DT usage.

Category:	CBA Aspect:	Interviewees:	Supporting literature:
Benefits	Failure Cost Avoidance	Central benefit mechanism; reducing costly errors, redesign, and misalignment in later stages.	Mentioned in broader built-environment research, but under-specified in DT CBAs. (Karunaratne et al., 2025; Kookalani et al., 2026; Moshood et al., 2024)
	Temporal Efficiency	Shorter planning timelines leading to lower financing and interest costs.	Some DT case-based reports mention substantial schedule reductions (mainly in construction/infrastructure) (Omrany et al., 2023). Reviews of economic impact also stress that results are highly context-dependent and constrained by integration effort and adoption maturity (Karunaratne et al., 2025)
	Decision Quality	DTs deliver more complete, integrated, and transparent insights.	Frameworks frame value as improved decision support and planning insight (Karunaratne et al., 2025; Lehtola et al., 2022; Peldon et al., 2024).
	Stakeholder Support	Potential for greater public support and understanding.	Visualization can enhance involvement, but raises equity and accessibility concerns (Luo et al., 2025).
Costs	Data Management	Focus on acquisition, quality improvement, and long-term management.	Reviews identify fragmented data ecosystems and connectivity as major barriers (Lei et al., 2023; Mazzetto, 2024).
	Governance	Costs for interdepartmental coordination and multi-actor arrangements.	Implementation is constrained by governance, policy, and privacy requirements (Azadi et al., 2025).
	Infrastructure	Platform licensing and IT infrastructure.	Economic-impact review repeatedly cites high initial investment and integration complexity as obstacles (Karunaratne et al., 2025). Treat infrastructure cost as a major upfront/lifecycle item but not necessarily net-negative if maturity is high (Omrany et al., 2023).
	Capacity	Organizational capacity needed and to run simulations and interpret results.	Reviews note human-factor complexities and socio-technical constraints in DTs which commonly include organizational readiness/capability (Karunaratne et al., 2025; Lehtola et al., 2022; Lei et al., 2023).
Refinements	Temporal Scope	Need for a plan-formation phase before formal design and a benefit review during realization phases, as extra effort in planning (e.g. several simulations) results in large benefits in later area and building realization phases.	Literature describes DT use as supporting planning and design processes; governance/maturity literature analyses how DTs are operationalized in planning processes (often early-stage and iterative) (Dembski et al., 2020).

4 Discussion and Future Work

Our findings indicate that DT value in governmental spatial planning is unlikely to be captured by a narrow financial business case alone. Many relevant costs are lifecycle- and governance-related (data, integration, security, capacity), while many benefits materialize indirectly through improved decision quality and reduced downstream failure costs. This helps explain why DT initiatives often remain pilot-based as downstream results, generated by other organizations, are often not taken into account in pilots. The evidence base required for formal CBA is still limited in practice. The proposed framework contributes by structuring hidden costs and implicit benefits across all process phases and stakeholders, highlighting the distribution of value across municipalities, consultants, developers, users and maintenance organizations. The results suggest a need to reconceptualize IS evaluation of public-sector DTs beyond financial CBA to explicitly incorporate digital governance and public-value measurement. Limitations include the small purposive sample and the largely qualitative nature of reported effects. Next steps are to operationalize key categories into measurable indicators (e.g., failure-cost reduction, redesign frequency, lead-time and financing impacts, stakeholder comprehension) and validate the framework across multiple cases with different governance arrangements. Where monetization is premature, the framework should be complemented with multi-criteria or public-value measures to support transparent decision-making about the use of DT-technology.

Acknowledgments

We would like to thank the experts and organizations that participated in this study for their time and expertise.

References

Abbas, O. H., Moretti, N., & Broyd, T. (2024). Assessing the benefits of digital twins towards the formulation of whole life cycle value propositions: A systematic review. Computing in Construction, 5, 0–0. https://doi.org/10.35490/EC3.2024.304

Asian Development Bank. (2025). Digital Twin Framework: A Practical Guide (0 edn). Asian Development Bank. https://doi.org/10.22617/SPR250158

Azadi, S., Kasraian, D., Nourian, P., & Van Wesemael, P. (2025). Formalizing-modelling-utilizing ontology: A semantic framework for adaptive stakeholder-specific urban digital twins in urban planning processes. Environment and Planning B: Urban Analytics and City Science, 23998083251329398. https://doi.org/10.1177/23998083251329398

Calzati, S. (2023). No longer hype, not yet mainstream? Recalibrating city digital twins' expectations and reality: a case study perspective. Frontiers in Big Data, 6, 1236397. https://doi.org/10.3389/fdata.2023.1236397

De Lathouwer, B., van Putten, G., Horbach, S., & Grothe, M. (2024). Policy processes and building blocks for Digital Twins.

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M., & Yamu, C. (2020). Urban digital twins for smart cities and citizens: The case study of herrenberg, germany. Sustainability, 12(6), 2307. https://doi.org/10.3390/su12062307

Green, A., Cosmas, A., Shenai, G., Moore, D., & Mehta, N. (2025). Digital twins: Boosting ROI of government infrastructure investments. https://www.mckinsey.com/industries/public-sector/our-insights/digital-twins-boosting-roi-of-government-infrastructure-investments

Hasan, S. K. K., & Crawford, C. W. (2025). A new era for digital twins: Progress and industry adoption. Digital Twin, 2(4), 2555877. https://doi.org/10.1080/27525783.2025.2555877

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems Research1. Management Information Systems Quarterly, 28(1), 75–106. https://doi.org/10.2307/25148625

Hiller, J., Mansour, M., Kremer, N., Crampen, D., & Von Behren, S. (2026). Mapping Urban Digital Twins Across Regions: An Exploratory Study of Maturity, Implementation Status, and Authority. Smart Cities, 9(3), 49. https://doi.org/10.3390/smartcities9030049

Karunaratne, T., Ajiero, I. R., Joseph, R., Farr, E., & Piroozfar, P. (2025). Evaluating the economic impact of digital twinning in the AEC industry: A systematic review. Buildings, 15(14), 1–29. https://doi.org/10.3390/buildings15142583

Kookalani, S., Green, S., Luo, P., Alavi, H., Parn, E., Sun, Z., & Brilakis, I. (2026). Mapping digital twin applications in infrastructure and the built environment across research types, methods, sectors, phases, and scales. Automation in Construction, 182, 106778. https://doi.org/10.1016/j.autcon.2026.106778

Lehtola, V. V., Koeva, M., Elberink, S. O., Raposo, P., Virtanen, J.-P., Vahdatikhaki, F., & Borsci, S. (2022). Digital twin of a city: Review of technology serving city needs. International Journal of Applied Earth Observation and Geoinformation, 114, 102915. https://doi.org/10.1016/j.jag.2022.102915

Lei, B., Janssen, P., Stoter, J., & Biljecki, F. (2023). Challenges of urban digital twins: A systematic review and a delphi expert survey. Automation in Construction, 147, 104716. https://doi.org/10.1016/j.autcon.2022.104716

Luo, J., Liu, P., Kong, X., Shen, J., Wu, Q., & Xu, D. (2025). Urban digital twins for citizen-centric planning: A systematic review of built environment perception and public participation. International Journal of Applied Earth Observation and Geoinformation, 143, 104746. https://doi.org/10.1016/j.jag.2025.104746

Mazzetto, S. (2024). A review of urban digital twins integration, challenges, and future directions in smart city development. Sustainability, 16(19), 1–33. https://doi.org/10.3390/su16198337

Moshood, T. D., Rotimi, J. OB., Shahzad, W., & Bamgbade, J. A. (2024). Infrastructure digital twin technology: A new paradigm for future construction industry. Technology in Society, 77, 102519. https://doi.org/10.1016/j.techsoc.2024.102519

Omrany, H., Al-Obaidi, K. M., Husain, A., & Ghaffarianhoseini, A. (2023). Digital twins in the construction industry: A comprehensive review of current implementations, enabling technologies, and future directions. Sustainability, 15(14), 1–26. https://doi.org/10.3390/su151410908

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. Information & Management, 52(2), 183–199. https://doi.org/10.1016/j.im.2014.08.008

Peldon, D., Banihashemi, S., LeNguyen, K., & Derrible, S. (2024). Navigating urban complexity: The transformative role of digital twins in smart city development. Sustainable Cities and Society, 111, 105583. https://doi.org/10.1016/j.scs.2024.105583

Smit, K., Peters, R., de Gelder, C., & Versendaal, J. (2023). Quality dimensions for digital twin maturity in the context of dutch public spatial planning. 154. https://doi.org/10.18690/um.fov.6.2023.9

Syed Abdul Rahman, S. A. F., Abdul Maulud, K. N., Ujang, U., Wan Mohd Jaafar, W. S., Shaharuddin, S., & Ab Rahman, A. A. (2024). The digital landscape of smart cities and digital twins: A systematic literature review of digital terrain and 3D city models in enhancing decision-making. Sage OPEN, 14(1), 21582440231220768. https://doi.org/10.1177/21582440231220768

Thomas, D. (2024). Economics of digital twins: Costs, benefits, and economic decision making (NIST AMS 100-61; p. NIST AMS 100-61). National Institute of Standards and Technology (U.S.). https://doi.org/10.6028/NIST.AMS.100-61

Tudor, B. H., Shargo, R., Gray, G. M., Fierstein, J. L., Kuo, F. H., Burton, R., Johnson, J. T., Scully, B. B., Asante-Korang, A., Rehman, M. A., & Ahumada, L. M. (2025). A scoping review of human digital twins in healthcare applications and usage patterns. Npj Digital Medicine, 8(1), 587. https://doi.org/10.1038/s41746-025-01910-w

Utrecht in 3D | gemeente utrecht. (2026). https://www.utrecht.nl/bestuur-en-organisatie/publicaties/utrecht-in-3d

Wolters, B. M. (2023). Exploring the potential of urban digital twins in climate adaptive development. 1–125.

World Economic Forum. (2022). Digital Twin Cities: Framework and Global Practices.

RESEARCH IN PROGRESS

EVALUATING THE FEASIBILITY OF LLM-BASED AUTOMATION OF MANUAL WCAG COMPLIANCE TESTING

ANNEMAE VAN DE HOEF, KOEN SMIT, SAM LEEWIS, DUAIN CASTRO, FABIAN HARTMAN, JULIANA TODOROVA, NICK KUIPER

HU University of Applied Science Utrecht, Utrecht, the Netherlands
annemae.vandehoef@hu.nl, koen.smit@hu.nl, sam.leewis@hu.nl, duain.castro@hu.nl, fabian.hartman@hu.nl, juliana.todorova@hu.nl, nick.kuiper@hu.nl

Digital inclusion requires that websites and online services are accessible to people with disabilities, yet verifying compliance with the Web Content Accessibility Guidelines (WCAG) remains resource-intensive. Many success criteria require manual auditing, as conventional automated tools detect only a limited subset of violations. This study evaluates the feasibility of using a fine-tuned enterprise large language model (LLM) to (semi-)automate such manual WCAG checks. We developed a Chrome extension that operationalizes selected criteria through structured prompting and tested it against prior expert audits of two benchmark pages. Results show moderate ability to localise problematic code, but weak performance in correctly assigning WCAG criteria and generating reliable improvenent suggestions. While the proof-of-concept demonstrates potential as a audit support tool, its probabilistic outputs currently lack the stability and accuracy required to properly support human accessibility auditing.

DOI
https://doi.org/10.18690/um.fov.4.2026.67

ISBN
978-961-299-152-4

Keywords:
large language models, WCAG, compliance testing, digital accessibility, research in progress

University of Maribor Press

1 Introduction

Web accessibility is essential for ensuring that digital content is usable by people with disabilities, which is an important aspect of a digital inclusive society. Governments and organizations worldwide mandate conformance to the Web Content Accessibility Guidelines (WCAG) (e.g. AA level for most regulations) (see e.g. (European Commission, 2025)). However, achieving and verifying WCAG compliance remains challenging. Full manual auditing by experts (often including people with lived disability experience) is time-consuming, costly, and requires specialized skills. For example, a comprehensive accessibility audit of a website can take hours or days, depending on complexity (López-Gil & Pereira, 2025; Smit et al., 2024). Human auditors may suffer fatigue or overlook issues during repetitive testing of dozens of pages and criteria, as well as the variability across raters when applying WCAG success criteria (Iniesto & Rodrigo, 2024). In contrast, automated tools can scan many pages quickly and consistently, but they only cover a subset of WCAG requirements. However, prior empirical data indicates that conventional scanners reliably detect perhaps only 15–30% of WCAG criteria, while the rest require human judgment (Fischer et al., 2025; Heine, 2025) and coverage was and is still narrow on certain success criteria (Vigo et al., 2013). The risk is, that in practice, a purely automated scan might report a perfect score, whereas a human audit uncovers many issues. Combining both is reported to be favorable.

Given these limitations, human auditors remain the gold standard. Humans can interpret semantics (e.g., whether alt text is meaningful) and assess usability beyond code patterns. But relying solely on experts is expensive and not scalable. To reduce this burden, researchers have investigated hybrid approaches. In particular, recent advances in large language models (LLMs, e.g., GPT-4) suggest a new avenue. LLMs, with their ability to parse and reason about text and code, could help (semi-)automate WCAG checks that currently require manual review. LLMs trained on vast web corpora have demonstrated surprising abilities on diverse tasks, and some studies have begun to apply them to accessibility. For example, López-Gil and Pereira (2025) used an LLM to evaluate three WCAG success criteria (Non-text Content 1.1.1, Link Purpose 2.4.4, Language of Parts 3.1.2) and found that their LLM-based scripts identified issues that traditional tools missed, achieving 87.2% detection across test cases. Similarly, Hinderks et al. (2025) compared GPT-4o to Lighthouse and reported that, with tailored prompts, the LLM found 88% of

induced violations on representative WCAG 2.1 criteria versus only 24% by Lighthouse. Andruccioli et al. (2025) showed that an LLM could reason about dynamically generated HTML and flag accessibility violations beyond the reach of static analyzers, though it also produced false positives ("hallucinated" issues). These early findings suggest LLMs could augment automated testing by bringing semantic understanding to the task (Bassi et al., 2025). To the knowledge of the authors, only the few selected studies identified in this paper focus on applying LLMs for accessibility testing.

However, LLMs are inherently probabilistic and not designed as validators. Unlike rule-based scanners, their outputs can vary with wording and context, and they may "hallucinate" or give inconsistent answers (Li et al., 2025). For example, AI-driven scans might incorrectly flag non-issues or miss subtle errors, requiring extra scrutiny. The question remains as follows: To what extent can a fine-tuned large language model reliably automate manual WCAG 2.x (AA-level) conformance checks? To answer this question, we conducted a feasibility study, for which we have developed a proof-of-concept (PoC) using a fine-tuned LLM to automate WCAG 2.x (primarily AA-level) checks.

2 Background and Related Work

Research on WCAG conformance testing can broadly be grouped into three methodological streams: 1) deterministic rule-based automation, 2) hybrid or AI-augmented extensions of static analysis, and 3) more recent large language model (LLM)-driven evaluative approaches. These streams differ not only in technical implementation but also in their assumptions about what constitutes testable accessibility knowledge.

The first stream consists of rule-based automated tools that operationalize WCAG success criteria as formalized code checks. Applications such as AXE and WAVE analyze the Document Object Model (DOM) and apply predefined rules to detect structural violations (Deque, 2026; WEBAIM, 2026). Their strength lies in determinism and reproducibility, given identical input, the output remains stable and predictable. However, this approach is inherently constrained to criteria that can be expressed as syntactic or programmatic rules. Thus, traditional automation is bounded less by computational power than by formal expressibility. A second

stream attempts to extend automation through hybrid or AI-supported methods. These approaches incorporate machine learning, dynamic rendering analysis, or heuristic reasoning to address limitations of purely static inspection. For example, Andruccioli et al. (2025) demonstrate that model-based systems can analyze dynamically generated HTML and identify accessibility violations that static tools miss (Andruccioli et al., 2025). While such systems broaden technical coverage, they still typically operate within constrained detection paradigms and do not fully replicate the contextual reasoning characteristic of expert human auditors. The third and most recent stream introduces LLMs as evaluative agents in accessibility testing. Unlike rule-based systems, LLMs do not rely on explicitly encoded validation rules; instead, they perform probabilistic inference based on learned representations of language, structure, and context. Empirical studies suggest that LLMs can outperform traditional automated tools on semantically complex WCAG criteria, such as evaluating the meaningfulness of alternative text or the contextual clarity of link purpose (Hinderks et al., 2025; López-Gil & Pereira, 2025). This positions LLMs as potentially capable of automating checks previously categorized as "manual."

However, this paradigm shift introduces methodological and maybe even ethical challenges. Because LLMs are probabilistic, their outputs may vary across runs, depend on prompt design, and occasionally generate hallucinated or unjustified findings (Andruccioli et al., 2025). In compliance-driven domains such as accessibility auditing, where reproducibility, traceability, and defensibility are essential, such variability raises concerns about reliability (Andruccioli et al., 2025; The American Foundation for the Blind, 2025). Consequently, the key research gap is not merely whether LLMs can detect violations, but whether fine-tuned enterprise-scale LLMs can do so with stability and accuracy comparable to human auditors under controlled conditions.

3 Research Method and Preliminary Results

This study adopts a design science approach (Hevner et al., 2004) to develop and evaluate a PoC Chrome extension that automates parts of manual WCAG 2.2 (AA) conformance checking using a locally deployed enterprise LLM. The extension scans a web page, extracts relevant DOM and textual context, and queries the model with criterion-specific prompts. A locally hosted version of Copilot was used to meet institutional security and privacy requirements. Implementation details and prompt

templates are documented in the project repository (Crasto et al., 2026), see Figure 1 for examples of the PoC after the first iteration.

Figure 1: Filter before starting test (left) and results after test (right)

The PoC targets success criteria that typically require manual auditing beyond rule-based tools (Vigo et al., 2013), listed in our dataset (Crasto et al., 2026). Each selected WCAG criterion was translated into a structured prompt containing: (1) a concise guideline description and reference, (2) a task specification, (3) common failure patterns, and (4) explicit pass/fail conditions (White et al., 2023). The model was required to return the relevant DOM location, the WCAG reference, a justification grounded in the guideline, and a suggestion on how to improve the code. Temperature settings were calibrated per analytical category to reduce output variability (lower for structural checks; slightly higher for context-sensitive checks) (Brown et al., 2020). Evaluation used two organizational web pages from the HU University of Applied Sciences that had previously undergone comprehensive manual WCAG audits, which served as the reference standard. For each page, the extension performed a full scan and generated criterion-specific findings. Outputs were evaluated along two dimensions: (1) correct classification of issues (location and WCAG success criterion) and (2) the quality of justification and suggested improvements. For each manually identified issue, it was assessed whether the LLM detected the same issue. Where outputs contained code, matching was performed using normalized HTML snippets to remove variable elements. Exact matches indicated correct localization, while partial matches were manually reviewed for semantic equivalence. If code matched, the corresponding WCAG criterion was verified. For results without explicit code, equivalence was assessed at the guideline

level based on substantive agreement. For correctly classified issues, the LLM's justification was evaluated against the relevant WCAG criterion, followed, if correct, by validation of the suggested improvement. Only criteria for which the manual audit identified at least one issue were included in performance calculations. Precision, recall, and F1-score were computed at criterion level (Sokolova & Lapalme, 2009). For findings where both location and criterion matched the manual audit, we additionally assessed the correctness of the model's justification and improvement advice.

Preliminary results

Across both benchmark pages, the PoC showed moderate ability to localise problematic code but substantially weaker performance in assigning the correct WCAG criterion, see Table 1.

Table 1: Testing results for both pages included in the first iteration of testing the PoC

Performance Dimension	Metric	Page 1 (ÉÉN HU)	Page 2 (HU Voltijdopleidingen)
A. Localisation Accuracy	% Correct Faulty DOM Locations Identified	57.8	60.4
B. Criterion Classification	% Correct WCAG Criterion Assignment	20.3	17.9
	% Precision	4.5	2.4
	% Recall	8.0	17.9
	% F1-score	5.7	4.2
C. Justification Quality*	# Fully Correct Justifications	3	89*
D. Improvement Quality*	# Fully Correct Improvement Suggestions	1	1

*Most correct justifications referred to the same recurring accessibility issue, namely a "favorite" button missing the required aria-checked attribute. Qualitative inspection suggests recurring mismatches between identified symptoms and the precise WCAG success criterion, as well as occasional hallucinated technical details. Overall, the PoC appears promising as a support tool for identifying potentially problematic code regions, but in its current configuration it does not yet provide sufficiently reliable criterion classification or remediation guidance to support or automate human auditing adequately.

4 Discussion and Future Research Directions

These preliminary findings suggest that LLM-based support for manual WCAG auditing is currently more promising for localising potentially problematic code than for making reliable compliance judgements. The PoC showed moderate localisation performance, but weak criterion assignment and limited remediation quality, indicating that the model may detect accessibility-relevant symptoms without consistently translating them into precise WCAG judgements. This aligns with earlier studies showing that LLMs can extend accessibility testing beyond rule-based tools, while still producing unstable or hallucinated results (Andruccioli et al., 2025; Hinderks et al., 2025; López-Gil & Pereira, 2025). At this stage, the PoC therefore seems better suited as a triage aid for human auditors than as a means of automating manual conformance assessment. From an IS perspective, this also suggests that effective AI-supported accessibility auditing depends on socio-technical alignment between AI outputs, human judgement, and compliance governance (Ahmi & Smidt, 2026). The study is limited by the small benchmark set, the time gap between manual and automated audits, and the use of prompting rather than fine-tuning. In addition, prompt calibration could not be performed due to the extensive processing time required for the LLM to handle a single page, combined with the need for subsequent manual validation. Future research should therefore test larger and concurrent datasets, separate localisation from criterion classification, explore criterion-specific or fine-tuned models, and examine whether multi-pass prompting and human-in-the-loop review improve reliability and practical usefulness.

References

Ahmi, A., & Smidt, L. (2026). AI-driven auditing: Trends, themes, and research trajectories. *EDPACS*, 1–20. https://doi.org/10.1080/07366981.2026.2637772

Andruccioli, M., Bassi, B., Delnevo, G., & Salomoni, P. (2025). Leveraging Large Language Models for Sustainable and Inclusive Web Accessibility. *Big Data and Cognitive Computing*, *9*(10), 247. https://doi.org/10.3390/bdcc9100247

Bassi, B., Delnevo, G., Franco, M., Gaggi, O., Gatto, S., Mirri, S., & Olaiya, K. (2025). An Assessment of LLM-Based Auditing and Validation for Web Accessibility. *Proceedings of the 2025 International Conference on Information Technology for Social Good*, 297–305. https://doi.org/10.1145/3748699.3749805

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners* (arXiv:2005.14165). arXiv. https://doi.org/10.48550/arXiv.2005.14165

Crasto, D., Kuiper, N., Hartman, F., & Todorova, J. (2026). *WCAG Tool* [Computer software]. https://github.com/wp7HUDT/AccessibilityCheck

Deque. (2026). The Axe Platform provides all your accessibility testing and monitoring tools in one place. *Deque*. https://www.deque.com/axe/

European Commission. (2025). *Web Accessibility Directive—Standards and harmonisation*. European Commission. https://digital-strategy.ec.europa.eu/en/policies/web-accessibility-directive-standards-and-harmonisation

Fischer, T., Lundell, B., & Gamalielsson, J. (2025). Coverage of web accessibility guidelines provided by automated checking tools. *Universal Access in the Information Society*, *24*(4), 3615–3637. https://doi.org/10.1007/s10209-025-01263-x

Heine, M. (2025, February 28). *Why Automated Accessibility Testing Isn't Enough*. B13. https://b13.com/blog/why-automated-accessibility-testing-isnt-enough

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–106. https://doi.org/10.2307/25148625

Hinderks, A., Davtyan, A., & Allerkamp, D. (2025). Automated accessibility testing with Large Language Models: A comparative analysis of GPT-40 and traditional tools for WCAG 2.1 conformance testing. *Mensch Und Computer 2025 - Usability Professionals*. Gesellschaft für Informatik eV.

Iniesto, F., & Rodrigo, C. (2024). The use of WCAG and automatic tools by computer science students: A case study evaluating MOOC accessibility. *JUCS - Journal of Universal Computer Science*, *30*(1), 85–105. https://doi.org/10.3897/jucs.101704

Li, Q., Leyang, C., Lingpeng, K., & Wei Bi. (2025). Exploring the Reliability of Large Language Models as Customized Evaluators for Diverse NLP Tasks. *Proceedings of the 31st International Conference on Computational Linguistics*, 10325–10344.

López-Gil, J.-M., & Pereira, J. (2025). Turning manual web accessibility success criteria into automatic: An LLM-based approach. *Universal Access in the Information Society*, *24*(1), 837–852. https://doi.org/10.1007/s10209-024-01108-z

Smit, K., Van De Hoef, A., Kuiper, N., Todorova, J., Hartman, F., Leewis, S., & Gevaert, R. (2024). Digital Accessibility in Higher Education: A Dutch case study. *Proceedings of the 2024 8th International Conference on Software and E-Business*, 76–86. https://doi.org/10.1145/3715885.3715890

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

The American Foundation for the Blind. (2025). *Missing a Human Touch: Why Automated Testing and AI Don't "Solve" Digital Accessibility… Yet*. The American Foundation for the Blind. https://afb.org/blog/entry/missing-human-touch-1

Vigo, M., Brown, J., & Conway, V. (2013). Benchmarking web accessibility evaluation tools: Measuring the harm of sole reliance on automated tests. *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility*, 1–10. https://doi.org/10.1145/2461121.2461124

WEBAIM. (2026). *WAVE Web Accessibility Evaluation Tools*. WAVE. https://wave.webaim.org/

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT* (arXiv:2302.11382). arXiv. https://doi.org/10.48550/arXiv.2302.11382

RESEARCH IN PROGRESS

SUBJECTIVE KNOWLEDGE AND AI PRIVACY RISKS

MARI PAALIMÄKI, JUHA MUNNUKKA

University of Jyväskylä, School of Business and Economics, Jyväskylä, Finland
mari.s.paalimaki@jyu.fi, juha.t.munnukka@jyu.fi

Artificial Intelligence (AI) systems are increasingly used in everyday decision-making, but the adaptive and opaque nature makes them difficult to understand, creating uncertainty in evaluating risks. Prior research has examined privacy risk through factors such as trust and perceived control but lacks an explanatory mechanism for how individuals assess such risk in AI context. This study explores the role of subjective knowledge, defined as individuals' perceived understanding of a system, in shaping privacy risk perception using a pilot experiment with two modalities (text vs. voice). The results suggest that subjective knowledge is associated with privacy risk perception and usage intention, providing preliminary insight of the underlying mechanism and highlighting its importance in AI-based interaction.

DOI
https://doi.org/10.18690/um.fov.4.2026.68

ISBN
978-961-299-152-4

Keywords:
AI assistant, modality, subjective knowledge, privacy risk, intention to use

University of Maribor Press

1 Introduction

The use of AI assistants in human decision-making has become increasingly common. However, the use of AI-based assistants involves many unidentified privacy risks due to their opaque and adaptive nature (Liu & Shi, 2025; Robin & Dandis, 2021) which may increase users' sense of uncertainty (Schneiderman, 2021). As uncertainty increases, the assessment of risks associated with using the system may increasingly rely on users' own perceptions of how the system works. Subjective knowledge is an individual's experience of their own knowledge and based on interpretations, observations, and experiences (Brucks, 1985; Flynn & Goldsmith, 1999). Thus, under conditions of uncertainty and complexity, subjective knowledge may play a vital role in how individuals evaluate privacy risks in using AI assistants.

Prior research on risk perception suggests that individuals often rely on heuristic and subjective interpretations when evaluating risks (Slovic, 1987). In the context of privacy, individuals' risk evaluations are shaped by their perceptions on concerns regarding data collection and use (Dinev & Hart, 2006; Malhotra, Kim & Agarwal, 2004; Xu, Dinev, Smith & Hart, 2011). In addition, prior studies have focused on privacy risk perception mechanisms with factors such as privacy concerns, trust, control, vulnerability, and transparent data collection (Karwatzki, Dytynko, Trenz & Veit, 2017; Liu et al., 2024; Robin et al., 2021). However, the role of subjective knowledge in privacy risk perception has received limited attention in previous privacy literature. Therefore, the significance of subjective knowledge in risk recognition in AI systems is noteworthy, as the inherent opacity of these systems makes it even more difficult to anticipate risks.

This pilot study aims to understand the role of subjective knowledge in assessing privacy risks in the use of AI assistants. This research is highly topical as the predictability of risks associated with AI systems is weak due to their adaptive nature. Study explores a mechanism whereby subjective knowledge shapes the privacy risks assessment in AI interactions and how it is associated with the intention to use AI assistant. Specifically, the study examines whether modality (text vs. voice) shapes this mechanism by influencing subjective knowledge. The pilot study provides preliminary findings of the effectiveness of the connections and enables the evaluation of the indicators used.

2 Conceptual framework

According to behavioral science, human reasoning and decision-making involve two information processing systems (e.g., Kahneman & Frederick, 2005; Schneider & Shiffrin 1977a, 1977b). System one is an automated processing mode with fast, intuitive, and reactive while system two is slower and reflective (Kahneman et al., 2003). A key concept related to intuitive information processing is processing fluency, which refers to the subjective experience of how effortless it is to understand information (Reber, Wurtz & Zimmermann, 2004). When information processing is fluent, people tend to rely on quick assessments (Kunst-Wilson & Zajonc, 1980). However, the fluency of information processing, combined with familiarity and repetition, can also lead to biases in knowledge (Alba & Hutchinson, 2000; Flynn & Goldsmith, 1999; Reber et al., 2004; Zajonc, 1968). This fluency-inducted bias in knowledge is particularly relevant in AI context, where system complexity makes objective evaluation difficulty.

Perceived understanding is closely related to the concept of subjective knowledge. According to Brucks (1985), knowledge that drives behavior can be divided into objective knowledge, i.e., what individual knows, and subjective knowledge which refers to individual's perception of their own knowledge. Subjective knowledge can strongly influence decision-making regardless of actual knowledge (Brucks, 1985; Flynn et al., 1999). When individuals perceive their understanding high, they tend to rely less on cognitive effort to process information and instead rely incorrectly on their own understanding (Alba et al., 2000; Brucks, 1985). Consequently, subjective knowledge may reinforce reliance on system one processing.

According to Hall and Mast (2007) modality refers to the channel of interaction through which information is interpreted. Different modalities can shape how individuals process information. Voice modality produces more spontaneous interpretations because it reduces cognitive load, while text modality requires the user to engage in more analytical (Hall et al., 2007; Poushneh & Gearhart, 2025). Thus, voice modality can be more interpretable and thus activates the quick information processing system one and this experience of fluency may reduce the need for analytical processing. In addition, when information is processed fluently, individuals may develop stronger perceptions that they understand the system. This perceived understanding can increase individuals' subjective knowledge, which in

turn may influence how they evaluate potential risks associated with the system. This suggests that modality does not merely shape information processing but also indirectly conditions how individuals assess associated privacy risks.

In risk assessment, risk predictability plays a significant role in risk identification (Slovic,1987). When risk is impossible to identify, an individual's uncertainty increases and often leads to distorted risk assessments (Lerner et al., 2000; Slovic, 1987). In AI-based systems, data processing is opaque for users, making privacy risks particularly difficult to identify. Prior studies have shown that perceived privacy risks negatively impact on behavioral intention (Alrawad et al., 2023; Song, Xing, Duan, Cohen & Mou, 2022). The use of AI-based systems involves data loss risks as well as psychological privacy risks, such as anxiety and control, and freedom-related privacy risks, such as manipulation by algorithms (Karwatzki, Trenz & Veit, 2022; Pizzi, Scarpi & Pantano, 2021). However, these invisible risks are often misjudged, which is why the nature of the risk becomes a matter of subjective interpretation (Mitchell, 1999). Therefore, subjective knowledge may create perception of predictability and thus plays a significant role in explaining the perception of privacy risks in complex AI systems. Based on the above literature, the following hypothesis are proposed:

- H1. Modality is associated with subjective knowledge
- H2. Subjective knowledge is negatively associated with privacy risk perception
- H3. Perceived privacy risk is negatively associated with intention to use

3 Methodology

The pilot study was conducted on Qualtrics online survey platform as a scenario-based pilot experiment in which respondents (n=86) from Prolific were randomly assigned to either a text (group 1, n=40) or voice (group 2, n=46) scenario group. The aim was to evaluate scenario effectiveness, assess indicator reliability, and conduct a preliminary examination of construct relationships. Attention check questions were included to ensure that participants understood what they were responding to. Demographic data will be collected in the main study.

Measures were adapted from existing scales on a 7-point Likert scale. Subjective knowledge was measured with three items (Flynn et al., 1999), privacy risk with seven items (Karwatzki et al., 2022), intention to use with three items (Hoehle & Venkatesh, 2015). The independent variable (IV) was formed by coding the test groups with dichotomous indicators (0=text-based mode, 1=voice-based mode). Processing fluency, used as manipulation check, was measured with three items (Alter & Oppenheimer, 2009; Reber et al., 2004; Song and Schwarz, 2008.) The data was analyzed using SPSS. The effectiveness of the manipulation was assessed using T-test, which calculated the differences in modality and fluency. Linear regression analyses were performed to verify the connection between variables.

4 Results

Table 1: Descriptive statistics and correlations

Variable	α	Mean	SD	1. Group	2. SK	3. PR	4. INT
1. Group	-	1,5	.502	-			
2. SK	.792	4.39	1.195	.230**	-		
3. PR	.861	2.89	1.040	-.111	-.360***	-	
4. INT	.956	5.63	1.318	.224**	.352***	-.427***	-

**p < .005, ***p < .001, n=86 (group 1. text, n =40, group 2. voice, n= 46)
SK=Subjective knowledge, PR=Privacy risk, INT=Intention to Use

Manipulation check. The manipulation check did not produce differences in the manipulation check variable between the test groups (text: M=5.13, SD=1.21, voice: M=5.33, SD=1.06, t(78) =-.781, p=.437). This suggests that modality effects may be more subtle of require refine operationalization. However, differences were observed in subjective knowledge (text: M=4.09, SD=1.18, voice: M=4.64, SD=1.16, t(84) =-2.162, p=.017) and intention to use (text: M=5.31, SD=1.4, voice: M=5.91, SD=1.2, t(84) =-2.095, p=.039). This suggests that modality can be linked to various interpretation processes that are not necessarily directly related to fluency. As presented in Table 1, the reliability of the measures was tested using Cronbach's alpha coefficient, and all measures exceeded the criteria of $\geq$ 0.7. Correlations were observed between the variables (see Table 1). The group (IV, 0=text-based group,

1=voice-based group) was positively correlated with subjective knowledge (r=.230, p=.033) and intention to use (r=.224, p=.038). In addition, subjective knowledge correlated negatively with perceived privacy risk (r=-.360, p<.001) and positively with intention to use (r=.352, p<.001). Moreover, perceived privacy risk correlated negatively with intention to use (r= -.427, p<.001).

Hypotheses testing: Regression analyses were performed to confirm the constructional logic of the mechanism. Modality (IV) was associated with subjective knowledge (β =.230, p=.033, R^2 =.042), subjective knowledge was negatively associated with privacy risk (β= -.360, p<.001, R^2=.119), and privacy risk negatively predicted intention to use (β= -.427, p<.001, R^2=.172). Thus, all hypotheses were supported (see Figure 1). Additionally, subjective knowledge was directly associated with intention to use (β=.352, p<.001, R^2 =.113), suggesting a potential mediating role of privacy risk. In further explore this, we tested the combined association of subjective knowledge (β=.228, p=.031, R^2 =.208) and privacy risks (β= -.345, p<.001, R^2=.208) on the intention to use and noticed that the association of subjective knowledge on the intention to use weakens when privacy risk is included in the model. Thus, this predicts that subjective knowledge is associated with the intention to use directly and indirectly through privacy risk.

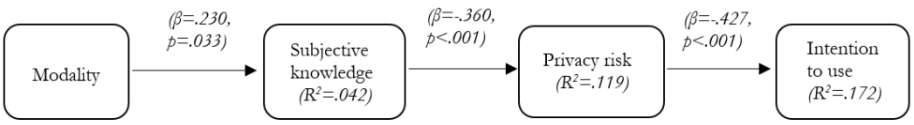

Figure 1: Research model with regression coefficients

5 Discussion

The preliminary findings provide initial insight into the relationships between modality, subjective knowledge, privacy risks, and intention to use. These findings are consistent with prior literature on subjective knowledge in risk perception (Brucks, 1985; Flynn & Goldsmith, 1999) and perceived privacy risk in behavioral intention (Alrawad et al., 2023; Song et al., 2022). The results offer tentative support for all three hypothesized relationships, suggesting that subjective knowledge plays a role in privacy risk evaluation and intentions to use AI assistants. Additionally, the findings hint at a potential mediation mechanism, where subjective knowledge

influence intention to use both directly and indirectly through privacy risk, which requires further research in the main study. Although manipulation did not produce the intended difference between conditions, the pilot study helped refine the experimental setup and provides a foundation for the main study. The main study will refine the manipulation, adjust the measured variables, and collect a larger sample size to test the proposed relationship more rigorously.

References

Alba, J., & Hutchinson, J. (2000). Knowledge Calibration: What Consumers Know and What They Think They Know. *The Journal of Consumer Research*, 27(2), 123-156. doi:10.1086/314317

Alrawad, M., Lutfi, A., Alyatama, S., Al Khattab, A., Alsoboa, S. S., Almaiah, M. A., . . . Al-Khasawneh, A. L. (2023). Assessing customers perception of online shopping risks: A structural equation modeling–based multigroup analysis. Journal of Retailing and Consumer Services, 71, 103188. doi:10.1016/j.jretconser.2022.103188

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review*, 13(3), 219-235. doi:10.1177/1088868309341564

Brucks, M. (1985). The Effects of Product Class Knowledge on Information Search Behavior. The *Journal of Consumer Research*, 12(1), 1-16. doi:10.1086/209031

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, 17(1), 61–80. doi:10.1287/isre.1060.0080

Flynn, L. R., & Goldsmith, R. E. (1999). A Short, Reliable Measure of Subjective Knowledge. *Journal of Business Research*, 46(1), 57-66. doi:10.1016/S0148-2963(98)00057-5

Hall, J. A., & Mast, M. S. (2007). Sources of Accuracy in the Empathic Accuracy Paradigm. *Emotion (Washington, D.C.),* 7(2), 438-446. doi:10.1037/1528-3542.7.2.438

Hoehle, H., & Venkatesh, V. (2015). Mobile Application Usability: Conceptualization and Instrument Development. *MIS Quarterly*, 39(2), 435-472. doi:10.25300/MISQ/2015/39.2.08

Kahneman, D. (2003). A Perspective on Judgment and Choice: Mapping Bounded Rationality. *The American Psychologist*, 58(9), 697-720. doi:10.1037/0003-066X.58.9.697

Kahneman, D. & Frederick, S. (2005). A model of heuristic judgment. In Holyoak, K., & Morrison, R. (Eds.) The Cambridge handbook of thinking and reasoning (pp. 267–293). *Cambridge University Press.*

Karwatzki, S., Dytynko, O., Trenz, M. & Veit, D. (2017). Beyond the Personalization-Privacy Paradox: Privacy Valuation, Transparency Features, and Service Personalization. *Journal of Management Information Systems,* 34(2), 369-400. doi:10.1080/07421222.2017.1334467

Karwatzki, S., Trenz, M., & Veit, D. (2022). The multidimensional nature of privacy risks: Conceptualisation, measurement and implications for digital services. *Information Systems Journal (Oxford, England),* 32(6), 1126-1157. doi:10.1111/isj.12386

Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective Discrimination of Stimuli that cannot be Recognized. *Science*, 207(4430), 557-558. doi:10.1126/science.7352271

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion*, 14(4), 473-493. doi:10.1080/026999300402763

Liu, X., & Shi, Z. (2025). Served or Exploited: The Impact of Data Capture Strategy on Users' Intention to Use Selected AI Systems. *Journal of Consumer Behaviour*, 24(2), 1003-1016. doi:10.1002/cb.2428

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15(4), 336–355. doi:10.1287/isre.1040.0032

Mitchell, V.W. (1999). Consumer perceived risk: Conceptualisations and models. *European Journal of Marketing*, 33(1/2),163–195. doi:10.1108/03090569910249229

Pizzi, G., Scarpi, D., & Pantano, E. (2021). Artificial intelligence and the new forms of interaction: Who has the control when interacting with a chatbot? *Journal of Business Research*, 129, 878-890. doi:10.1016/j.jbusres.2020.11.006

Poushneh, A., & Gearhart, R. S. (2025). Semi-Human Artificial Intelligence: The Effect of Search Modality (Text vs. Voice) on Perceived Eeriness and Consumers' Purchase Intention. *Journal of Consumer Behaviour,* 24(4), 1854-1876. doi:10.1002/cb.2496

Reber, R., Wurtz, P., & Zimmermann, T. D. (2004). Exploring "fringe" consciousness: The subjective experience of perceptual fluency and its objective bases. *Consciousness and Cognition*, 13(1), 47–60. doi:10.1016/S1053-8100(03)00049-7

Robin, R., & Dandis, A. O. (2021). Business as usual through contact tracing app: What influences intention to download? *Journal of Marketing Management*, 37(17-18), 1903-1932. doi:10.1080/0267257X.2021.2017323

Schneiderman, B. (2021). Human-Centered AI. *Issues in Science and Technology*, 37(2), 56-61. < https://issues.org/wp-content/uploads/2021/01/56%E2%80%9361-Shneiderman-Human-Centered-AI-Winter-2021.pdf>

Schneider, W., & Shiffrin, R. M. (1977a). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review*, 84, 1–66. doi:10.1037/0033-295X.84.1.1

Schneider, W. & Shiffrin, R. M. (1977b). Controlled and automatic human information processing: 2. Perceptual learning, automatic attending and general theory. *Psychological Review,* 84, 127–190.

Slovic, P. (1987). Perception of risk. Science. 236(4799), 280-285. doi:10.1037/0033-295X.84.2.127

Song, H., & Schwarz, N. (2008). If It's Hard to Read, It's Hard to Do: Processing Fluency Affects Effort Prediction and Motivation. *Psychological Science*, 19(10), 986-988. doi:10.1111/j.1467-9280.2008.02189.x

Song, M., Xing, X., Duan, Y., Cohen, J., & Mou, J. (2022). Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption intention. Journal of retailing and consumer services, 66, 102900. doi:10.1016/j.jretconser.2021.102900

Xu, H., Dinev, T., Smith, J., & Hart, P. (2011). Information privacy concerns: Linking individual perceptions with institutional privacy assurances. *Journal of the Association for Information Systems*, 12(12), 798–824. doi:10.17705/1jais.00281

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9(2), 1-27. doi:10.1037/h0025848

About the authors:

M.Sc. **Mari Paalimäki** is a Doctoral Researcher of Marketing at Jyväskylä University School of Business and Economics (Finland). Her research focuses on privacy decision-making in the age of AI systems, with a particular emphasis on the perspective of consumer behaviour. Alongside her research work, she teaches and supervises theses in the Department of Marketing at Jyväskylä University of School of Business and Economics.

Dr **Juha Munnukka** is a Senior Researcher of Marketing at Jyväskylä University School of Business and Economics (Finland). His research interests include consumer behaviour, digital and AI-powered marketing communication, and influencer marketing. He has led and participated in several research projects and is currently leading a project on AI-powered virtual influencers. Munnukka has published in various academic journals such as the Journal of Interactive Marketing, Industrial Marketing Management, Computers in Human Behavior, Journal of Marketing Management, Journal of Service Theory and Practice, and Journal of Product and Brand Management.

RESEARCH IN PROGRESS

THE DIGITAL PATH TO SUSTAINABILITY: A CONFIGURATIONAL INVESTIGATION

ROHIT NISHANT,[1] MARIE-CLAUDE BOUDREAU[2]

[1] Queen's Business School, Queen's University Belfast, Belfast, United Kingdom of Great Britain and Northern Ireland
r.nishant@qub.ac.uk
[2] University of Georgia's Terry College of Business, Georgia, United States of America
mcboudre@uga.edu

Firms are increasingly leveraging information technology (IT) not only to enhance productivity, but also to improve their environmental performance. In parallel, organizations have begun creating IT-specific C-suite roles and assigning them responsibilities related to sustainability agendas. Against this backdrop, this preliminary study empirically explores whether and how different configurations of IT use for sustainability, IT-specific executive roles, and their environmental responsibilities may jointly shape a firm's environmental sustainability. We conduct a crisp-set Qualitative Comparative Analysis to uncover multiple pathways leading to high environmental sustainability. To deepen and validate our interpretation of these preliminary configurational findings, we plan to complement the analysis with text analysis of firms' reports. Taken together, our preliminary findings suggest that the IT–sustainability relationship is complex and configurational, offering an initial basis for future theory development and practitioner guidance on how digital resources and leadership structures may support environmental goals.

DOI
https://doi.org/10.18690/um.fov.4.2026.69

ISBN
978-961-299-152-4

Keywords:
sustainability,
B Corp,
CIO,
leadership,
QCA

University of Maribor Press

1 Introduction

Environmental sustainability has emerged as a critical priority for firms worldwide, driven by increasing regulatory scrutiny, heightened stakeholder expectations, and the broader societal imperative to mitigate environmental degradation (Kocoglu & Mithani, 2025). Simultaneously, organizations are undergoing substantial digital transformation, investing in information technology (IT) to enhance operational efficiency, improve productivity, and enable data-driven decision-making (Wiener et al., 2025; Vial, 2019). These two developments, environmental sustainability imperatives and digital transformation trajectories, are increasingly intersecting. Firms are not only deploying IT to improve productivity but are also leveraging digital tools and data infrastructures to achieve environmental sustainability (Kotlarsky et al., 2023). In parallel, organizational structures surrounding IT leadership have evolved considerably. Many organizations have instituted IT-specific executive roles, including Chief Information Officers (CIOs), Chief Digital Officers (CDOs), and associated senior positions (Bendig et al., 2023). These leaders are increasingly entrusted with responsibilities related to environmental performance and sustainability governance, while also playing a growing role in digital innovation aligned with environmental objectives (Viets & von Delft, 2025; Zhang & Liu, 2025). Yet, despite these developments, existing literature offers limited insight into whether using IT for sustainability or elevating IT executives to sustainability-related decision-making roles contributes to improved environmental outcomes.

While research on digital sustainability highlights digital technology's role in supporting sustainable operations, such as analytics for environmental monitoring (e.g., Pan et al., 2021), it focuses on technology in isolation and overlooks the broader organizational context. Similarly, research on sustainability governance often neglects the influence of IT executives, despite the role of digital infrastructure in sustainability efforts. Thus, we lack a full understanding of how IT use and IT leadership jointly impact environmental outcomes.

To address this gap, our study examines the interplay between three critical elements: (1) a firm's use of IT for sustainability-related purposes, (2) the presence of IT-specific executive roles, and (3) the environmental responsibilities of IT-specific C-suite executives. We conceptualize environmental sustainability as the joint outcome of IT use for sustainability, IT-specific executive roles, and their

environmental responsibilities. Specifically, we address the following overarching research question: *What configurations of IT-for-sustainability, IT leadership, and IT governance drive high environmental performance?*

We posit that the relationship between IT and environmental sustainability is unlikely to be linear or uniform across firms. For example, investments in sustainability-oriented IT may be insufficient without proper leadership structures to guide, coordinate, and institutionalize their use. Conversely, appointing CIOs or CDOs with sustainability responsibilities may not produce significant results if firms lack the technological resources to implement meaningful environmental practices. Additionally, different combinations of conditions can lead to multiple viable paths for firms to achieve strong environmental performance.

To build the theoretical foundation for our configurational inquiry, we next review three streams of IS research that inform our understanding of how IT use, IT leadership, and IT governance jointly shape environmental sustainability outcomes.

2 Theoretical Background

Our study is at the intersection of three streams of information systems (IS) research: digital sustainability, IT leadership, and IT governance. While each stream offers important insights into how organizations can leverage IT to create value, they have largely developed in parallel, providing fragmented explanations of how IT contributes to environmental performance. Bringing these streams together allows for a more integrated understanding of not only what IT can achieve in the sustainability domain, but also who enables these outcomes and how they are orchestrated within organizations.

The first stream of research is digital sustainability (historically called "green IT/IS"). Prior work in this stream has shown, for instance, that green IT announcements generate positive shareholder reactions (e.g., Nishant et al., 2017) and that carefully designed persuasive digital systems can nudge employees toward pro-environmental behaviors even across work-home boundaries (e.g., Corbett et al., 2025). Yet, despite these important insights into IT's signaling value and behavioral influence, this stream of literature has not disentangled the specific pathways through which IT/IS contributes to environmental sustainability within organizations.

The second stream of research focuses on IT leadership, examining topics such as the critical success factors that enable effective IT leadership (e.g., Kaufeld et al., 2009) and the contribution of IT leadership roles to organizational performance (e.g., Hütter & Reidl, 2017). However, existing research tends to examine leadership effects in isolation, without considering how leadership interacts with broader organizational structures and governance mechanisms to influence sustainability outcomes.

The third area of research is IT governance, which examines how organizations extract value from IT investments while managing risks, how they align technology with strategy, and how they maintain accountability (Wilkin & Chenhall, 2020; Gregory et al., 2018). This area of research is concerned with the mechanisms through which organizations control IT, including oversight processes, decision-making structures, and the allocation of responsibilities for IT-related initiatives.

Taken together, these streams offer complementary but incomplete explanations: digital sustainability research emphasizes what IT can achieve, IT leadership identifies who drives these initiatives, and IT governance explains how they are structured and controlled. We argue that their relationship is best understood through causal complexity, where environmental performance arises from interdependent and intersectional rather than independent effects. Specifically, the impact of digital sustainability initiatives depends on the alignment between IT leadership and governance structures, which jointly enable and constrain their implementation. These relationships are configurational, such that multiple combinations of conditions can produce similar outcomes (equifinality), while misalignment may undermine them. Accordingly, we conceptualize environmental performance as the outcome of interacting organizational conditions, including the use of IT for environmental sustainability, the presence and positioning of C-suite IT leaders, their responsibility for sustainability, the presence of C-suite sustainability leaders, and exposure to ESG risk.

3 Method

Our sample for preliminary analysis consists of 40 certified B Corps drawn from across the globe. We selected B Corps because their certification provides external validation of firms' environmental performance, accountability, and transparency,

thereby reducing the likelihood that the sample would include firms primarily engaged in symbolic sustainability communication or greenwashing. This made B Corps an appropriate empirical setting for investigating organizations with a demonstrated commitment to generating genuine environmental impact.

Consistent with the approach used by Lou et al. (2025), we use the S&P Capital IQ Pro database and its integrated chatbot, chatIQ (late October version), which helps us retrieve information from various published reports through prompts, to capture whether the firm used IT for sustainability-related purposes, the presence of IT-specific executive roles, and the environmental responsibilities of IT-specific C-suite executives. In addition, we considered whether each company operates in sectors with high environmental risk exposure (e.g., energy, mining, etc.). Finally, we also considered, for each company, the environmental dimension of their ESG score, i.e., whether it was higher or lower than that of their peers. To ensure the reliability of results through prompts, we checked the outputs' sources and also regenerated the outputs multiple times.

To ensure data robustness, we triangulated our data with that obtained from Qinsights, an AI qualitative research GDPR-compliant platform, by analyzing annual and sustainability reports published between 2022 and 2025. We also checked the ESG score on the LSEG workspace to verify the consistency of the outcome variable. Regarding operationalization as a categorical variable, no issues were observed.

For our analysis, we use crisp-set Qualitative Comparative Analysis (csQCA) (Rihoux & Ragin, 2009; Nishant & Ravishankar, 2020). We operationalize all variables as categorical indicators, coding each as 0 when a firm does not exhibit the characteristic of interest and 1 when it does. While constructs such as IT use for sustainability and executive responsibilities are inherently nuanced, we adopt crisp-set calibration to focus on the presence or absence of key conditions, consistent with our configurational approach. This is appropriate given the difficulty of deriving reliable continuous measures from heterogeneous, text-based disclosures. We use the fsQCA 4.1 software and set a minimum frequency threshold of 1 case, and raw and PRI consistency thresholds of 0.80, for our preliminary analysis.

4 Preliminary Results

In our preliminary analysis, we identify three distinct causal pathways explaining high environmental performance, with IT-for-sustainability appearing as a core condition across most high-performance configurations. A C-suite IT leader responsible for sustainability emerges as a contributing condition in some configurations but accounts for very low unique coverage or explanation, suggesting limited relevance.

To deepen our interpretation of the configurational findings, we plan to supplement the above analysis with a qualitative text analysis of firms' disclosures, sustainability reports, and public statements, along with targeted interviews with the leadership of firms within each configuration. We expect this additional analysis to enable us to identify the underlying mechanisms that may explain why specific configurations yield superior environmental performance.

5 Potential Contributions and Limitations

This study offers three contributions. First, we advance green IT/IS research by showing that the IT–sustainability relationship is configurational: meaningful environmental outcomes arise from specific combinations of technological, structural, and governance conditions, not from IT adoption alone. Second, we contribute to IT governance and leadership research by demonstrating that CIO and CDO roles improve sustainability only when coupled with appropriate organizational conditions. Third, we offer practitioners actionable guidance by identifying multiple viable pathways to high environmental sustainability, recognizing that no single blueprint fits all firms.

Our study has several limitations. First, our sample of 40 B Corp-certified firms, while appropriate for csQCA, limits the generalizability of our findings to firms with demonstrated sustainability commitments. Second, our reliance on AI-assisted data extraction from publicly available documents may introduce measurement error, as the completeness and consistency of firm disclosures vary. Third, our binary coding of conditions, inherent to csQCA, may obscure meaningful variation within categories. Future research should replicate our configurational approach with larger, more diverse samples and explore finer-grained measures of IT use and leadership responsibilities.

References

Bendig, D., Wagner, R., Piening, E. P., & Foege, J. N. (2023). Attention to digital innovation: exploring the impact of a chief information officer in the top management team. *MIS Quarterly*, 47(4), 1487–1516.

Corbett, J., Nishant, R., Cherki El Idrissi, S., & Chung, A. Q. (2025). Taking Things Out of Context: A Configuration Analysis of the Persuasiveness of Digital Technologies for Sustainability Across Work-Home Boundaries. *Communications of the Association for Information Systems*, 57(1), 33

Gregory, R. W., Kaganer, E., Henfridsson, O., & Ruch, T. J. (2018). IT consumerization and the transformation of IT governance. *MIS Quarterly, 42*(4), 1225–1253.

Hütter, A., & Riedl, R. (2017). *Chief information officer role effectiveness: literature review and implications for research and practice* (pp. 1-30). Springer, Cham.

von Urff Kaufeld, N., Chari, V., & Freeme, D. (2009). Critical Success Factors for Effective IT Leadership. *Electronic Journal of Information Systems Evaluation*, 12(1).

Kocoglu, I., & Mithani, M. A. (2025). Digital sustainability solutions: Expanding the toolkit for environmentally sustainable outcomes. *Academy of Management Perspectives*, (ja), amp-2023.

Kotlarsky, J., Oshri, I., & Sekulic, N. (2023). Digital sustainability in information systems research: Conceptual foundations and future directions. *Journal of the Association for Information Systems*, 24(4), 936–952.

Lou, B., Ma, X., & Wu, L. (2025). Artificial Intelligence, CEO Turnover, and Exploration Orientation in Firm Innovation. *Information Systems Research*. https://doi.org/10.1287/isre.2024.0970

Nishant, R., Teo, T. S., & Goh, M. (2017). Do shareholders value green information technology announcements? *Journal of the Association for Information Systems*, 18(8), 3.

Nishant, R., & Ravishankar, M. N. (2020). QCA and the harnessing of unstructured qualitative data. *Information Systems Journal*, 30(5), 845-865.

Pan, S. L., Li, M., Pee, L. G., & Sandeep, M. S. (2021). Sustainability design principles for a wildlife management analytics system: An action design research. *European Journal of Information Systems*, 30(4), 452–473.

Rihoux, B., & Ragin, C. C. (Eds.). (2009). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vol. 51). Sage.

Wilkin, C. L., & Chenhall, R. H. (2020). Information technology governance: Reflections on the past and future directions. *Journal of Information Systems*, 34(2), 257-292.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems, 28*(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Viets, A., & von Delft, S. (2025). Chief Digital Officer Presence and Corporate Environmental Performance: Assessing Organizational Task Environments Under Which Their Presence Affects GHG Emissions. *Business Strategy and the Environment*, 34(8), 10220–10238.

Wiener, M., Strahringer, S., & Kotlarsky, J. (2025). Where are the processes in IS research on digital transformation? A critical literature review and future research directions. *Journal of Strategic Information Systems, 34*(2), 101900. https://doi.org/10.1016/j.jsis.2025.101900

Zhang, X., & Liu, L. Y. (2025). CIO leadership, employee digital ability, and corporate green innovation performance: Moderating effect of organizational agility and environmental culture. *Environment, Development and Sustainability, 27*, 24585–24628.

RESEARCH IN PROGRESS

FRUGAL DIGITAL TRANSFORMATION FOR RESILIENT FOOD BANKS: A MULTIPLE-CASE STUDY OF DISTRIBUTED NONPROFITS

OUSMANE SY,[1] ELAINE MOSCONI,[1]
MARIE-CLAUDE BOUDREAU[2]

[1] Université de Sherbrooke, Ecole de gestion, Québec, Canada
ousmane.sy@usherbrooke.ca, elaine.mosconi@usherbrooke.ca
[2] University of Georgia's Terry College of Business, Georgia, United States of America
mcboudre@uga.edu

In the context of accelerating technological advancement, digital transformation is becoming a major strategic issue. Yet, it remains understudied in non-profit organizations. Food banks, operating in distributed networks with limited resources, clearly illustrate this gap. Despite their crucial social role, they struggle to leverage the potential of information systems and technology to improve the coordination and efficiency of their operations. This research-in-progress paper draws on the framework of Frugal Information Systems (Frugal IS) guided by the principles of IT (Information Technology) for Good, to examine how simple and inclusive digital solutions can strengthen coordination, improve social value creation, and enhance the resilience of food banks. Adopting a qualitative approach based on a multiple case study within the "Moissons du Québec" network, we aim to formulate recommendations for community organizations seeking to reconcile operational efficiency and social value through digital transformation.

DOI
https://doi.org/10.18690/um.fov.4.2026.70

ISBN
978-961-299-152-4

Keywords:
digital transformation, food banks, distributed organizations, frugal IS, IT for good

University of Maribor Press

1 Introduction

In today's era, marked by an unprecedented acceleration of technological advancements (Omol, 2024), the importance of digital transformation has become almost universally acknowledged. This evolution is transforming the way organizations operate worldwide (Ashrafi et al., 2025; Carroll et al., 2023; Omol, 2024; Wessel et al., 2025). This largely explains why information systems researchers are focusing so intently on the study of digital transformation, approaching the subject from multiple perspectives (Wessel et al., 2025).

A perspective that has been neglected is one that foregrounds human realities (Ashrafi et al., 2025; Carroll et al., 2023), prioritizes societal over purely economic outcomes (Wessel et al., 2025), and accommodates diverse organizational contexts (Grover & Niederman, 2021). From this perspective, it is essential to examine how a human-centered, societally oriented, and context-sensitive approach to digital transformation is enacted within specific organizational structures.

Indeed, in 2023, 2.33 billion people faced food insecurity, including 733 million suffering from chronic hunger (FAO, 2024). Food banks play a key role by redistributing surplus food to vulnerable populations, while also reducing food waste (Martins et al., 2019), which contributes to 8–10% of global greenhouse gas emissions (UNEP, 2021).

However, food banks face several operational and logistical challenges (Rivera et al., 2023; Wu & Tai, 2024). Difficulties exist in information sharing between different food banks, which are generally charitable organizations, a gap in digital literacy between volunteers and partner companies, and significant food losses (Wu & Tai, 2024). Furthermore, inefficient logistics management and increasing public demand exacerbate these problems (Martin et al., 2022). While digital transformation is widely studied in IS research (Ashrafi et al., 2025; Wessel et al., 2025), most work focuses on for-profit organizations, leaving nonprofits largely overlooked (Cipriano & Za, 2025). This gap matters because food banks address food insecurity and reduce waste, key goals for sustainability and resilience (FAO, 2024; UNEP, 2021), yet operate under severe resource constraints and distributed structures that make conventional approaches impractical (Grover & Niederman, 2021). Current literature offers little guidance on how frugal, inclusive, and socially oriented IS

solutions can improve coordination, efficiency, and adaptive capacity in such settings (Radjou, 2024; Watson et al., 2013). Closing this gap is essential for designing digital practices that align operational performance with social value and sustainability objectives.

Our research focuses on digital transformation in the internal operations of food banks. We aim to answer the following research question:

How can digital transformation guided by Frugal IS and IT for Good principles enhance coordination, social value and resilience in distributed nonprofit organizations?

To answer this question, we use a qualitative multiple-case study (Eisenhardt, 1989) of four food banks in the "Moissons du Québec" network. Data collection combines 12 interviews across four food banks, observations, and document analysis for triangulation. Thematic analysis (Braun & Clarke, 2006) guides interpretation, consistent with an inductive approach. Our research contributes to IS research in three ways. Theoretically, it bridges Frugal IS, IT for Good, and stakeholder theory to explore digital transformation in distributed nonprofit organizations, a context often overlooked in the literature. Empirically, it offers insights from a multiple-case study of food banks, revealing how frugal digital practices shape coordination, social value, and resilience under resource constraints. Practically, it proposes design-oriented recommendations to help community organizations implement inclusive, low-cost technologies that support sustainable and socially valuable operations.

2 Theoretical Background

Frugal Information System: The concept of a Frugal Information System (Frugal IS) emerged in the information systems literature following the pioneering work of Watson et al. (2013). These authors define a frugal IS as "an information system that is developed and deployed with minimal resources to meet the pre-eminent goal of the client" (p. 2). This definition rests on two fundamental ideas: minimization of financial, human, and technical resources, and focus on a single priority objective. Performance is measured by the ability to meet essential needs in contexts of scarcity. Ahuja & Chan (2016) extend this perspective through Frugal IT Innovation

Capability (FITIC), emphasizing co-creation and collaborative ecosystems. Radjou (2024) highlights simplicity, efficiency, and creativity under constraints.

Proposition 1: Frugal IS supports simple and adaptable digital solutions that improve coordination and efficiency in food banks under resource constraints.

IT for Good: Although widely used in research, the concept of IT for Good remains relatively unformalized in scientific literature, with few benchmark studies. It is based on a vision of technology as a lever for creating social value, inclusion, cohesion, and sustainability, directing information systems toward the common good rather than economic performance alone. Barradas et al. (2019) illustrate this by showing how an information system for social tourism generates social value by improving access for marginalized older people. Similarly, Wessel et al. (2025) call for rethinking digital transformation to unlock new value propositions. Thus, digital technologies should not only enhance performance but also enable inclusive and sustainable social value, in line with IT for Good.

Proposition 2: IT for Good aligns digital transformation with social value, strengthening cohesion and coordination within organizations.

Digital Transformation in Distributed Nonprofits: Digital transformation is generally defined as a profound and continuous process of integrating digital technologies into organizational activities, structures, and models, transforming practices, relationships, and objectives (Ashrafi et al., 2025; Carroll et al., 2023; Wessel et al., 2025). The literature highlights issues such as distributed innovation (Ashrafi et al., 2025), the consideration of human dimensions (Carroll et al., 2023), and societal impacts (Wessel et al., 2025), but this research remains largely rooted in resource-rich and performance-oriented contexts, leaving the constrained environments of non-profit organizations largely unexplored (Cipriano & Za, 2025; Yang et al., 2025). Within this framework, food banks represent a particularly relevant field of study. As distributed organizations, they rely on the coordination of autonomous actors through digital infrastructures that facilitate interactions (Vergne, 2020). However, their digital transformation is not driven by massive technological investments (Martin et al., 2022; Ng et al., 2025), but by the need to coordinate dispersed operations in a context of scarcity (Rivera et al., 2023). This leads to a preference for accessible, adaptable, and inclusive digital solutions, thus revealing logics where coordination, frugality, and social value take precedence.

Proposition 3: In distributed food banks, digital transformation guided by frugal IS and It for good enhances social value creation and strengthens organizational resilience.

3 Methods

In accordance with the framework of (Edmondson & Mcmanus, 2007), our study falls within an intermediate knowledge domain, making a qualitative approach particularly well-suited for producing rich and contextualized theoretical contributions. This choice is justified by the exploratory nature of the research question. We adopt a multiple case study strategy, as proposed by Eisenhardt (1989), to capture the diversity of practices, organizational contexts, and technological uses within the network of 18 food banks in Quebec, Canada. Four cases were purposefully selected to ensure variation in size, geographical location, and level of technological adoption.

Data collection: Three data sources were used. First, 12 semi-structured interviews were conducted with directors, logistics managers, coordinators, and volunteers across four food banks. Second, observations were carried out during in-person coordination meetings and logistics activities. Third, internal documents and external reports from the "Banques Alimentaires du Québec" and partners supported triangulation. Data collection continued until data adequacy was reached, ensuring sufficient depth and variation to support robust theoretical insights (Abdalla Mikhaeil & Robey, 2024).

Data analysis : We conducted a thematic analysis (Braun & Clarke, 2006) aligned with an interpretative and inductive stance. This approach is particularly well-suited to the in-depth exploration of the meanings that actors attribute to their practices, as well as to the progressive construction of analytical categories based on fieldwork.

4 Preliminary Findings

Our preliminary findings highlight several dynamics related to coordination, technology use, and organizational constraints in food banks.

First, our data shows that coordination relies heavily on informal mechanisms. Participants describe exchanges based on trust, interpersonal relationships, and tacit knowledge, rather than structured formal systems:

"*…he knows from experience, mostly from habit… everyone knows how things work… we have enough experience in transportation here*" (B3 INT3)

"*…drivers with 15, 20, 25 years of experience in transportation… the telephone is sufficient*" (B3 INT2)

"*If we had it in stock, but it was more in our heads*" (B1 INT1).

Second, our data reveals that the technological solutions adopted are mostly simple, adapted, and developed according to local needs. Participants frequently mention the use of unsophisticated or "homegrown" tools, compatible with their limited resources: "*…tailor-made…adapted to the needs of reality*" (B1 INT3)

Furthermore, the uses of digital technology are primarily geared towards supporting the social mission. Technologies are mobilized to improve food redistribution and reduce food waste, rather than to optimize economic performance.

"*Digital technology contributes to strengthening the social mission*" (B1 INT1)

"*Feed as many people as possible… A platform where all surpluses could be stored"* (B2 INT2).

Our data also show that organizational practices rely on hybrid systems combining digital tools and tacit knowledge, particularly in inventory management, as illustrated by these quotes:

"Because I was in the warehouse, I knew if I had enough for the organization. There was no inventory system" (B1 INT1)

"We don't have an inventory control system because we don't need one... we use a push system, we push the food... The organization doesn't choose what it wants." (B3 INT3)

Finally, food banks appear as distributed structures, characterized by coordination between autonomous entities. However, digital transformation remains constrained by limitations in resources and skills, leading to gradual and pragmatic adoption processes.

5 Discussion

This ongoing research allow us to move beyond a purely technological interpretation of digital transformation. First, our data shows that coordination does not rely exclusively on digital tools, but on pre-existing social mechanisms. Technologies thus appear as support mechanisms for already established relational dynamics. Second, frugality emerges as an organizational logic that structures technological choices and promotes the appropriation of tools, in line with local capacities (Proposition 1). Third, unlike for-profit organizations, the use of digital technology here is driven by social mission. This confirms that digital transformation can be guided by logics of social value rather than by objectives of economic efficiency (Proposition 2). Furthermore, the observed hybrid systems suggest that digital transformation does not replace existing practices, but rather integrates with them, combining technologies and tacit knowledge. Finally, these results highlight a tension between network coordination and local autonomy. Digital transformation in these organizations appears as a situated, constrained, and progressively constructed process, rather than a linear and standardized change. These results support distributed food banks, digital transformation guided by frugal IS and IT for good help to strengthen social value creation and organizational resilience (Proposition 3).

6 Conclusion

This research highlights the role of frugal digital transformation in food banks as distributed organizations. The findings show that simple, tailored, and inclusive solutions support coordination, enhance operational efficiency, and facilitate the creation of social value. Beyond technology, digital transformation is embedded in social and organizational dynamics, where governance, volunteers, and resource constraints play a central role. By combining frugality, inclusion, and social mission, digital practices contribute to the resilience of the food system. Thus, this study

underscores the importance of designing technologies aligned with the realities and needs of non-profit organizations.

References

Abdalla Mikhaeil, C., & Robey, D. (2024). When is enough enough? A critical assessment of data adequacy in IS qualitative research. Information and Organization, 34(4), 100540. https://doi.org/10.1016/j.infoandorg.2024.100540

Ahuja, S., & Chan, Y. E. (2016). Digital Innovation : A Frugal Ecosystem Perspective. Communications of the Association for Information Systems, 39, 35-48.

Ashrafi, A., Constantinides, P., Mehandjiev, N., & Thatcher, J. B. (2025). Mobilising new frontiers in digital transformation research : A problematization review. Information Systems Journal, 35(1), 97-139. https://doi.org/10.1111/isj.12531

Barradas, L. C., Nascimento, J., Martins, L., Oliveira, S., Leal, S., Vivas, C., & João, R. S. (2019). Developing an Information System for Social Value Creation. Procedia Computer Science, 164, 546-551. https://doi.org/10.1016/j.procs.2019.12.218

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Carroll, N., Conboy, K., University of Galway, Hassan, N. R., University of Minnesota, Hess, T., Ludwig Maximilian University of Munich, Junglas, I., College of Charleston, Morgan, L., & University of Galway. (2023). Problematizing Assumptions on Digital Transformation Research in the Information Systems Field. Communications of the Association for Information Systems, 53, 508-531. https://doi.org/10.17705/1CAIS.05322

Cipriano, M., & Za, S. (2025). Digital Transformation of Non-profit Organizations : A Research Agenda. Communications of the Association for Information Systems, 56. https://doi.org/doi.org/10.17705/%25201CAIS.05643

Edmondson, A. C., & Mcmanus, S. E. (2007). Methodological fit in management field research. Academy of Management Review, 32(4), 1246-1264. https://doi.org/10.5465/amr.2007.26586086

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. The Academy of Management Review, 14(4), 532-550. https://doi.org/10.2307/258557

FAO. (2024). The State of Food Security and Nutrition in the World 2024. FAO; IFAD; UNICEF; WFP; WHO; https://doi.org/10.4060/cd1254en

Grover, V., & Niederman, F. (2021). Research Perspectives : The Quest for Innovation in Information Systems Research: Recognizing, Stimulating, and Promoting Nkovel and Useful Knowledge. Journal of the Association for Information Systems, 22(6), 1753-1782. https://doi.org/10.17705/1jais.00705

Martin, N. M., Barnett, D. J., Poirier, L., Sundermeir, S. M., Reznar, M. M., & Gittelsohn, J. (2022). Moving Food Assistance into the Digital Age : A Scoping Review. International Journal of Environmental Research and Public Health, 19(3), 1328. https://doi.org/10.3390/ijerph19031328

Ng, Y., Fraser, K. T., Stamou, A., Rosenthal, A., & Cohen, N. (2025). Using Digital Technology to Facilitate Choice for Food Pantry Customers : An Evaluation of a Pilot Program. Journal of Nutrition Education and Behavior, S1499404625000521. https://doi.org/10.1016/j.jneb.2025.02.011

Omol, E. J. (2024). Organizational digital transformation : From evolution to future trends. Digital Transformation and Society, 3(3), 240-256. https://doi.org/10.1108/DTS-08-2023-0061

Radjou, N. (2024). The Frugal Economy : A Guide to Building a Better World With Less (1re éd.). Wiley. https://doi.org/10.1002/9781394308644

Rivera, A. F., Smith, N. R., & Ruiz, A. (2023). A systematic literature review of food banks' supply chain operations with a focus on optimization models. Journal of Humanitarian Logistics and Supply Chain Management, 13(1), 10-25. https://doi.org/10.1108/JHLSCM-09-2021-0087

Schallmo, D., Williams, C. A., & Boardman, L. (2017). DIGITAL TRANSFORMATION OF BUSINESS MODELS — BEST PRACTICE, ENABLERS, AND ROADMAP. International Journal of Innovation Management, 21(08), 1740014. https://doi.org/10.1142/S136391961740014X

UNEP. (2021, avril 3). UNEP Food Waste Index Report 2021 | UNEP - UN Environment Programme. https://www.unep.org/resources/report/unep-food-waste-index-report-2021

Vergne, J. (2020). Decentralized vs. Distributed Organization : Blockchain, Machine Learning and the Future of the Digital Platform. Organization Theory, 1(4), 2631787720977052. https://doi.org/10.1177/2631787720977052

Vial, G. (2019). Understanding digital transformation : A review and a research agenda. The Journal of Strategic Information Systems, SI: Review issue, 28(2), 118-144. https://doi.org/10.1016/j.jsis.2019.01.003

Watson, R. T., Kunene, K. N., & Islam, M. S. (2013). Frugal information systems (IS). Information Technology for Development, 19(2), 176-187. https://doi.org/10.1080/02681102.2012.714349

Wessel, L., Mosconi, E., Indulska, M., & Baiyere, A. (2025). Digital Transformation : Quo Vadit? Information Systems Journal, isj.12578. https://doi.org/10.1111/isj.12578

Wu, P.-J., & Tai, Y.-C. (2024). Artificial intelligence-based food-quality and warehousing management for food banks' inbound logistics. Journal of Enterprise Information Management, 37(1), 307-325. https://doi.org/10.1108/JEIM-10-2022-0398

Yang, Y., An, R., Fang, C., & Ferris, D. (2025). Artificial Intelligence in Food Bank and Pantry Services : A Systematic Review. Nutrients, 17(9), 1461. https://doi.org/10.3390/nu17091461

Author Revision Statement

We hereby confirm that the manuscript entitled:

"Frugal digital transformation for resilient food banks: a multiple-case study of distributed nonprofits"

submitted to the 30th Bled eConference has been carefully revised following the comments and recommendations provided by the reviewers.

We confirm that no substantial content has been altered beyond the scope of the reviewers' recommendations without proper justification.

RESEARCH IN PROGRESS

Evaluating an AI Act Regulatory Sandbox Framework Against Pilot Cases: Towards a Methodologically Grounded Three-Round Process Framework

Jukka Heikkilä,[1] Altti Lagstedt,[2] Seppo Heikura,[1] Anastasiia Dardykina[3]

[1] University of Turku, Turku School of Economics, Turku, Finland
jups@utu.fi, seppo.heikura@utu.fi
[2] Haaga-Helia University of Applied Science, Helsinki, Finland
altti.Lagstedt@haaga-helia.fi
[3] University of Vaasa, School of Accounting and Law, Vaasa, Finland
anastasiia.dardykina@uwasa.fi

This research-in-progress paper advances out tentative AI Regulatory Sandbox Framework (AIRSF) in two directions. Empirically, it evaluates 46 pilot cases and with five workshops of Nordic national/regional AI ecosystems. Conceptually, it argues that the AI Act's dual objective of innovation and compliance can only be operationalised in a business-friendly manner if grounded in participatory, situation-adaptive method development with dispute resolution procedures on controversial issues. Earlier we have identified four methodological steps: stakeholder mapping, conflict surfacing, iterative KPI-building, institutionalised conflict resolution, and show their concrete design consequences for a three-round sandbox related sub-processes: (1) a revised use case and risk category definition procedures with company-paced entry; (2) multi-authority evaluation with agile guidance and co-ordination; (3) post-market surveillance options with iterative re-entry. Preliminary findings surface three structural challenges: entry friction for determining use case purpose, borderline high-risk cases, capabilities altering the AI System design and purpose while in sandbox, and a post-market visibility gap in Round 3.

DOI
https://doi.org/10.18690/um.fov.4.2026.71

ISBN
978-961-299-152-4

Keywords:
AI act, regulatory sandbox, stage-gate model, participatory method engineering, triple helix, post-market surveillance, SMEs

University of Maribor Press

1 Introduction and method

The EU AI Act (AIA) mandates at least one AI Regulatory Sandbox (AIRSF) in each Member State by August 2026, with a dual objective encoded in Art. 57(9): foster innovation and ensure regulatory compliance (EC, 2025a). For AI providers and deployers, particularly SMEs, the sandbox is potentially the difference between market access and exclusion. The AIRSF situates itself within the IS research stream on regulatory governance infrastructures (Tilson et al., 2010) addressing the identified gap that existing IT governance frameworks address internal organisational alignment but rarely incorporate dynamic, multi-party conflict resolution (Sambamurthy & Zmud, 1999). Art. 57(9)'s objectives (b) and (d) structurally serve regulator interests while (a) and (c) serve company interests (Dardykina et al., 2026), an explicit conflict-resolution method is a design necessity.

Building on the AIRSF presented at Bled 2025 (Heikkilä et al., 2025), a seven-gate Stage-Gate model as shared boundary object for sandbox co-design, this paper evaluates the tentative framework against 46 pilot cases and five workshops. The AIA's dual objective is only achievable if sandbox design is grounded in participatory, situation-adaptive method engineering. Without explicit conflict resolution, implementations default to regulator-centric or company-centric designs, neither of which realises the Triple Helix ideal (Etzkowitz & Leydesdorff, 2000) the AIA requires. We address three research questions: RQ1 How do the AIRSF's design requirements hold up against pilot case evidence? RQ2 What methodological grounding is necessary to resolve Art. 57(9)'s conflicting interests? RQ3 What are the eligibility/entry criteria for AI Systems in the pre-entry phase and for substantially changed systems post-market? The pilot case set comprises 46 AI providers, deployers, and National Competent Authorities (NCAs) across the EU, selected through selective sampling prioritising geographic spread, actor-role diversity, and system-type coverage, supplemented by 4 regional workshops combining Design Science Research with Action Research (Iivari & Venable, 2009).

2 From Triple Helix to Revising Tentative model, preliminary findings

Dardykina et al. (2026) identify four methodological steps for sandbox governance: (1) map all affected parties and interests; (2) surface and negotiate conflicts via structured scenario methods (Godet, 2000); (3) build solutions iteratively with

transparent KPIs; and (4) establish institutionalised conflict-resolution mechanisms. These steps correspond to three theoretical groundings. Participatory method engineering (Checkland, 2000) provides the conflict-surfacing tools for Steps 2 and 4. Design Science Research (Van Aken & Romme, 2009; Baskerville et al., 2018; Hevner et al., 2004) supplies the iterative evaluation logic for Steps 1 and 3, enabling agile refinement of compliance criteria without full process restart. The Triple Helix model (Etzkowitz & Leydesdorff, 2000) provides the institutional coordination frame, repositioning the sandbox from a one-directional checkpoint into a middle-out learning space where business, authority, and innovation ecosystem actors co-produce compliance. Without this integration, a purely process-oriented sandbox remains a sequential inspection station rather than a generative institution for trustworthy AI governance.

Our suggested approach serves a triple helix activity among the businesses, value adding parties of the innovation ecosystem and the authorities during and after the sandbox. We use the following simplified figure (Figure 1) of the tentative process (Heikkilä et al., 2025) to condense the stage-gate model to the minimum (c.f. Heikkilä, 2025; Heikkilä & Lagstedt, 2025).

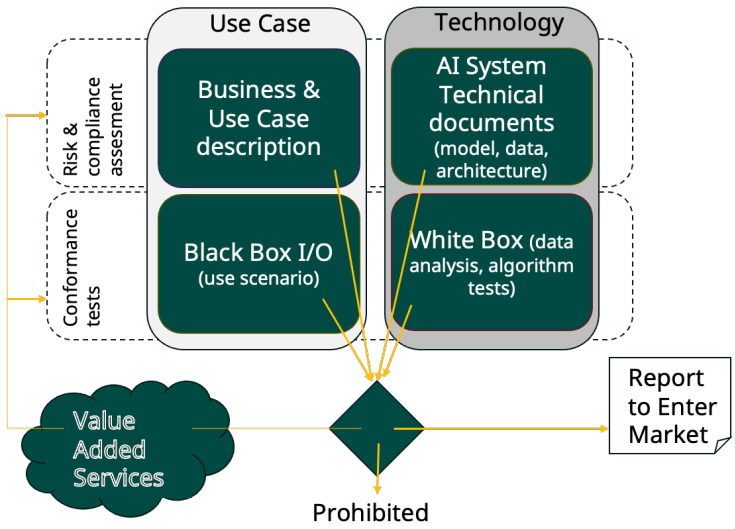

Figure 1: The Simplified Sandbox & Value Added Services Representation (Use Case × Technology evaluation matrix)

Source: Authors

3 Round 1: Use Case Definition and Company-Paced Entry

Building on Heikkilä et al. (2025), sandbox participation begins with AI provider self-evaluation (EC, 2025b) covering use case documentation, initial AIA compliance, and risk categorisation. Pilot cases reveal high demand for guidance on borderline-risk systems (medical devices, biometric tools, digital services), where a minor design and use-case change can shift the risk category, particularly for SMEs facing overlapping regulations (GDPR, Medical Device regulation). System maturity must exceed Technology Readiness Level 3 (TRL 3) (Martínez-Plumed et al., 2021), with a clearly defined use case, training and input data process, documented compliance attempts, and a credible go-to-market strategy, agentic AI's dynamism demands additional scrutiny (Yousefi et al., 2025). Company and business validity, adequate funding, viable business model, and value chain, are also assessed, while cross-domain AI systems (e.g. object identification across healthcare and autonomous driving) require deep use-case and technical analysis (Fig. 1) supported by innovation ecosystem services.

Pilot analysis reveals a more interactive guidance process than anticipated: company representatives are supported by legal, technical, and business experts to map systems along the AIRSF's two-dimensional use case typology (number of models × domain breadth) and produce a preliminary risk estimate (Use Case column, Fig. 1). This has prompted a key AIRSF revision: Gate 1 has been split into a lightweight 'pre-entry' sub-process with company-paced documentation support, distinct from the formal NCA eligibility decision, reducing entry friction without compromising evaluation integrity.

4 Round 2: Multi-Authority Evaluation of Compliance with/without Conformance testing

Pilot evidence shows the number of involved authorities can grow to tens (e.g. GDPR, IPR, Cyber Security, Product safety and its Quality Systems), raising an open question of whether product-safety testing with AI components should be embedded in the sandbox or treated as an auxiliary activity, with co-ordination potentially residing with innovation ecosystem parties providing services to start-ups and SMEs. Three design consequences follow. First, entry must include a rapid self-assessment followed by a formal NCA eligibility decision with transparent criteria,

in days rather than weeks. Second, testing must accommodate agile, in-turn commentary and development, since a sandbox functioning as a final-stage inspection station is incompatible with the fail-fast model of most start-ups (Dardykina et al., 2026). This may require the use of some external experts at appropriate stages of the process. Third, re-entry after modification must not require full process restart, only the changes requested by the supervisory body are re-verified. These findings have prompted revision of the AIRSF Round 2 gate structure: a rolling-wave evaluation protocol now replaces the original single compliance checkpoint, allowing in-turn authority commentary and partial re-evaluation of changed components without restarting the full multi-authority review.

5 Post-Market Surveillance and Re-Entry to the sandbox.

The third sub-process extends the AIRSF beyond the exit report. Three further sub-processes are required: (a) continuous monitoring of the deployed system against the exit-report compliance baseline, with an AI Educational Center (Dardykina et al., 2026) as the institutional repository for accumulated learning; (b) detection and flagging of changes crossing the AIA Art. 83 'substantial modification' threshold, which are in a clear case a change of the purpose of the AI System, change of data, model or both (c) conditional re-entry at the gate appropriate to the modification, depending on the exit state for revision, or value-added services expertise and guidance. The open questions concern who operationalises the trigger criteria and who bears observation responsibility – questions that methodological step 4 demands be explicitly assigned, since provider incentive structures do not reliably generate timely self-reporting of compliance-relevant changes. This gap has prompted revision of the AIRSF's Round 3 design: explicit accountability assignments for post-market observation are now embedded as a negotiated output of Round 2, with the AI Educational Center (Dardykina et al., 2026) designated as the institutional repository for monitoring and re-entry triggering.

Finally, NCAs increasingly rely on RegTech tools to guide, document, and analyse AI systems, tools that companies already use. As the number of market solutions grows, Market Surveillance Authorities (MSAs) will need adapted versions of these same tools for effective surveillance.

6 Discussion and Open Questions

Table 1 shows the round-specific open questions and their methodological stakes. Two key tensions frame the discussion. First, a specification–development mismatch: the AIA expects systems to be well-defined at sandbox entry, whereas the target actors' iterative practices assume the opposite. This creates a clash between regulatory epistemology (stable systems enable risk classification) and innovation epistemology (specification emerges through iteration). The Round 2 agile-compatibility measures respond to this but implementing them requires the participatory method's ability to surface such conflicts, not process design alone.

Secondly, the dual-objective tension cannot be solved merely by allocating functions across the three rounds unless institutional roles in each are explicitly negotiated. The paper's methodological contribution is to show that the AIA's Triple Helix ambition requires a participatory method as its operational base; without it, the sandbox fails both companies and regulators, an outcome by 46 pilot cases. The full paper will present the thematic analysis, a revised AIRSF diagram with the three-round structure and agile re-entry pathways, and design propositions refined through stakeholder co-design.

Table 1: Open design and governance questions by round

Round	Open Question1	Methodological Stake
R1	Separate deployer entry lane or provider-led process? (AIA Art. 25 accountability)	Stakeholder mapping; all affected parties identified
R1	What business significance criteria determine priority access for borderline high-risk cases? Can 3–4 day NCA assessment be operationalised?	Company-paced entry; SME problems tackled in advance
R2	Is notified-body conformance testing inside or outside the sandbox boundary? (cf. Fig. 1 conformance track)	Institutional scope; avoiding redundant evaluation silos
R2	How to accommodate agile/fail-fast cycles within the gate structure without undermining evaluation integrity?	Rolling-wave planning; sandbox as early engagement, not final checkpoint
R3	What operationalises AIA Art. 83 'substantial modification' as an observable re-entry trigger?	Legal predictability of change of purpose; re-entry without full process restart in case of tech/data change.
R3	Who observes and flags changed AI behaviour post-deployment — provider, deployer, NCA, or MSA?	Accountability assignment; institutionalised conflict resolution.

References

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design science research contributions: Finding a balance between artifact and theory. In *Journal of the Association for Information Systems* (Vol. 19, Number 5, pp. 358–376). Association for Information Systems. https://doi.org/10.17705/1jais.00495

Checkland, P. (2000). Soft Systems Methodology: A Thirty Year Retrospective a. *Systems Research and Behavioral Science Syst. Res*, *17*, 11–58. https://doi.org/10.1002/1099-1743(200011)17:1

Dardykina, A., Heikkilä, J., Lagstedt, A., Heikura, S., & Jokinen, J. (2026). *Who is Really Playing in the Sandbox?* https://cybersecnatlab.it/wp-

EC. (2025a). *AI act*. Https://Digital-Strategy.Ec.Europa.Eu/En/Policies/Regulatory-Framework-Ai.

EC. (2025b). *AI Act Single Information Platform*. Https://Ai-Act-Service-Desk.Ec.Europa.Eu/En.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university-industry-government relations. In *Research Policy* (Vol. 29). www.elsevier.nlrlocatereconbase

Godet, M. (2000). *The Art of Scenarios and Strategic Planning: Tools and Pitfalls*. http://www.cnam.fr/lips/,

Heikkilä, J. (2025). *A presentation at ASCEND Confrence under the Danish Presidency of the Council of the EU: Strategy to Action: Closing the European Innovation and Technology Gap*.

Heikkilä, J., & Lagstedt, A. (2025). *A tutorial at ECAI*. https://sites.google.com/view/sandbox-ecai2025-tutorial

Heikkilä, J., Lagstedt, A., Heikura, S., Kaakkurivaara, E., Dardykina, A., & de Chardin, A. T. (2025). *From Simple Sandbox Process to Regulatory Sandbox Framework: Serving the Dual Objectives of AI Regulation*. 643–660. https://doi.org/10.18690/um.fov.4.2026.40

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). DESIGN SCIENCE IN INFORMATION SYSTEMS RESEARCH 1. In *Design Science in IS Research MIS Quarterly* (Vol. 28, Number 1).

Iivari, J., & Venable, J. R. (2009). *Association for Information Systems AIS Electronic Library (AISeL) Action research and design science research-Seemingly similar but decisively dissimilar Recommended Citation*. http://aisel.aisnet.org/ecis2009/73

Martínez-Plumed, F., Gómez, E., & Hernández-Orallo, J. (2021). Futures of artificial intelligence through technology readiness levels. *Telematics and Informatics*, *58*. https://doi.org/10.1016/j.tele.2020.101525

Sambamurthy, V., & Zmud, R. W. (1999). *ARRANGEMENTS FOR INFORMATION TECHNOLOGY GOVERNANCE: A THEORY OF MULTIPLE CONTINGENCIES^*.

Tilson, D., Lyytinen, K., & Sørensen, C. (2010). Digital infrastructures: The missing IS research agenda. *Information Systems Research*, *21*(4), 748–759. https://doi.org/10.1287/isre.1100.0318

Van Aken, J. E., & Romme, G. (2009). Reinventing the future: Adding design science to the repertoire of organization and management studies. *Organisation Management Journal*, *6*(1), 5–12. https://doi.org/10.1057/omj.2009.1

Yousefi, Y., Billi, M., & Rotolo, A. (2025). *Agentic AI: An EU AI Act Paradigm Shift?* https://doi.org/https://dx.doi.org/10.2139/ssrn.5731424

DOCTORAL CONSORTIUM

STRATEGIC ADOPTION OF AI-ENABLED DECISION-MAKING SYSTEMS: DESIGNING FOR HUMAN AGENCY IN DATA-DRIVEN ORGANIZATIONS

BARRAK ALBABTAIN

University College Dublin, UCD Michael Smurfit Graduate Business School, Dublin, Ireland
barrak.albabtain@ucdconnect.ie

DOI
https://doi.org/10.18690/um.fov.4.2026.72

ISBN
978-961-299-152-4

Artificial intelligence is increasingly embedded in organizational decision-making, reshaping how managers exercise discretion and responsibility. This doctoral research examines how AI-enabled decision systems affect human agency in data-driven organizations. It focuses on how technological design features, including transparency and override flexibility, interact with governance structures such as accountability and incentive systems. Using a sequential multi-phase design combining experiments and qualitative fieldwork, the study investigates both perceived and enacted managerial agency. The intended contribution is an Information Systems framework explaining when AI supports human augmentation and when it produces functional substitution. The research aims to inform the design of responsible sociotechnical systems that preserve meaningful human involvement while retaining the efficiency benefits of AI-enabled decision support.

Keywords:
artificial intelligence;
human-AI interaction;
decision support systems;
algorithmic governance;
managerial discretion;
information systems design

University of Maribor Press

1 Introduction

Artificial intelligence is increasingly embedded in organizational decision systems across hiring, performance evaluation, financial risk assessment, forecasting, and resource allocation (Faraj et al., 2018; Kellogg et al., 2020). Organizations adopt AI-enabled systems to improve efficiency, scalability, and predictive accuracy. However, these systems do more than support decisions. They shape how decisions are structured, justified, and governed (Shrestha et al., 2019; Raisch & Krakowski, 2021).

Information Systems research has examined algorithmic transparency, explainability, trust calibration, bias mitigation, and human-AI collaboration (Glikson & Woolley, 2020; Burton et al., 2020; Jarrahi, 2018). While these perspectives are valuable, much existing work treats AI mainly as an input into decision-making rather than as a structural component of decision systems. When AI systems generate recommendations, rankings, or classifications, they introduce new forms of epistemic authority that may alter how managerial discretion is exercised (Berente et al., 2021).

Recent AI governance developments intensify this issue. The European Union Artificial Intelligence Act establishes risk-based requirements for transparency, accountability, and human oversight in high-impact AI systems (European Union, 2024; Cancela-Outeda, 2024). Organizational responsible AI frameworks similarly emphasize documentation, auditability, and structured oversight (Papagiannidis et al., 2025). These governance arrangements do not simply constrain AI use. They shape how decision systems are designed and how authority is distributed between humans and algorithms.

This doctoral research asks how AI-enabled decision systems reshape human agency in organizational contexts. Human agency is understood as the capacity to exercise meaningful discretion, interpret information, and assume responsibility for outcomes. The project examines when AI functions as augmentation, preserving interpretive authority, and when it becomes functional substitution, where human actors formally approve algorithmic outputs while exercising limited substantive discretion.

2 Problem Definition

Much of the current literature focuses on AI adoption, trust, fairness, or performance outcomes. Less attention has been given to how AI integration reshapes managerial discretion itself. This gap matters because managers may remain formally accountable for decisions even when algorithmic outputs strongly shape what choices appear legitimate, rational, or defensible.

Human-in-the-loop arrangements are often treated as sufficient to preserve agency. However, the mere presence of a human actor in a workflow does not guarantee meaningful discretion. Override mechanisms may exist but be procedurally burdensome. Performance metrics may reward adherence to algorithmic outputs. Accountability structures may diffuse responsibility and discourage deviation. In these situations, substitution can emerge indirectly through system design and governance architecture rather than explicit automation.

This research therefore conceptualizes AI integration along an augmentation-substitution continuum (Raisch & Krakowski, 2021). Augmentation refers to configurations in which AI expands informational capacity while preserving meaningful human interpretation and override authority. Substitution refers to configurations in which AI effectively determines outcomes, with humans serving primarily as validators of system outputs.

The central research question is: How do AI-enabled decision systems reshape human agency in organizational contexts, and under what conditions do they function as augmentation rather than substitution?

The study is guided by three sub-questions. First, how do design features such as algorithmic transparency and override flexibility influence perceived and enacted managerial agency? Second, how do accountability structures and incentive systems shape the legitimacy and frequency of deviation from AI recommendations? Third, how do managers interpret and enact discretion in AI-supported strategic decisions?

3 Methodology

This doctoral research adopts a sequential multi-phase design. The phased approach allows the study to move from controlled mechanism testing to contextualized organizational understanding and finally to integrative theory development. The design consists of three phases: an experiment on perceived agency, qualitative fieldwork on enacted agency, and integrative model development.

Phase 1 uses experimental methods to identify causal mechanisms linking AI system design features to perceived managerial agency. Participants will assume managerial roles in simulated decision environments and evaluate cases in which AI systems provide recommendations. The experiment will manipulate algorithmic transparency, strength of recommendation, override flexibility, and accountability framing. Dependent variables will include perceived control, responsibility attribution, confidence in decision quality, and likelihood of override.

The purpose of Phase 1 is to strengthen internal validity by isolating how specific design features influence agency-related outcomes. For example, transparency may increase perceived control by making algorithmic outputs more interpretable, while high-friction override procedures may reduce willingness to deviate even when participants disagree with the recommendation.

Phase 2 uses qualitative case studies to examine enacted agency in organizations implementing AI-enabled decision systems. Data collection will include semi-structured interviews with managers, analysts, and governance officers. Where possible, it will also include policy documents, workflow diagrams, audit procedures, and decision logs. The qualitative phase will examine how managers interpret algorithmic outputs, when they override or follow recommendations, and how organizational norms shape the legitimacy of deviation.

Phase 3 integrates findings from the experimental and qualitative phases into a conceptual framework. This framework will link technological design features and institutional governance structures to perceived and enacted agency outcomes. The sequential logic is important: the experiment provides mechanism-level clarity, while fieldwork examines how those mechanisms operate within real organizational settings.

The design explicitly addresses validity. Internal validity is supported through controlled manipulation in Phase 1. Construct validity is strengthened by distinguishing perceived agency from enacted agency. External validity is supported through field-based evidence from organizational contexts. Triangulation across experimental and qualitative data also reduces reliance on a single source of evidence.

4 Preliminary and Expected Results

Although data collection is not yet complete, the research anticipates several patterns. First, algorithmic transparency is expected to increase perceived agency by enabling managers to interpret recommendations rather than merely receive them (Bansal et al., 2021). However, transparency alone may not ensure enacted discretion if organizational norms discourage deviation from algorithmic outputs.

Second, override flexibility is expected to shape enacted agency. When deviation from AI recommendations is easy, legitimate, and low-friction, managers may be more willing to integrate contextual judgment. When override requires extensive documentation, approval, or exposes managers to scrutiny, they may default to algorithmic recommendations despite reservations. This suggests that substitution may arise from procedural design rather than direct automation.

Third, accountability structures are expected to moderate the relationship between AI design and agency. If managers remain clearly responsible for outcomes, they may engage more actively with AI outputs. If responsibility is diffused or attributed to algorithmic objectivity, managers may defer to system recommendations. This implies that responsible AI governance requires more than formal oversight; it requires governance structures that preserve legitimate discretion.

The expected theoretical contribution is a refined understanding of the augmentation-substitution continuum. Rather than treating augmentation or substitution as inherent properties of AI, the research argues that they are sociotechnically constructed through interactions between design, governance, and managerial practice. The distinction between perceived and enacted agency is also expected to be important, since managers may feel included in decision loops while still having limited practical discretion.

Practically, the research is expected to inform the design of AI-enabled decision systems that preserve meaningful human involvement. Potential implications include explanation interfaces that support interpretation, low-friction override pathways, explicit accountability mapping, and performance metrics that recognize contextual judgment rather than rigid adherence to algorithmic outputs.

5 Future Development

The next stage of this doctoral research is to finalize the experimental design and conduct pilot testing. Pilot work will refine the transparency manipulation, calibrate override conditions, and strengthen the measurement of perceived control and responsibility attribution. The main experiment will then form the basis of the first dissertation paper.

Following the experimental phase, the project will move toward organizational fieldwork. The field phase will require access to organizations using AI-enabled decision systems in areas such as hiring, performance evaluation, financial forecasting, or resource allocation. Theoretical sampling will be used to capture variation in governance structures and override protocols.

In the medium term, findings from the experimental and field phases will be integrated into a framework linking AI design, governance, perceived agency, and enacted discretion. This integrative work is intended to form a third dissertation paper and to support journal submissions in Information Systems and digital innovation outlets. Over the longer term, the research may be extended to compare regulatory or industry contexts and to develop prescriptive design guidance for responsible AI implementation.

Overall, the research trajectory moves from mechanism identification to contextual explanation and then to integrative theory building. This progression is designed to ensure both conceptual rigor and practical relevance within the doctoral timeframe.

6 Conclusion

This doctoral project contributes to Information Systems research by clarifying how technological design and institutional governance jointly shape human agency in AI-enabled decision systems. By combining experimental and field-based approaches, it aims to move beyond generic calls for human-in-the-loop systems toward a structured understanding of meaningful human involvement. The central argument is that AI does not automatically augment or substitute human judgment. Instead, these outcomes depend on how systems are designed, governed, and enacted in organizational practice.

References

Bansal, G., Wu, M., Zhou, J., Fok, R., & Kamar, E. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary human-AI performance. Proceedings of the ACM Conference on Human Factors in Computing Systems, 1-16. https://doi.org/10.1145/3411764.3445717

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. MIS Quarterly, 45(3), 1433-1450. https://doi.org/10.25300/MISQ/2021/16274

Burton, J. W., Stein, M. K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. Journal of Behavioral Decision Making, 33(2), 220-239. https://doi.org/10.1002/bdm.2155

Cancela-Outeda, C. (2024). The EU's AI Act: A framework for collaborative governance. Technology in Society, 78, 102565. https://doi.org/10.1016/j.techsoc.2024.102565

European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. Information and Organization, 28(1), 62-70. https://doi.org/10.1016/j.infoandorg.2018.02.005

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals, 14(2), 627-660. https://doi.org/10.5465/annals.2018.0057

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business Horizons, 61(4), 577-586. https://doi.org/10.1016/j.bushor.2018.03.007

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. Academy of Management Annals, 14(1), 366-410. https://doi.org/10.5465/annals.2018.0174

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90-103. https://doi.org/10.1016/j.obhdp.2018.12.005

Meijerink, J., Boons, M., Keegan, A., & Marler, J. (2021). Algorithmic human resource management: Synthesizing developments and cross-disciplinary insights on digital HRM. The International Journal of Human Resource Management, 32(12), 2545-2562. https://doi.org/10.1080/09585192.2021.1925326

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. Journal of Strategic Information Systems, 34(2), 101885. https://doi.org/10.1016/j.jsis.2024.101885

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation-augmentation paradox. Academy of Management Review, 46(1), 192-210. https://doi.org/10.5465/amr.2018.0072

Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. California Management Review, 61(4), 66-83. https://doi.org/10.1177/0008125619862257

DOCTORAL CONSORTIUM

CREATING BUSINESS VALUE IN UNPREDICTABLE TIMES AIMING FOR ANTIFRAGILITY

EDZO A. BOTJES, TIM HUYGH, LAURY BOLLEN, REMKO W. HELMS

Open Universiteit, Heerlen, the Netherlands
edzo.botjes@ou.nl, tim.huygh@ou.nl, laury.bollen@ou.nl, remko.helms@ou.nl

Cyber resilience is conceptualized as a socio-technical form of risk management aimed at mitigating the business impact of both foreseen and unforeseen disruptions. While traditional cybersecurity overlaps with this domain, true resilience requires more than mere impact reduction; organizations must extract strategic value from the unknown, a state characterized as antifragility. Organizational learning serves as the fundamental mechanism for fostering this antifragile posture. This PhD research project proposes the development of a Cyber Resilience Maturity Model (CRMM) designed to facilitate organizational learning and enable firms to gain value from uncertainty.

DOI
https://doi.org/10.18690/um.fov.4.2026.73

ISBN
978-961-299-152-4

Keywords:
antifragility, cyber resilience, cyber security, enterprise risk management, enterprise architecture, maturity models

University of Maribor Press

1 Introduction

Modern society is characterized by the seamless integration of organizations and individuals within a complex network of information and processes, creating an environment defined by volatility, uncertainty, complexity, and ambiguity (VUCA) (Bennett & Lemoine, 2014). Disruptions within a value chain result in significant impact at global, national, and individual levels. A series of events has contributed to the creation of multiple stringent legislative frameworks, e.g., NIS2, RED, CRA, DORA, PLR, and the MDR (Reeder & Hall, 2021; Tasheva & Kunkel, 2022), each of which prescribes the application of (existing) risk management and cyber security controls and practices.

Virtually all enterprise risk management and cyber security frameworks incorporate the fundamental steps of identification, treatment and monitoring of possible threats (Diefenbach et al., 2019). However, this approach often proves insufficient due to the inherent unpredictability of an organization's internal and external contexts, which cannot be adequately addressed through deterministic checklists (Aven, 2015; Boeing, 2016; Metcalf et al., 2021). In addition, risk management and cyber security traditionally start by identifying value, followed by managing the protection of this value. Security is achieved when expectations regarding a specific value overlap with actual observations (Huang et al., 2010; Schneier, 2008). Hence, the security of value is dynamic in a VUCA context, as is the value itself (Van Gils, 2023), rendering investments to secure value essentially unfalsifiable (Herley, 2016).

The primary goal of both risk management and cybersecurity is to ensure business continuity. Both disciplines encompass the organization and its environment, where the scope of causality is dominated by the interplay between human and machine, also addressed as 'cyber'. These commonalities lead to the unification of risk management and cyber security into cyber resilience (Eling et al., 2021; Linkov & Kott, 2019).

2 Problem definition

Where traditional approaches focus on managing potential threats (the "known unknown"), cyber resilience needs to address the "unknown unknown" (Aven, 2017). Recognizing that our reality is inherently unpredictable, Taleb (2012) coined

the term antifragile as the antithesis of fragile. Whereas fragility leads to diminished value caused by stress, antifragility leads to enhanced value caused by stress, and is therefore closely aligned with resilience (e.g., Axenie et al., 2024).

Martin-Breen and Anderies (2011) distinguish three subtypes of resilience within the literature. First, Engineering Resilience where the function (what) and the construction (how) are well engineered to uphold the continuity of value delivery. Second, Systems Resilience where the construction can alter, to uphold the function to deliver value. Finally, Complex Adaptive Systems (CAS) Resilience where the function and construction both can be altered to deliver a more suitable value. The concept of antifragility extends beyond these forms of resilience (Hillson, 2023), by optimizing the conditions for CAS Resilience while simultaneously integrating the two other subtypes of resilience to mitigate fragility toward "known" and "known unknown" events. Although various studies have investigated the requirements for achieving antifragility (e.g., Adobor and Kudonoo, 2025; Axenie et al., 2024; Botjes et al., 2021), none have yet been incorporated into cyber resilience guidance, such as a framework or maturity model.

Maturity models are a popular tool for guiding organizational management (van Looy et al., 2017). These models define discrete maturity levels for specific classes of objects, representing qualitative and/or quantitative stages of capability (Becker et al., 2009). By evaluating the current state, these models provide insights into progress and provide insights used to create roadmaps for transitioning from the present to a desired future state. Maturity models are widely used as sensemaking tools, across hundreds of disciplines in both business and academia (Adekunle et al., 2022).

While existing Cyber Resilience Maturity Models (CRMMs) offer valuable guidance, such as roadmapping, they often overlook two critical concepts. Firstly, the concept of antifragility as the capacity to not only withstand disruptions, but to gain value from them, thereby optimizing resilience. Secondly, the concept of the learning organization, which emphasizes fostering a culture that advances in resilience maturity, and prepares for unknown challenges. These omissions limit the effectiveness of CRMMs in addressing the dynamic and evolving nature of both internal and external organizational contexts.

Critiques on maturity models, and knowledge artifacts in general, often center on the lack of replication regarding causal relationships (Adekunle et al., 2022; Gräubig & Bley, 2023; Monteiro & Maciel, 2020) and the subjective perspective in sensemaking, namely, the notion that what is unknown cannot be observed (Aven, 2015; Lasrado, 2018; Uppström, 2017; Vom Brocke et al., 2020). These two points of critique are fundamentally rooted in the volatile, uncertain, complex, and ambiguous (VUCA) nature of both internal and external organizational contexts. This research systematically accounts for these concerns within its methodology.

Our research objective is to address the following central research question: *How can organizations exploit a chaotic context to generate more business value?* We formulated the following research questions;

RQ1: *What is a Cyber Resilience Maturity Model (CRMM) that adequately addresses the exploitation of unforeseen events?*

RQ2: *How does the (extended) Cyber Resilience Maturity Model (CRMM) offer guidance and value when applied by organizations?*

RQ3: *What are the key lessons learned from the application of the Cyber Resilience Maturity Model (CRMM) and how can these lessons be incorporated into a refined version of the CRMM?*

Upon answering these questions, this PhD research will deliver a validated Cyber Resilience Maturity Model (CRMM) designed to facilitate organizational learning in the face of chaos. This novel framework transcends traditional, static risk management approaches by providing a pathway for organizations to navigate 'unknown unknowns'. Ultimately, this research will demonstrate that by embracing antifragile behavior and enabling organizational learning, organizations can do more than merely withstand disruptions; they can actively gain value from exposure to the 'unknown unknowns'.

3 Methodology

Embarking onto research of cyber resilience implies an inquiry in the socio-technical domain, which presents significant challenges regarding generalization and replication (Kaleliglu, 2021; Pyle, 2023; Stevenson, 2023). Replication is essential to

prove the causal relationship between actions and results. The lack of research on the causality between actions and results is a central critique on maturity models (Monteiro & Maciel, 2020).

Critical realism as a research lens offers a way to address socio-technical reality while accounting for the replication crisis and the generalization challenge (Edwards et al., 2014; Tsang & Kwan, 1999; Uppström, 2017). Critical realism maintains that an objective reality exists independently of our subjective observations; consequently, retroduction is applied to uncover underlying, unobservable structures.

In line with this perspective, and respecting the 'wicked' nature of the problem domain, we apply iterative Design Science Research (DSR), by integrating the eADR (Mullarkey & Hevner, 2019) and the eDSR (Tuunanen et al., 2024) into five phases: (1) Diagnosis, (2) Design Objectives and Requirements, (3) Design and Development, (4) Implementation, and (5) Evolution. Each phase follows five steps: (a) Problem Formulation, (b) Artifact Creation, (c) Evaluation, (d) Reflection, and (e) Formalization of Learning (Figure 1). The aim of this approach is to optimize learning during the research and deliver intermediate (small) knowledge artifacts to be validated and reused (Hevner et al., 2024; Maedche et al., 2024).

Figure 1: Applied research method based on eADR (Mullarkey & Hevner, 2019) and on eDSR (Tuunanen et al., 2024)

We embrace open science (UNESCO & Canadian National Commission for UNESCO, 2022) and the FAIR-principles (Wilkinson et al., 2016) in support of replicability, independence, precision, and falsifiability (Recker, 2021). Therefore, our research infrastructure (Botjes et al., 2026b) consists out of primarily open-source tooling for each of the four phases from the open science framework: search and discover, design study, collect and diagnose data, and publish reports (Open Science Collaboration, 2015).

4 Preliminary/Expected results

During our iterations in and between the *diagnosis phase* and the *design objectives and requirements phase* (Figure 1), we formalized CRMM design goals and design principles based on relevant constructs and kernel theories. These form requirements for a CRMM that are used for the iterations of the *design and development phase*. In multiple iterations, in which we searched in literature and among experts, we were unable to find an existing CRMM (related) framework (Botjes et al., 2026a) which adheres to the formulated design goals and principles.

The focus of existing frameworks is on keeping the status quo, by determining the variety of the external context and to dampen it for the internal context. Our current objective for the next iterations of the *design and development phase* is how to capture these observations in a rigorous way to be replicated by others. Key challenges in achieving this include a lack of clear documentation on how the identified frameworks were constructed, and few experts are available to validate our observations for multiple frameworks.

Currently we are running various experimental iterations of the *design and development phase* to determine the viability of extending existing CRMM (related) frameworks. We do this together with masters students at multiple universities to determine the viability of extending existing models to adhere to the requirements from the *design objectives and requirements phase*. If not, the conclusion will be that it is inevitable to create a new CRMM that adheres to the requirements (caputered as design goals and design principles) and is grounded in the academic knowledge base.

In support of the *evaluation step* in the *design and development phase* and future replication iterations of this phase, we created an open data set of existing frameworks for cyber resilience practitioners to collaborate on.

5 Future development

Based on our analysis in the *diagnosis phase* and on the preliminary requirements being the design goals and principles as deliverable from the *design objective and requirements phase* we are currently of the opinion that it will probably not be viable to alter existing CRMM (related) frameworks, to cover all the formulated requirements. The existing frameworks are focused on not only minimizing the impact, but also focused on a specific expert domain, and thus not designed to offer guidance to non-experts. Extending existing frameworks to cover as many as possible of the CRMM requirements has the upside of utilizing the current adoption in the field by experts and practitioners of these frameworks. This will improve the existing resilience of organizations, which in itself adds to the purpose of business continuity.

We hence foresee the need for creating a new CRMM which enables decision-makers of varying experience and backgrounds, to make sense of their current domain and envision future steps to evolve towards antifragility. This might also imply adopting language in the CRMM that resonates well with practitioners, to encourage adoption and application. Therefore, the research might need to deliver multiple models where one is the core model, which adheres to a high level of scientific rigor, and other models which are derivatives of the core model that are tailored to relevant domains in practice.

We also anticipate a challenge in shifting the mental models of current practitioners. For instance, many risk managers are conditioned, or primed, to associate risk solely with negative outcomes. As a result, even if risk management is defined to inlcude opportunities, as in ISO 31000, this may not automatically translate into the integration of opportunity management within the actual risk management practices. This is an example of the tension between scientific rigor, and model adoption and application.

Therefore, the *design and development phase* will deliver multiple models, of which each will have their own evaluation step for validation. We will select multiple models (at least three) from the *design and development phase* to enter the *implementation phase*. In this phase we will apply each model in multiple groups, to measure the perceived value to reach the defined design goals.

As the proposed design goals and principles do not prescribe specific sectors or criteria for applying the CRMM to improve maturity in resilience, we will favor a randomized approach for group selection model application. The model(s) will be evaluated on their perceived overall value, their contribution to the design goals, their coverage of the design principles, and the frequency of learning that is supported. The outcomes of the formulation of learning steps of the various iterations over the *implementation phase*, will be used as input to the *evolution phase* looping back to the *diagnosis phase* for re-evaluation of the problem and solution domain.

During our research for a CRMM, which enables organizations to move towards antifragility, we will provide various validated building blocks to add to the design knowledge on how to achieve antifragility as an organization to uphold business continuity in a VUCA context.

References

Adekunle, S. A., Aigbavboa, C., Ejohwomu, O., Ikuabe, M., & Ogunbayo, B. (2022). A critical review of maturity model development in the digitisation era [Publisher: MDPI AG Publisher: Multidisciplinary Digital Publishing Institute]. Buildings, 12(6), 858. https://doi.org/10.3390/buildings12060858

Adobor, H., & Kudonoo, E. C. (2025). Antifragility and organizations: An organizational design perspective. Leadership & Organization Development Journal. https://doi.org/10.1108/LODJ-03-2024-0185

Aven, T. (2015). Implications of black swans to the foundations and practice of risk assessment and management [Publisher: Elsevier]. Reliability Engineering & System Safety, 134, 83–91. https://doi.org/10.1016/j.ress.2014.10.004

Aven, T. (2017). A conceptual foundation for assessing and managing risk, surprises and black swans [ISSN: 2191-530X]. In The illusion of risk control: What does it take to live with uncertainty? (pp. 23–39). Springer International Publishing. https://doi.org/10.1007/978-3-319-32939-0

Axenie, C., López-Corona, O., Makridis, M. A., Akbarzadeh, M., Saveriano, M., Stancu, A., & West, J. (2024). Antifragility in complex dynamical systems [Publisher: Springer Science and Business Media LLC]. Nature Partner Journals (NPJ) Complexity, 1(12). https://doi.org/10.1038/s44260-024-00014-y

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing maturity models for IT management: A procedure model and its application [Publisher: Springer Nature]. Business & Information Systems Engineering, 1(3), 213–222. https://doi.org/10.1007/s12599-009-0044-5

Bennett, N., & Lemoine, G. J. (2014). What a difference a word makes: Understanding threats to performance in a VUCA world. [Publisher: Elsevier BV]. Business Horizons, 57(3), 311–317. https://doi.org/10.2139/ssrn.2406676

Boeing, G. (2016). Visual analysis of nonlinear dynamical systems: Chaos, fractals, self-similarity and the limits of prediction [Publisher: Multidisciplinary Digital Publishing Institute]. Systems, 4(4), 37. https://doi.org/10.3390/systems4040037

Botjes, E. A., Huygh, T., Bollen, L., & Helms, R. (2026a, January 21). Cyber resilience maturity model keywords. https://doi.org/10.5281/ZENODO.18323583

Botjes, E. A., Huygh, T., Bollen, L., & Helms, R. (2026b, February 4). Open research infrastructure - CRMM. Zenodo. https://doi.org/10.5281/ZENODO.18487444

Botjes, E. A., van den Berg, M., van Gils, B., & Mulder, H. (2021). Attributes relevant to antifragile organizations. In J. P. A. Almeida, D. Bork, G. Guizzardi, M. Montali, H. A. Proper, & T. P. Sales (Eds.), 2021 IEEE 23nd conference on business informatics (CBI) (pp. 62–71, Vol. 01). IEEE. https://doi.org/10.1109/CBI52690.2021.00017

Diefenbach, T., Lucke, C., & Lechner, U. (2019). Towards an integration of information security management, risk management and enterprise architecture management – a literature review [Place: Sydney, Australia Journal Abbreviation: 2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)]. 2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), 326–333. https://doi.org/10.1109/CloudCom.2019.00057

Edwards, P. K., O'Mahoney, J., & Vincent, S. (2014). Studying organizations using critical realism: A practical guide (First edition, Vol. 15) [ISSN: 1476-7430 Publication Title: Journal of Critical Realism]. Oxford University Press. https://search.worldcat.org/en/title/847725011

Eling, M., McShane, M., & Nguyen, T. (2021). Cyber risk management: History and future research directions [Publisher: Blackwell Publishing Ltd]. Risk Management and Insurance Review, 24(1), 93–125. https://doi.org/10.1111/rmir.12169

Gräubig, D. M., & Bley, K. (2023). A configurational set-theoretic approach to an innovation capability maturity model. AMCIS 2023 Proceedings, 13. https://aisel.aisnet.org/amcis2023/sig_scuidt/sig_scuidt/13/

Herley, C. (2016). Unfalsifiability of security claims. Proceedings of the National Academy of Sciences, 113(23), 6415–6420. https://doi.org/10.1073/pnas.1517797113

Hevner, A. R., Parsons, J., Brendel, A. B., Lukyanenko, R., Tiefenbeck, V., Tremblay, M. C., & Vom Brocke, J. (2024). Transparency in design science research. Decision Support Systems, 182, 114236. https://doi.org/10.1016/j.dss.2024.114236

Hillson, D. (2023). Beyond resilience: Towards antifragility? Continuity & Resilience Review, 5(2), 210–226. https://doi.org/10.1108/CRR-10-2022-0026

Huang, D.-L., Rau, P.-L. P., & Salvendy, G. (2010). Perception of information security [Publisher: Informa UK Limited]. Behaviour & Information Technology, 29(3), 221–232. https://doi.org/10.1080/01449290701679361

Kaleliglu, U. B. (2021). Replicability: 21st century crisis of the positivist social sciences. International Journal of New Approaches in Social Studies, 2021, 5(2), 400–425. https://doi.org/10.38015/sbyy.1003103

Lasrado, L. A. (2018). Set-theoretic approach to maturity models [Doctoral dissertation, Copenhagen Business School] [OCLC: 1175940135]. https://www.econstor.eu/bitstream/10419/209064/1/cbs-phd2018-15.pdf

Linkov, I., & Kott, A. (2019). Fundamental concepts of cyber resilience: Introduction and overview [Journal Abbreviation: Cyber Resilience of Systems and Networks]. In A. Kott & I. Linkov (Eds.), Cyber resilience of systems and networks (pp. 1–25, Vol. abs/1806.02852). Springer International Publishing. https://doi.org/10.1007/978-3-319-77492-3_1

Maedche, A., Elshan, E., Höhle, H., Lehrer, C., Recker, J., Sunyaev, A., Sturm, B., & Werth, O. (2024). Open science: Towards greater transparency and openness in science [Publisher: Springer Nature]. Business & Information Systems Engineering. https://doi.org/10.1007/s12599-024-00858-7

Martin-Breen, P., & Anderies, J. M. (2011). The bellagio initiative, background paper, resilience: A literature review [event-place: Brighton:IDS]. Resilience: A Literature Review. http://opendocs.ids.ac.uk/opendocs/handle/123456789/3692

Metcalf, G. S., Kijima, K., & Deguchi, H. (Eds.). (2021). Handbook of systems sciences. Springer Singapore. https://doi.org/10.1007/978- 981- 15-0720-5

Monteiro, E. L., & Maciel, R. S. P. (2020). Maturity models architecture: A large systematic mapping [Publisher: Sociedade Brasileira de Computacao - SB]. iSys - Brazilian Journal of Information Systems, 13(2), 110–140. https://doi.org/10.5753/isys.2020.761

Mullarkey, M. T., & Hevner, A. R. (2019). An elaborated action design research process model (P. Ågerfalk, Ed.) [Publisher: Informa UK Limited]. European Journal of Information Systems, 28(1), 6–20. https://doi.org/10.1080/0960085X.2018.1451811

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science [Publisher: American Association for the Advancement of Science]. Science, 349(6251), aac4716. https://doi.org/10.1126/science.aac4716

Pyle, B. (2023). Knowledge generation and uncertainty in an unpredictable social world. Boston University Law Review, 103(7), 2049–2063. https://www.bu.edu/bulawreview/files/2023/12/PYLE-.pdf

Recker, J. (2021). Scientific research in information systems: A beginner's guide [ISSN: 2196-8705 Publication Title: Progress in IS]. Springer International Publishing. https://doi.org/10.1007/978-3-030-85436-2

Reeder, J. R., & Hall, T. (2021). Cybersecurity's pearl harbor moment: Lessons learned from the colonial pipeline ransomware attack. The Cyber Defense Review, 6(3), 15–40. https://www.jstor.org/stable/48631153

Schneier, B. (2008). The psychology of security [Series Title: Lecture Notes in Computer Science]. In S. Vaudenay (Ed.), Progress in cryptology –AFRICACRYPT 2008 (pp. 50–79, Vol. 5023). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68164-9_5

Stevenson, M. T. (2023). Cause, effect, and the structure of the social world. Boston University Law Review, 103, 2001–2047. https://doi.org/10.21428/cb6ab371.77eda2fb

Taleb, N. N. (2012, November). Antifragile: Things that gain from disorder [ISSN: 1059-5422 Publication Title: Competitiveness Review: An International Business Journal]. Random House. https://doi.org/10.1108/10595421311319852

Tasheva, I., & Kunkel, I. (2022). In a hyperconnected world, is the EU cybersecurity framework connected? [Publisher: SAGE Publications]. European View, 21(2), 186–195. https://doi.org/10.1177/17816858221136106

Tsang, E. W. K., & Kwan, K.-M. (1999). Replication and theory development in organizational science: A critical realist perspective [Publisher: Academy of Management]. The Academy of Management Review, 24(4), 759. https://doi.org/10.2307/259353

Tuunanen, T., Winter, R., & Vom Brocke, J. (2024). Dealing with complexity in design science research: A methodology using design echelons. MIS Quarterly, 48(2), 427–458. https://doi.org/10.25300/MISQ/2023/16700

UNESCO & Canadian National Commission for UNESCO. (2022, November). An introduction to the UNESCO recommendation on open science. UNESCO. https://doi.org/10.54677/XOIR1696

Uppström, E. (2017). Designing, theorizing, and reflecting on information systems artifacts and value co-creation in e-government [Doctoral dissertation, Department of Computer and Systems Sciences, Stockholm University]. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1092699&dswid=-2989

van Looy, A., Poels, G., & Snoeck, M. (2017). Evaluating business process maturity models. Journal of the Association for Information Systems, 18(6), 461–486. https://doi.org/10.17705/1jais.00460

Van Gils, B. (2023). Data in context: Models as enablers for managing and using data. Springer Nature Switzerland. https://doi.org/10.1007/978- 3-031-35539-4

Vom Brocke, J., Winter, R., Hevner, A., & Maedche, A. (2020). Special issue editorial – accumulation and evolution of design knowledge in design science research: A journey through time and space. Journal of the Association for Information Systems, 21(3), 520–544. https://doi.org/10.17705/1jais.00611

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3(1), 160018. https://doi.org/10.1038/sdata.2016.18

About the authors

Edzo A. Botjes MSc is a PhD candidate in cyber resilience at the Open Universiteit, the Netherlands. His research interests are on how to improve business continuity and focuses on the overlap between risk management, cyber security, enterprise architecture, enterprise governance, resilience, and is specialised in antifragility. He is supervising several MSc students at the Antwerp Management School, and is an external teacher for the MSc course Enterprise Architecture at the University of Applied Sciences Utrecht. His research activities resulted in (cited) publications.

Prof.dr.ir. **Remko Helms** is a professor in Business Analytics at Open Universiteit (the Netherlands) in the Information Systems group. His teaching and research interests are mainly focused on how organizations realize value from exploring and exploiting data, including topics such as data analytics processes, data analytics strategy, data governance and analytics governance. He is supervising several Phd students and groups of master students on these topics. His work appeared in different Information Systems and Knowledge Management conferences as well as journals. He is a member of Association for Information Systems (AIS).

Dr. **Laury Bollen** is an associate professor of Information Management at the Open Universiteit, the Netherlands. He holds a PhD in Business Economics from Maastricht University. His research interest focuses on IT Governance, Information Security Governance and Platform Governance.
His research activities have resulted in publications in a broad range of international academic journals including Information & Management, the Journal of Information Systems, Information Systems Management, the Journal of Cleaner Production, the European Accounting Review, and the Journal of Business Ethics.

Dr. **Tim Huygh** is assistant professor at the department of information science at the faculty of science of the Open Universiteit, the Netherlands. He is also a visiting professor at the Antwerp Management School. He received his PhD in 2019 on IT governance from the University of Antwerp. His research interests include the governance and (strategic) management of information and technology, information security governance, and cyber resilience. His research has been published in SCI-indexed journals like Information Systems Journal (ISJ) and Decision Support Systems (DSS), and various conference proceedings including the International Conference on Information Systems (ICIS) and the Hawaii International Conference on System Sciences (HICSS). He also co-authored three books on the topic of IT governance (published by Springer). Since 2020, he is co-chairing the minitrack "IT governance and its mechanisms" at HICSS.

DOCTORAL CONSORTIUM

CONCEPT MINING FOR DATA MANAGEMENT

DOI
https://doi.org/10.18690/um.fov.4.2026.74

ISBN
978-961-299-152-4

SWARUPA HARDIKAR[1, 2]

[1] HAN University of Applied Sciences, Arnhem, the Netherlands
swarupa.hardikar@han.nl
[2] Radboud University, Nijmegen, the Netherlands
swarupa.hardikar@han.nl

Keywords:
concept mining,
data management,
text mining,
terminology extraction,
term disambiguation

The PhD project addresses data ambiguity, a core issue hampering the management of organisational data, and aims to support data management through concept mining. We intend to (semi-)automatically analyse perceived meanings of terms and concepts within an organisation's data domain to aid terminological disambiguation. The concept mining system will deploy text mining and information retrieval techniques within unstructured data. We adopt a 3-phase approach of exploring the problem and requirements, iterative development and testing, and finally, an extrinsic evaluation in a real-life setting. Every component of the system will be intrinsically evaluated with metrics best suited to that component. The toolkit will also be evaluated qualitatively through studies with data managers for efficiency, quality and real-world applicability. The project will bridge the gap between domain expertise and automated methodologies, with expected contributions such as an evaluation framework for concept mining and advancement towards the research of terminology extraction.

University of Maribor Press

1 Introduction

1.1 The Importance of Data Meaning

Across organisational domains, people use language differently, and data owners may frequently develop their own interpretations of key terms (Alexopoulos, 2020). This may cause ambiguity in communication and a lack of data overview - for example, the term "model" might describe a machine learning system, a conceptual representation, or a 3D simulation depending on the domain the person using it operates in. It has been estimated by many studies that decision-making caused by the poor quality of information resulted in losses of thousands of dollars within organisations (Castillo et al., 2017). Per a McKinsey report, an average of 30% of time is spent by organisations on non-value added tasks due to poor data quality and availability (Petzold et al., 2020).

To reduce ambiguity and streamline communication, managing business vocabularies and clarifying the definitions of terms is crucial (Ross, 2020; Voskuil, 2015; Hoppenbrouwers, 2003). Consequently, to communicate about data meaning, having a description of the semantics is key.

1.2 The Challenge

Curating lexical assets such as business glossaries is typically the responsibility of a data steward or manager - however, the amount of organisational data can make the task resource-intensive (Henderson et al., 2017).

From multiple interviews with data management professionals, it was confirmed that data managers require insight into organisational term usage and meaning to make informed decisions about useful terms, concepts, and semantic descriptions. It was articulated more specifically that organisations are currently largely facing an issue with data semantics. Interviewees stated that reconciling terminological ambiguity was an active pain point which could potentially result in monetary losses.

As effective data management relies on the combination of qualitative domain expertise and automated methodologies, our approach respects the DAMA Data Management Body of Knowledge principles (Henderson et al., 2017) by focusing on the structured extraction and validation of domain-specific lexical assets.

Currently, such ambiguity is resolved by way of multiple consultations and meetings with domain experts. On the other hand, such domain experts are often a scarce resource (personal correspondence, Data Architect), and in addition, interviewees noted that such prolonged practices can contribute to employee fatigue (personal correspondence, Domain Architect). As such, we believe that (semi-)automated approaches within concept mining may expedite the process as well as better reflect the ground reality.

1.3 Concept Mining for Data Management

Efforts into the extraction and analysis of domain-specific terms within text data have been observed within the fields of natural language processing and terminology extraction (a subtask of information extraction). However, the latter still suffers from issues such as polysemy (usage of terms within different contexts) and domain adaptability (generalisability across a variety of domains) (Xu et al., 2025).

Given that effective data management relies on the marriage of qualitative domain expertise and automated methodologies, our approach aligns with the DAMA DM-BOK (Data Management Body of Knowledge) principles (Henderson et al., 2017) by focusing on the structured extraction and validation of domain-specific lexical assets, while also drawing upon components from FAIR (findable, accessible, interoperable and reproducible; Wilkinson et al., 2015) guidelines. The project will bridge the gap between qualitative domain expertise and automated methodologies.

Thus, we believe Concept Mining (CM) could be used to support data management. CM has been applied in other contexts such as formalisation (Yang et al., 2008), information retrieval (Liu et al., 2019) or data sharing for medical applications (Si et al., 2019). In this project, we focus specifically on methods to support data managers to gain overview of perceived interpretations by data users, and to support selecting and recording definitions.

2 Proposed Research

Given the above considerations, we formulate the following research question:

How can a concept mining system be designed, applied, and evaluated to support data managers in analysing concepts based on perceived interpretations of terms by data users?

The research question addresses (a) the translation of the needs of data managers into requirements and selection of text mining techniques ("*how to design*"), (b) the development, comparison and evaluation of the detection and/or generation of synonyms, homonyms and definitions for a given term using CM ("*how to apply*"), and (c) an evaluation of the concept mining system based on efficiency, effectiveness and usefulness ("*how to evaluate*").

The research question is supported by the following sub-questions:

1. How can we extract information from relevant data to determine the relevance of a concept given a specific goal of a data manager?
2. What information does a data manager need to adequately manage concepts with multiple interpretations (homonyms)?
   How can we extract homonymous concepts from data?
3. What information does a data manager need to find related concepts (synonyms)?
   How can we identify synonyms in data?
4. What information does a data manager need to determine which definition fits a concept in a specific context?
   How can we select, link or generate possible definitions from data?
5. How should information be represented to support different types of meaning descriptions?

Figure 1 draws upon the DIKW (data-information-knowledge-wisdom) model (Wien et al., 2006) to encapsulate our contribution in the development of data and information into meaning and knowledge, with subject matter expertise contributing to contextualisation and validation (note: we do not specifically adhere to the model's 'wisdom' layer; it is out of scope of the project).

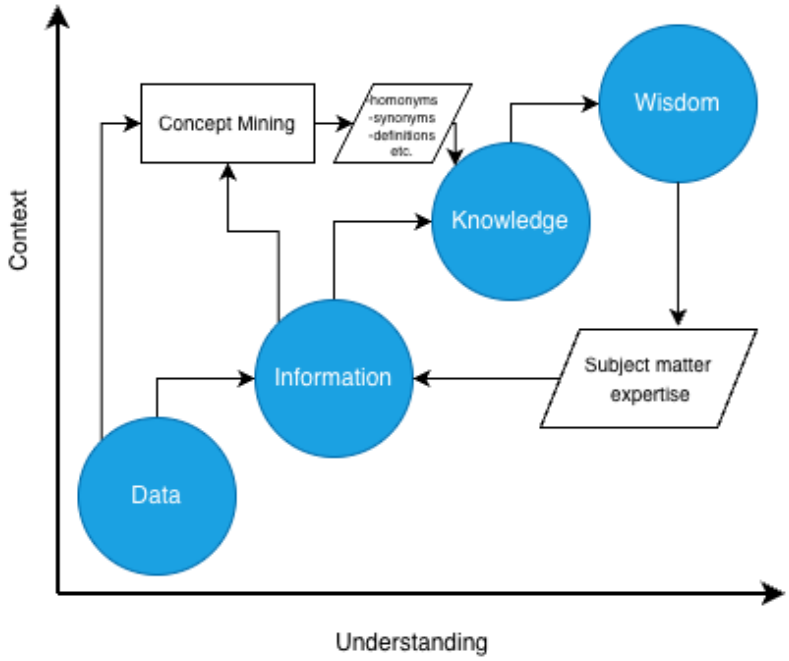

Figure 1: Concept Mining within the DIKW model
Source: Own

3 Research methodology and proposed experiments

Our research design is rooted in the Design Science Research (DSR) paradigm, aiming to improve scientific/technological bases by creating innovative artifacts that empower the working environments where they are implemented (Vom Brocke et al., 2020). In this case, the artifact would be a tool(kit) that deploys CM within organisations to support data managers.

The approach and methodology are also based on the DSR process model (Peffers et al., 2007). We first define the problem and objectives, i.e. the requirements (subtasks) of the solution (the concept mining toolkit). We then identify potential techniques that can help satisfy the requirements and test these techniques on the data. Finally, we evaluate the system quantitatively and qualitatively.

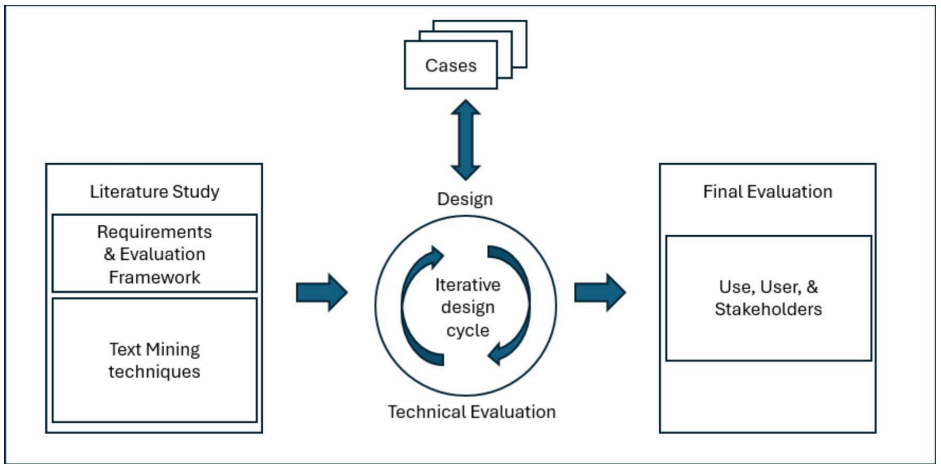

Figure 2: Phased design of the research approach based on the DSR model.
Source: Own

The three-phase approach spanning the duration of the PhD is shown below:

Phase 1: Exploring requirements and industry input

The phase is primarily a preparatory one. We have conducted an initial literature review to identify the issues surrounding good data management within organisations. We have also conducted research interviews with data professionals such as data architects, information managers and domain architects to validate the challenges of data management, the necessity of the informational value of data and the potential need for automating the extraction and analysis of concepts. Taking these into account, we have defined the research questions.

Phase 1 Steps:

i) A study of the relevant literature and analysis of SOTA text mining and IR techniques:
   a. Exploring pain points of data management and previous attempts at resolving them: Overview of literature
   b. Hypothesising and defining the problem

ii) Research interviews with data professionals (completed): We match the pain points discovered in the literature with findings from the industry, thereby defining the final problem statement

Phase 2: Iterative toolkit development via use cases and evaluation (Ongoing)

To implement the analyses conducted in the previous phase, we conduct a series of use case studies using concept mining. Each study consists of applying the aforementioned techniques, followed by evaluations. After every study, we aim to implement the findings into the next one, in combination to a revisitation of state-of-the-art (SOTA) techniques. In addition, we test the setup after each study for transferability and generalisability.

Phase 2 steps:

(i) Energy domain: abbreviation disambiguation for Dutch and English data (completed);
(ii) (ii-iii) Education and Banking domains. We also test the CM system for generalisability across domains.

Phase 3: Extrinsic evaluation of concept mining in a practical setting

In Phase 3, we plan to conduct a qualitative study involving experience, comments, feedback and advice from data management professionals from various organisations, pertaining to the utility and efficiency of the concept mining toolkit in various downstream data management tasks. The presentation of the concept mining system will be in the form of demos and live workshops where the user (typically a data professional) can implement it with their data. This will be followed by surveys and questionnaires that can be answered by the user, assessing the accuracy, efficiency and usability of the system.

Phase 3 steps:

(i) Qualitative assessment of concept mining
   a. Surveys and questionnaires to gather post-demo or post-live workshop responses from users

(ii) Test strategies based on the requirements of previous phases

4 Preliminary Results from Use Case #1: Acronym Disambiguation

Our first study explored the disambiguation of acronyms within domain-specific data with Alliander, a major energy company in the Netherlands. We investigate the applicability of state-of-the-art methods for multilingual acronym disambiguation (AD) on an open benchmark and an authentic domain-specific bilingual dataset to establish transfer from theory to practice.

Acronyms are often homonymous, with expansions that have similar or completely different meanings. It can be difficult to disambiguate them within organisational text, which may cause misinterpretation in downstream tasks, especially with in-the-wild domain- or company-specific acronyms. For example, the acronym 'AC' may stand for "Alternating Current" or "*Asbestcement*" (*Dutch: Asbestos cement*). The company lists numerous acronyms in use, with many highly domain-specific and company-specific usages. We found 176 acronyms with multiple long forms (expansions) each, which has raised concerns regarding their interpretation within the company (especially their usage within large language models).

Thus, we conducted a two-phase exploration for highly domain-specific and bilingual AD, creating a dataset of 176 examples for domain-specific bilingual (Dutch+English) data to help us reflect in-the-wild acronym usage in research.

Experimental setup:

We first curated a bilingual AD dataset with original Dutch and English examples from the company data. For a second dataset, we used an open benchmark for scientific acronym disambiguation in English.

We then tested the performance of two language models on the two datasets: AcroBERT (a pre-trained English AD language model) (Chen et al., 2023) and Ministral-3 8B (an open generative language model) (Liu et al., 2026), focusing on AD on highly domain-specific acronyms.

As an alternative resource-friendly tactic in the possible absence of an internal knowledge base, we also evaluated the use of Web search result pages in locating the correct acronym expansion given context keywords. This was done using DDGS,[1] a meta-search Python library for retrieving results across multiple search engines.

Results and Conclusion:

Table 1: Breakdown of performance on the Company Dataset by domain specificity and language

Category	Subset	Model	Accuracy
Domain-specificity	Company-specific	Ministral	0.5753
		AcroBERT	0.5342
	Domain-specific	Ministral	0.7869
		AcroBERT	0.5738
	Global	Ministral	0.9048
		AcroBERT	0.5952
Language	English	Ministral	0.8077
		AcroBERT	0.7692
	Dutch	Ministral	0.6867
		AcroBERT	0.5267

Source: Own

Table 2: Acronym expansion matches found in Web search result pages (max. 50 aggregated results)

Category	Subset	Model	Recall
Overall	All acronyms	39/103	37.86
Domain-specificity	Global	22/42	52.38
	Domain-specific	17/61	27.87
Language	English	15/23	65.22
	Dutch	24/80	30.00

Source: Own

For the first phase, we found that Ministral-3 8B can often identify the appropriate expansion in a non-English, domain-specific setting without finetuning, but it cannot generalise well enough to unseen data. For AcroBERT, despite its lower

[1] https://pypi.org/project/ddgs/

scores, its comparable performance in correctly identifying company-specific acronym expansions could indicate that finetuning it with a Dutch AD dataset would already improve performance, without the need for training on company-specific data, making it a resource-efficient alternative that handles unseen data better.

For the usage of search results from the Web, we found that our attempt at the extraction of domain-specific acronym expansions did not yield satisfactory recall. However, we believe that there may be promise in using it as a low-resource tool in the absence of an internal knowledge base to get diverse perspectives on the real-world usage of a given acronym.

We thus concluded that there remains a gap between theoretical and practical implementations of AD tools for multilingual domain-specific text.

5 Future Development

The above study indicated that in the case of non-English and domain-specific text data, there is still room for improvement in getting theoretically tested models ready for application in practice. On the other hand, it was an experiment on a smaller scale, examining acronyms as homonymous terms encountered on a daily basis. For the next study, we aim to implement these findings to scale beyond acronym disambiguation and study term-level ambiguity in unstructured company text. In doing so, we expect to analyse homonymous (and synonymous) usages of terms of interest and provide an overview of these usages to aid terminological disambiguation.

As highlighted in Section 3, we plan on conducting analyses across domains, iteratively developing the concept mining system to ensure generalisability across domains. The next implementation will address the education domain, focusing on commonly used terms and concepts across departments (domains) within the same organisation. Consequently, we will compare the official definitions of the terms with their actual usage in documentation, thereby analysing diverse interpretations. The interpretations will then help expand the overview of the terms for a data manager and potentially aid disambiguation efforts.

Acknowledgment

The PhD project is supported by the DEMAND project under SVB/ SPR.ALG.02.015 and the HAN Doctoral Grant.

References

Alexopoulos, P. (2020). *Semantic modeling for data: avoiding pitfalls and breaking dilemmas* (First edition). Beijing Boston Farnham Sebastopol Tokyo: O'Reilly.

Castillo, L. F., Raymundo, C., & Mateos, F. D. (2017). Information Architecture Model for Data Governance Initiatives in Peruvian Universities. *Proceedings of the 9th International Conference on Management of Digital EcoSystems*, 104–107. Presented at the Bangkok, Thailand. doi:10.1145/3167020.3167036

Chen, L., Varoquaux, G., & Suchanek, F. M. (2023). GLADIS: A general and large acronym disambiguation benchmark. In Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023).

Henderson, D., Earley, S., & Data Administration Management Association (Eds). (2017). *DAMA-DMBOK: data management body of knowledge* (Second edition). Basking Ridge, New Jersey: Technics Publications.

Hoppenbrouwers, S. J., Proper, H. A., & van der Weide, T. P. (2005). A fundamental view on the process of conceptual modeling. *International Conference on Conceptual Modeling*, 128–143. Springer.

Liu, A. H., Khandelwal, K., Subramanian, S., Jouault, V., Rastogi, A., Sadé, A., Jeffares, A., Jiang, A., Cahill, A., Gavaudan, A., Sablayrolles, A., Héliou, A., You, A., Ehrenberg, A., Lo, A., Eliseev, A., Calvi, A., Sooriyarachchi, A., Bout, B., ... Ramzi, Z. (2026). Ministral 3. arXiv. doi:10.48550/arXiv.2601.08584Liu, B., Guo, W., Niu, D., Wang, C., Xu, S., Lin, J., … Xu, Y. (2019). A User-Centered Concept Mining System for Query and Document Understanding at Tencent. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1831–1841. doi:10.1145/3292500.3330727

Liu, B., Guo, W., Niu, D., Wang, C., Xu, S., Lin, J., … Xu, Y. (2019, July). A User-Centered Concept Mining System for Query and Document Understanding at Tencent. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1831–1841. doi:10.1145/3292500.3330727

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. doi:10.2753/MIS0742-1222240302

Petzold, Z., Roggendorf, M., Rowshankish, K., & Sporleder, C. (2020). Designing data governance that

delivers value. Retrieved from https://www.mckinsey.com/business-functions/mckinsey-digital/ourinsights/designing-data-governance-that-delivers-value#

Ross, R. G. (2020). *Business knowledge blueprints: enabling your data to speak the language of the business* (Second edition). Place of publication not identified: Business Rule Solutions, LLC.

Si, Y., Wang, J., Xu, H., & Roberts, K. (2019). Enhancing clinical concept extraction with contextual embeddings. *Journal of the American Medical Informatics Association*, *26*(11), 1297–1304. doi:10.1093/jamia/ocz096

Voskuil, J. (07 2015). Enterprise vocabulary management A lexicographic view. *Wacana*, *16*, 411. doi:10.17510/wacana.v16i2.384

Veyseh, A. P. B., Dernoncourt, F., Tran, Q. H., & Nguyen, T. H. (2020). What does this acronym mean? Introducing a new dataset for acronym identification and disambiguation. In D. Scott, N. Bel, & C. Zong (Eds.), Proceedings of the 28th International Conference on Computational Linguistics (pp. 3285–3301). International Committee on Computational Linguistics. doi:10.18653/v1/2020.coling-main.292

Vom Brocke, J., Hevner, A., & Maedche, A. (2020). Introduction to design science research. In *Design science research. Cases* (pp. 1–13). Springer.

Wien, J. J. F., Jansen, J. M. L., Zeeuw, K., & Roosenschoon, O. R. (01 2006). Frameworks and Interactive Tools for Scientific Knowledge Systematization. *MODSIM07 - Land, Water and Environmental Management: Integrated Systems for Sustainability, Proceedings*.

Xu, K., Feng, Y., Li, Q., Dong, Z., & Wei, J. (2025). Survey on terminology extraction from texts. *Journal of Big Data*, *12*(1), 29. doi:10.1186/s40537-025-01077-x

Yang, K.-M., Kim, E.-H., Hwang, S.-H., & Choi, S.-H. (2008). *Fuzzy concept mining based on formal concept analysis. 2*, 279–290.

About the author

I am a 2nd year PhD candidate at Radboud University (RU) and HAN University of Applied Sciences (where I also work as a junior researcher). I have a background in text mining, and my areas of interest include natural language processing, language modelling, information retrieval and data semantics. I am currently supervised by Prof.dr. Arjen de Vries (promoter, RU), Lec.dr. Stijn Hoppenbrouwers (co-promoter, HAN), and Dr. Maya Sappelli (daily supervisor, HAN).

Summary

The research addresses the problem of data ambiguity within organisations, which is known to be a core issue hampering data management. We aim to support data management through concept mining: (semi)-automatically analysing concepts within an organisation's data domain, given a list of terms of interest. In doing so, we intend to analyse perceived meanings of terms and concepts, support the detection of synonyms, homonyms and find or generate suitable definitions. The concept mining system is a "toolkit" for data managers, incorporating text mining techniques. It will be evaluated for efficiency, quality and real-world applicability.

DOCTORAL CONSORTIUM

VALUE-SENSITIVE DESIGN FOR RESPONSIBLE DATA SHARING IN GOVERNMENTAL ECOSYSTEMS

CHRISTEL VAN DE WAL

Open Universiteit, Faculty of Science, Heerlen, the Netherlands
christel.vandewal@ou.nl

Advances in data-sharing technologies offer governments opportunities to improve efficiency and public service delivery. However, ethical risks, legal constraints, and concerns about public trust often limit the adoption of data-sharing practices within governmental organisations. This study examines how Value-Sensitive Design (VSD) and collaboration technologies, including Distributed Ledger Technology (DLT) and Privacy Enhancing Technologies (PETs), can support responsible and trustworthy data sharing in public governance. The research addresses ethical, legal, and organisational tensions while safeguarding public values and fundamental rights. A mixed-method research design combines literature reviews, qualitative case studies, and empirical validation, structured according to design science research principles. Building on analyses of governance frameworks and case studies in Dutch governmental organisations, the study develops and operationalises a methodological and normative framework for responsible data sharing. The research adapts VSD to governmental data ecosystems and provides a structured design approach supporting ethical and privacy-preserving data-sharing practices in public administration.

DOI
https://doi.org/10.18690/um.fov.4.2026.75

ISBN
978-961-299-152-4

Keywords:
public value theory, value sensitive design, data sharing, multi party computation, distributed ledger technology, government, ecosystem

University of Maribor Press

1 Introduction

Data sharing holds the potential to enhance both economic and public value; however, concerns regarding privacy risks, data security, and public trust pose significant challenges[1]. In the Netherlands, these concerns pose significant barriers to the implementation of data-sharing initiatives, thereby limiting the ability of government organisations to leverage data for effective public service delivery (Staat van de Uitvoering, 2023). This observation is reinforced by Mark Vermeer, Director of Digital Government at the Ministry of the Interior and Kingdom Relations, who states: "When sharing data, you have to weigh the risks, and government organisations are not very fond of that... they are naturally risk-averse" (Vermeer, 2023, p. 48). A fundamental issue underlying these barriers is the complexity of continuously aligning ethical, legal, technical, societal, and organisational requirements for data sharing (Agahari et al., 2022; Rukanova et al., 2023; The Royal Society, n.d.; Susha et al., 2023). While policies, IT processes, and systems are often designed with stability in mind, evolving societal values and shifting contexts can disrupt their alignment with public needs (Van de Poel, 2021, 2024). In such cases, previously unconsidered norms or design specifications may become necessary to ensure that organisational processes and IT infrastructures remain relevant and effective. The complexity of balancing ethical principles arises from the often-conflicting interests of stakeholders (Agahari et al., 2022; The Royal Society, n.d.; Wirth et al., 2021). Alignment is therefore not a static goal but a dynamic process, continually shaped by evolving societal values (Van de Poel, 2021). Shifts, such as changing value priorities or conceptualizations, are particularly evident in digital contexts, where trustworthiness depends on dynamic user expectations, perceived vulnerabilities, and socio-technical interactions (Schuetz et al., 2025; The Royal Society, n.d.; Van de Poel, 2024). Addressing these challenges requires ongoing alignment of digital innovations with public values to foster trust, safety, and inclusiveness in data-sharing ecosystems (Panagiotopoulos et al., 2019).

Public values, shaped through deliberation and consultation, play a pivotal role in guiding system and policy development (Benington & Moore, 2011; Moore, 1995). Effective governance strategies must reflect citizens' expectations regarding government data use, ensuring that public values are embedded in decision-making

[1] COM/2020/66 final

(Noije & Ridder, 2024). A range of methods exists to address ethical challenges in technological design, often referred to as design for values (Van de Poel, 2024). This concept, originally introduced by Van den Hoven et al. (2015), is used by Van de Poel (2024) as "a term for a number of approaches that aim at integrating values of moral importance in the design of new technological products, services, and systems in a systematic and proactive way" (p. 301). Four noteworthy approaches that explicitly respect values during design, are Value Sensitive Design (VSD), Value-Based Engineering (VBE), Z-Inspection, and the Guidance Ethics Approach (GEA).

The VSD- approach systematically incorporates human values in a specific context into policy and system design through a formative framework of conceptual, empirical, and technical phases (Friedman & Hendry, 2019; Hendry et al., 2021). VSD actively involves stakeholders and balances competing public values, supporting the alignment of digital infrastructures with societal expectations (Friedman et al., 2006; Lee-Geiller & Lee, 2019). The application of VSD can be operationalized through various methods, each tailored to a specific purpose and adapted to the context in which it is applied (Friedman & Hendry, 2019; Van de Poel, 2021, 2024). In contrast, Value-Based Engineering (VBE) offers a standardized approach, made possible through the IEEE 7000™ framework (2022), which provides a structured process for integrating values into system design (Bednar & Spiekermann, 2024). Where VSD emphasizes reflection and dialogue, VBE "specifically assesses negative character effects of IT systems" (Bednar & Spiekermann, 2024, p. 184). Z-Inspection functions as an ethical audit tool during or after the development of AI systems, assessing whether ethical considerations have been adequately addressed in practice (Zicari et al., 2021). Its retrospective character makes it less suitable for shaping values during earlier design stages, but it can add value in the assessment to what extent ethical requirements have been met. The Guidance Ethics Approach (GEA), spans the entire lifecycle from design to implementation and use, fostering societal dialogue to ensure that both the technology and its application context remain aligned with public values (Verbeek & Tijink, 2020). This study adopts a Value Sensitive Design (VSD) approach, as the realization of public values and the context-dependent nature of system and process design require a methodology that can flexibly respond through a diverse set of methods. VSD was also chosen to offer a complementary perspective to predominantly risk-averse practices of VBE, creating space for reflection and a focus

on public values within broader societal contexts. Z-Inspection merely concentrates on assessment and less on design and development, while GEA was deemed less suitable given its intensive stakeholder engagement requirements across the full technology lifecycle. Responsible and trustworthy data sharing in governmental ecosystems requires technologies that can minimise, and secure information transferred between parties. Emerging technologies, including Multi-Party Computing (MPC) in general and Privacy-Enhancing Technologies (PETs) in particular, as well as Distributed Ledger Technology (DLT), provide relevant capabilities to enable the secure exchange and analysis of data while supporting public and economic value creation (ICO, 2023; Wirth et al., 2021). This Office identifies MPC as a privacy-enhancing technology that "provides data minimisation and security by allowing different parties to jointly perform processing on their combined information, without any party needing to share all of its information with each of the other parties" (p. 17). On the other hand, DLT refers to a family of technologies that enable decentralised data management by distributing control among multiple participants. Its most prominent example, blockchain, demonstrates how DLT can foster transparency, traceability, and trust in digital governance (Bons et al., 2020). Together, MPC and DLT are examined as key technologies to support secure, privacy-compliant, and trustworthy data sharing in governmental ecosystems.

2 Research objective

This study investigates how Value Sensitive Design (VSD) can guide the implementation of Multi-Party Computation (MPC) and Distributed Ledger Technology (DLT) to support public values in governmental data-sharing ecosystems. The research question of this research proposal is:

How can Value Sensitive Design be applied in the implementation of Distributed Ledger Technology (DLT) and Multi Party Computation (MPC) to strengthen responsible and trustworthy data sharing in governmental ecosystems?

To systematically address this question, the research is structured around four key sub-questions with mentioned studies, each reflecting a complementary perspective on ethical, legal, technical, societal, and organisational requirements:

1. What are the current mechanisms or practices of Value-Sensitive Design (VSD) in use, and in what ways are Distributed Ledger Technology (DLT), and Multi Party Computation (MPC) being explored within governmental data-sharing initiatives? This sub-question is addressed through a narrative literature review and an exploratory case study, focusing on the application of VSD, DLT, and MPC in an existing public-sector data-sharing project.
2. How can VSD, DLT, and MPC be applied ex-ante in the design of trustworthy and responsible data-sharing processes in governmental ecosystems, addressing value tensions? To explore this, a pilot case study is conducted, examining how the continuous consideration of public values can be embedded in the early stages of technological design and governance for data sharing.
3. How can a normative and methodological framework be designed and constructed to integrate public values into the use of emerging technologies and governance strategies for responsible and trustworthy data sharing? This question is approached through a theoretical literature review combined with the development of a governance framework.
4. How can the proposed framework be operationalized and empirically assessed in terms of responsible and trustworthy data sharing? This final sub-question is examined through empirical validation, using either action research or a multi-case study approach. These methods are intended to evaluate the practical applicability and effectiveness of the framework in real-world settings.

3 Elaboration of the core concepts

To operationalise public values, a Value Sensitive Design (VSD) approach systematically integrates and aligns these values within both policy and system design. Concurrently, the technologies MPC and DLT promise to enhance the governmental technological capacity to mitigate risks, promote transparency, and ensure adherence to legal and ethical standards. The interplay between VSD, DLT, and MPC is vital to enabling responsible, secure, and trustworthy data sharing.

3.1 Public Value Theory

According to Horner and Hutton in Benington and Moore (2011), public value is "an interrelated process of authorization, creation and measurement that determines what the public values through a process of consultation and deliberation" (p. 122). These elements underpin Moore's (1995) *Strategic Triangle*, which balances value creation, legitimacy, and operational capacity. Public services are shaped by democratic processes and must align with societal expectations, especially in areas like data sharing. Public value strategies depend on the level of abstraction, risk and legal certainty (e.g. GDPR[2]), and the actors involved. Effective strategies require alignment between ethical values, legal authorization, and operational feasibility.

3.2 Value Sensitive Design

Value Sensitive Design (VSD) integrates conceptual, empirical, and technical analyses to embed public values into system design (Friedman et al., 2006). VSD balances services, delivery methods, and public expectations. Conceptual analysis identifies values and stakeholders, empirical analysis aligns with real-world contexts, and technical analysis ensures implementation. VSD follows iterative steps such as identifying values, stakeholders, risks, and resolving conflicts. Collaborative governance is central, involving stakeholders in co-creation. This approach supports the development of ethical, trustworthy public systems.

3.3 Emerging technologies

This study focuses on Privacy-Enhancing Technologies (PETs), such as Multi-Party Computation (MPC) and Private Set Intersection (PSI), which enable privacy-preserving data sharing (Gao et al., 2024). MPC allows joint computation without revealing inputs, while PSI identifies overlaps between datasets without exposing other data (Bay et al., 2022). Despite their potential, challenges remain in scalability, accessibility, and governance (Agarwal et al., 2022). Distributed Ledger Technology (DLT), including blockchain, enhances transparency and trust through decentralisation and immutability. Smart contracts can automate agreements but also introduce legal and operational complexities (Zou, 2020).

[2] http://data.europa.eu/eli/reg/2016/679/2016-05-04

3.4 Aligning Ethical, Legal, and Socio-Technical Requirements

Effective data sharing depends on governance, data quality, organisational culture, and expertise. Technologies like MPC and blockchain reduce the need for full trust while maintaining data control. However, trustworthy use requires standards, monitoring, and secure access mechanisms. Technologies are value-laden, meaning ethical considerations must be embedded in both design and implementation. Achieving public value requires dynamic capabilities within organisations to adapt to change. Alignment between values, legal frameworks, and operational capacity is essential and must be continuously maintained.

4 Conceptual model, propositions and expectations

The conceptual model is based on three theoretical perspectives. Firstly, Public Value Theory, introduced by Moore (1995), describes how public values emerge through deliberative processes and influence governance structures, policy decisions, and institutional trust. Secondly, Value Sensitive Design (VSD) provides a structured approach for incorporating ethical, legal, and societal values into technological and policy development. This approach consists of conceptual, empirical, and technical phases that integrate stakeholder values and interests into system and governance design (Friedman et al., 2006). Thirdly, blockchain and Multi-Party Computation (MPC) or Private Set Intersection (PSI) enable secure data processing in multi-actor infrastructures while supporting trust and compliance with legal and ethical frameworks (Bay et al., 2022; ICO, 2023). The conceptual model integrates these perspectives to provide a structured approach to digital governance. Public Value Theory informs the normative foundations of governance, VSD provides a methodology for embedding stakeholder values into policies and systems, and PETs and blockchain function.

4.1 Structure and components of the conceptual model

Public value identification, derived from deliberative and consultative processes, serves as the foundation for governance strategies regulating data sharing within the data ecosystem. VSD translates these values by guiding policy and infrastructure design choices, while emerging technologies in a multi-actor infrastructure, including blockchain and privacy-enhancing techniques, enable responsible and trustworthy data-sharing mechanisms that align with regulatory requirements. The interaction

between the VSD and de emerging technologies influences trust and compliance, shaping responsible and ethical data-sharing practices within a governmental ecosystem. To ensure continuous alignment with public values, the model incorporates a feedback loop where governance mechanisms are evaluated against predefined criteria, stakeholder expectations, and regulatory frameworks.

4.2 Structure and components of the conceptual model

Based on the theoretical framework, the following research propositions are formulated:

P1: The integration of Value Sensitive Design (VSD) and the deployment of blockchain and privacy-enhancing technologies (PETs) in multi-actor infrastructures positively affects responsible and trustworthy data sharing in governmental ecosystems.

P2: The alignment of ethical, legal, technical, societal, and organizational requirements moderates the relationship between VSD with blockchain and PETs and responsible and trustworthy data sharing in governmental ecosystems.

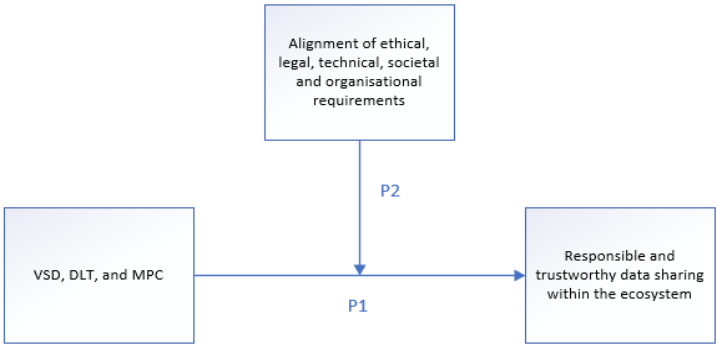

Figure 1: Conceptual model for the research

5 Technical research design and methods

This study adopts a primarily bottom-up approach, complemented by a top-down perspective, to identify relevant norms and values from academic and grey literature, including Dutch governmental sources. These norms inform the identification of

underlying values that guide semi-structured interviews. The research follows a multidisciplinary, iterative methodology involving stakeholders from policy, technology, IT architecture, project management, and legal domains, enabling a comprehensive analysis of safeguarding public values and fundamental rights in governmental data-sharing. Grounded in the GDPR, related legislation, and principles of data protection and good governance, each research question is addressed through tailored methods (Section 5.1), including literature reviews, interviews, surveys, focus groups, and Design Science Research, providing conceptual, normative, and empirical insights.

5.1 Structure and components of the conceptual model

To address Research Questions 1 and 2, this study adopts a qualitative case study approach combining bottom-up and top-down methods. For RQ1, a literature review identifies public values, norms, and stakeholders in governmental data sharing, followed by semi-structured interviews on Value Sensitive Design (VSD) and the use of Distributed Ledger Technology (DLT) and Multi-Party Computation (MPC). For RQ2, qualitative interviews examine how ethical, legal, technical, societal, and organisational factors shape trust and reliability, complemented where possible by quantitative insights. For RQ3, a Design Science Research (DSR) approach guides the development of a normative and methodological framework. A theoretical review and conceptual analysis, combined with earlier findings, synthesise VSD, DLT, MPC, and governance into design principles for embedding public values. Normative grounding draws on Public Value Theory, Value Sensitive Design, and Values Hierarchies, enriched by deontological, consequentialist, and virtue-ethical perspectives. For RQ4, the framework is operationalised and evaluated in two phases: refinement through focus groups and expert panels, followed by a pilot within a governmental organisation (if feasible), assessing effects on privacy, trust, security, and inclusiveness through interviews or surveys.

5.2 Data-analysis

This study employs qualitative and quantitative analysis methods aligned with the nature of the collected data. Thematic analysis (is used to identify patterns ("themes") in qualitative data from interviews, focus groups, expert panels, and open-ended survey responses (Braun et al., 2019). Depending on the research

question and data characteristics, the analysis follows either a deductive approach using predefined codes or an inductive approach in which codes emerge from the data. Quantitative survey data, including Likert-scale and ranking items, are analysed using descriptive statistics to examine trends in the alignment of ethical, legal, technical, societal, and organisational requirements and their perceived influence on trust and reliability in governmental data-sharing processes. This mixed-methods approach supports analytical depth, rigour, and triangulation across data sources.

6 Scientific impact, outputs, and dissemination

This PhD research generates theoretical, empirical, and methodological insights to advance value-driven governance, trust, and ethical ICT design. It delivers a conceptual framework for integrating public values into digital governance, empirical evidence from Dutch governmental case studies, and a methodological approach to value-sensitive data-sharing design.

The research proceeds in phases: two retrospective case studies (2024–2026) examine value identification, design practices, and the role of emerging technologies; a design study (2026–2027) develops a normative and methodological framework; and two empirical studies (2026–2028) operationalise and evaluate this framework in terms of trust, privacy, and institutional legitimacy. The final phase (2028–2029) synthesises the findings into the dissertation. Results will be disseminated through journal publications, conferences, and policy engagement.

7 Relevance for business and educational or training practice

This research advances value-driven governance, emerging technologies, and ethical ICT design, contributing to professional and educational practice. It supports policymakers and organisations in integrating public values into digital governance and promoting responsible, privacy-preserving data sharing. Through policy briefings, workshops, and stakeholder engagement, the study provides actionable insights to enhance trust, compliance, and ethical decision-making. In education, it offers empirical findings, case studies, and methodological insights for curricula and training in public administration, data governance, and ICT ethics. Findings will be disseminated via open-access publications, expert panels, and public engagement, strengthening the link between academic research, policy, and practice.

References

Agahari, W., Ofe, H., & de Reuver, M. (2022). It is not (only) about privacy: How multi-party computation redefines control, trust, and risk in data sharing. *Electronic Markets*, *32*(3), 1577-1602. https://doi.org/10.1007/s12525-022-00572-w

Agarwal, N., Sikka, G., & Awasthi, L. K. (2022). A systematic literature review on web service clustering approaches to enhance service discovery, selection and recommendation. *Computer Science Review*, *45*, 100498. https://doi.org/https://doi.org/10.1016/j.cosrev.2022.100498

Bay, A., Erkin, Z., Hoepman, J. H., Samardjiska, S., & Vos, J. (2022). Practical Multi-Party Private Set Intersection Protocols. *IEEE Transactions on Information Forensics and Security*, *17*, 1-15. https://doi.org/10.1109/TIFS.2021.3118879

Bednar, K., & Spiekermann, S. (2024). The Power of Ethics: Uncovering Technology Risks and Positive Value Potentials in IT Innovation Planning. *Business & Information Systems Engineering*, *66*(2), 181-201. https://doi.org/10.1007/s12599-023-00837-4

Benington, J., & Moore, M. H. (2011). *Public Value. Theory & Practice*. Palgrave Macmillan.

Bons, R. W. H., Versendaal, J., Zavolokina, L., & Shi, W. L. (2020). Potential and limits of Blockchain technology for networked businesses. *Electronic Markets*, *30*(2), 189-194. https://doi.org/10.1007/s12525-020-00421-8

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic Analysis. In P. Liamputtong (Ed.), *Handbook of Research Methods in Health Social Sciences* (pp. 843-860). Springer Singapore. https://doi.org/10.1007/978-981-10-5251-4_103

Durst, S., & Leyer, M. (2022). The influence of institutional conditions on firms' process innovation – evidence from firms based on a multi-country analysis. *The Bottom Line*, *35*(4), 161-184. https://doi.org/10.1108/BL-11-2021-0105

Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design*. The MIT Press.

Friedman, B., Kahn, P., Borning, A., Zhang, P., & Galletta, D. (2006). Value Sensitive Design and Information Systems. In. https://doi.org/10.1007/978-94-007-7844-3_4

Gao, D., Wu, C., Zhang, X., Yao, X., & Yang, Q. (2024, 2024//). Secure Dataset Condensation for Privacy-Preserving and Efficient Vertical Federated Learning. Machine Learning and Knowledge Discovery in Databases. Research Track, Cham.

Hendry, D. G., Friedman, B., & Ballard, S. (2021). Value sensitive design as a formative framework. *Ethics and Information Technology*, *23*(1), 39-44. https://doi.org/10.1007/s10676-021-09579-x

ICO, I. C. s. O. (2023). *Privacy-enhancing technologies (PETs)*. https://ico.org.uk/media/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/privacy-enhancing-technologies-1-0.pdf

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, *37*(3), 101493. https://doi.org/https://doi.org/10.1016/j.giq.2020.101493

Jiang, Z. H., Zhu, H., Zhou, B. B., Lu, C. H., Sun, M. F., Ma, X. J., Fan, X. L., Wang, C., & Chen, L. B. (2023). CrowdPatrol: A Mobile Crowdsensing Framework for Traffic Violation Hotspot Patrolling. *Ieee Transactions on Mobile Computing*, *22*(3), 1401-1416. https://doi.org/10.1109/tmc.2021.3110592

Lee-Geiller, S., & Lee, T. (2019). Using government websites to enhance democratic E-governance: A conceptual model for evaluation. *Government Information Quarterly*, *36*(2), 208-225. https://doi.org/https://doi.org/10.1016/j.giq.2019.01.003

Moore, M. H. (1995). *Creating Public Value*. Harvard University Press.

Noije, L., van, & Ridder, J., den. (2024). *Burgers over democratie en rechtsstaat Kennisnotitie voor de Staatscommissie rechtsstaat*. https://www.scp.nl/binaries/scp/documenten/publicaties/2024/01/22/kennisnotitie-scp-voor-de-staatscommissie-rechtsstaat/Kennisnotitie+Staatscommissie+rechtsstaat.pdf

Panagiotopoulos, P., Klievink, B., & Cordella, A. (2019). Public value creation in digital government. *Government Information Quarterly*, *36*(4), 101421. https://doi.org/https://doi.org/10.1016/j.giq.2019.101421

Rukanova, B., van Engelenburg, S., Ubacht, J., Tan, Y.-H., Geurts, M., Sies, M., Molenhuis, M., Slegt, M., & van Dijk, D. (2023). Public value creation through voluntary business to government information sharing enabled by digital infrastructure innovations: a framework for analysis. *Government Information Quarterly*, *40*(2), 101786. https://doi.org/10.1016/j.giq.2022.101786

Schuetz, S., Kuai, L., Lacity, M. C., & Steelman, Z. A qualitative systematic review of trust in technology. *Journal of Information Technology*, *0*(0), 02683962241254392. https://doi.org/10.1177/02683962241254392

Schuetz, S., Kuai, L., Lacity, M. C., & Steelman, Z. (2025). A qualitative systematic review of trust in technology. *Journal of Information Technology*, *40*(1), 55-76. https://doi.org/10.1177/02683962241254392

Society, T. R. (n.d.). *Privacy Enhancing Technologies: Market Readiness, Enabling and Limiting Factors in the UK public sector*. https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/privacy-enhancing-technologies-market-readiness-enabling-and-limiting-factors.pdf

Susha, I., Rukanova, B., Zuiderwijk, A., Gil-Garcia, J. R., & Gasco Hernandez, M. (2023). Achieving voluntary data sharing in cross sector partnerships: Three partnership models. *Information and Organization*, *33*(1), 100448. https://doi.org/https://doi.org/10.1016/j.infoandorg.2023.100448

Treiber, A., Müllmann, D., Schneider, T., & Döhmann, I. S. g. (2022). *Data Protection Law and Multi-Party Computation: Applications to Information Exchange between Law Enforcement Agencies* Proceedings of the 21st Workshop on Privacy in the Electronic Society, Los Angeles, CA, USA. https://doi-org.ezproxy.elib11.ub.unimaas.nl/10.1145/3559613.3563192

Uitvoering, S. v. d. (2023). *The state of our public service delivery 2022*. https://staatvandeuitvoering.nl/app/uploads/2023/06/ICTU-SvdU-2022-EN-interactive-v03.pdf

Van de Poel, I. (2021). Design for value change. *Ethics and Information Technology*, *23*(1), 27-31. https://doi.org/10.1007/s10676-018-9461-9

Van de Poel, I. (2024). Design for Value. In A. Grunwald (Ed.), *Handbook of Technology Assessment* (pp. 301-309). Edward Elgar Publishing https://doi.org/https://doi.org/10.4337/9781035310685

Van den Hoven, J. E., Vermaas, P. E., & Van de Poel, I. (2015). Springer. https://doi.org/10.1007/978-94-007-6970-0

Verbeek, P.-P., & Tijink, D. (2020). *Guidance ethics approach*. ECP | Platform voor de InformatieSamenleving. https://ecp.nl/wp-content/uploads/2020/11/Guidance-ethics-approach.pdf

Vermeer, M. (2023). *Eigen cocon versus leentjebuur* [Interview]. Binnenlands Bestuur.

Wirth, F. N., Meurers, T., Johns, M., & Prasser, F. (2021). Privacy-preserving data sharing infrastructures for medical research: systematization and comparison. *BMC Medical Informatics and Decision Making*, *21*(1), 242. https://doi.org/10.1186/s12911-021-01602-x

Zicari, R. V., Brodersen, J., Brusseau, J., Düdder, B., Eichhorn, T., Ivanov, T., Kararigas, G., Kringen, P., McCullough, M., Möslein, F., Mushtaq, N., Roig, G., Stürtz, N., Tolle, K., Tithi, J. J., Halem, I. v., & Westerlund, M. (2021). Z-Inspection®: A Process to Assess Trustworthy AI. *IEEE Transactions on Technology and Society*, *2*(2), 83-97. https://doi.org/10.1109/TTS.2021.3066209

Zou, M. (2020). Code, and Other Laws of Blockchain [Blockchain and the Law: The Rule of Code, Primavera De Filippi, Aaron Wright]. *Oxford Journal of Legal Studies*, *40*(3), 645-665. https://www.jstor.org/stable/48595050

DOCTORAL CONSORTIUM

BUSINESS SYSTEM LONG-TERM VIABILITY ACROSS LIFE-CYCLE PHASES

SARA KREMSAR, EVA KRHAČ ANDRAŠEC, TOMAŽ KERN
University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
sara.kremsar@student.um.si; eva.krhac1@um.si; tomaz.kern@um.si

This research addresses how business systems can achieve short-term success while sustaining long-term viability. Although the organizational literature already examines performance, resilience, and life-cycle dynamics, it lacks an integrated, empirically grounded framework that systematically connects internal and external factors to long-term viability across developmental phases. The research is conceptualized as a proposed Design Science Research study and will adopt a mixed-methods approach, combining qualitative and quantitative methods to identify, classify, and analyse factors influencing both short-term success and long-term viability. Using a five-phase life-cycle typology as an analytical framework, the research is expected to develop a reference model and accompanying analytical factor cards as its primary artefact. The original scientific contribution of the research is expected to lie in distinguishing short-term success from long-term viability, classifying factors across life-cycle phases and integrating them into a unified model. The results are expected to advance organizational theory and provide a structured analytical foundation for examining long-term viability in complex environments.

DOI
https://doi.org/10.18690/um.fov.4.2026.76

ISBN
978-961-299-152-4

Keywords:
long-term viability, business system, life-cycle, internal factors, external factors

University of Maribor Press

1 Introduction

The central issue addressed in this research is how to design a business system that ensures success at a given point in time while also maintaining preparedness for future challenges and long-term viability. From an organizational perspective, this challenge requires reflection on how to design processes, organizational structure, and culture so that the business system performs effectively today while remaining resilient and adaptable in the future (Burns & Stalker, 1961; Chandler, 1962; Mintzberg, 1979; Duchek, 2019; Hillmann et al., 2022; Florez-Jimenez et al., 2024). This research, therefore, focuses on identifying and classifying the internal and external factors that enable long-term viability, and on explaining how their relevance changes across different phases of a business system's life-cycle.

The research is based on the thesis that long-term viability is statistically dependent on the interconnected operation of internal and external factors, and that the significance of these factors varies across life-cycle phases (Napolitano et al., 2015; Williams et al., 2019; Miceli et al., 2021; Padovano & Ivanov, 2025). In line with this, the research distinguishes between short-term success, a results-oriented dimension, and long-term viability, a capability-oriented, systemic dimension. Short-term success may support long-term viability, but it does not guarantee it. This distinction is critical because many business systems demonstrate strong short-term performance while gradually losing their capacity to remain functional, legitimate, and value-creating (Blasini et al., 2013; Hitt et al., 2015; Ortiz-de-Mandojana & Bansal, 2015; Kaplan, 2012; Kayani et al., 2023; Sun et al., 2024).

Every business system is inherently vulnerable to decline or termination. However, through deliberate organizational approaches and strategic interventions, this process can be slowed, redirected, or in some cases prevented. Unlike biological organisms, business systems can persist for extended periods when appropriately designed and institutionally supported (Collins & Porras, 1994; Elkington, 1997; Barney & Hesterly, 2015; Georgescu et al., 2024; Dickson, 2025).

Although long-term viability is often assessed through financial or market indicators, this research proceeds on the assumption that it is not primarily an economic phenomenon but an organizational one. It fundamentally depends on how business systems develop and renew appropriate structures, decision mechanisms, routines,

culture, and adaptive capacity in a changing environment (Makkonen et al., 2014; Koronis & Ponis, 2018; Duchek, 2019; Miceli et al., 2021; Florez-Jimenez et al., 2024). Consequently, long-term viability is treated as an emergent systemic property, arising from the interaction of multiple organizational and environmental domains rather than from a single isolated factor (Barasa et al., 2018; Napolitano et al., 2015; Williams et al., 2019; Miceli et al., 2021).

To address this problem, the research develops a research artefact in the form of a phase-sensitive reference model of factors and analytical factor cards. The factor cards serve as an operational instrument for structured analysis and comparison of factor influence across contexts and phases.

The research is guided by research questions that focus on identifying factors of long-term viability, distinguishing them from factors of short-term success, examining their overlaps and tensions, and analysing how their relevance changes across life-cycle phases. By integrating these components into a coherent analytical framework, the research aims to strengthen theoretical clarity and provide a transferable basis for empirical testing and practical diagnosis.

2 Problem definition

The long-term viability of business systems is a crucial yet insufficiently integrated field within organizational studies, particularly amid increasing environmental volatility, technological disruption, and institutional instability. Business systems today operate in a world characterized by rapid digitalization, global interdependence, regulatory shifts, and recurrent crises. These dynamics continuously challenge their stability, adaptive capacity, and long-term continuity.

Most of the existing research predominantly focuses on factors of short-term success, such as profitability, innovation output, efficiency, competitive positioning, and market share (Blasini et al., 2013; Kaplan, 2012; Hitt et al., 2015; Sun et al., 2024; Hillmann et al., 2022; Kayani et al., 2023). While these contributions provide valuable insights into performance dynamics, comparatively less attention has been devoted to understanding the systemic conditions that enable business systems to remain functional and legitimate over extended periods.

Although long-term viability is frequently evaluated through financial indicators or longevity measures, this research proceeds on the assumption that it is not primarily an economic outcome but an organizational phenomenon. It depends on the capability of a business system to configure and reconfigure its structures, processes, governance mechanisms, and cultural patterns in alignment with changing environmental conditions (Makkonen et al., 2013; Koronis and Ponis, 2018; Duchek, 2019; Miceli et al., 2021; Florez-Jimenez et al., 2024). In this sense, long-term viability is conceptualized as an emergent systemic property arising from the interaction of internal and external factors.

Foundational organizational theories emphasize the alignment between strategy and structure (Chandler, 1962), the importance of organizational design (Mintzberg, 1979), and the need to adapt structural configurations to environmental uncertainty (Burns and Stalker, 1961). Across different streams of research, various factors of organizational stability and long-term functioning have been identified. Innovation capability has been consistently identified as a key survival factor in dynamic industries (Christensen, 1997). Vision and organizational culture (Collins and Porras, 1994), sustainability-oriented approaches (Elkington, 1997), and organizational health (McKinsey, 2024) further contribute to stability and resilience. However, these contributions remain fragmented across conceptual domains and are rarely integrated within a unified, phase-sensitive analytical framework.

Empirical evidence demonstrates that short-term success does not necessarily translate into long-term viability. The case of Eastman Kodak illustrates how strong historical performance can coexist with structural rigidity and strategic misalignment in the face of disruptive technological change (Lucas et al., 2009; Mui, 2012). Similarly, Nokia's rapid decline during the smartphone transition highlights the consequences of delayed strategic adaptation despite prior market dominance (Vuorti and Huy, 2016; Kornelakis and Petrakaki, 2024). These cases reveal a structural gap between operational success and long-term systemic adaptability.

At a systemic level, this dynamic is further illuminated by the Gartner Hype Cycle, which describes fluctuating expectations surrounding emerging technologies. Organizations often make strategic decisions under conditions of uncertainty, misaligned timing, or distorted expectations (Shi and Herniman, 2022; Mora-López et al., 2025). Failures, therefore, frequently result not from lack of competence or

resources but from temporal misalignment between environmental shifts and organizational response capacity.

In contrast, the case of Pliva and the development of azithromycin (Sumamed) illustrates a trajectory of sustained long-term viability through the integration of technological innovation, intellectual property protection, regulatory positioning, and international licensing (Tomišić and Inić, 2024). This example demonstrates that long-term viability may depend not only on innovation itself but on the institutional and strategic mechanisms that enable prolonged value appropriation.

Despite the growing emphasis on resilience, sustainability, and adaptability in management discourse, the academic literature still lacks a comprehensive, empirically grounded framework that systematically integrates internal and external factors to ensure long-term viability across the different phases of the business system life-cycle. Existing research remains dispersed across performance studies, resilience research, life-cycle models, and strategic management theories, limiting cumulative knowledge and comparative analysis (Napolitano et al., 2015; Williams et al., 2019; Miceli et al., 2021; Padovano and Ivanov, 2025).

The central research problem, therefore, arises from the need to systematically identify and classify the factors that enable the long-term viability of business systems, distinguish them from factors of short-term success, and examine how their relative importance changes across life-cycle phases. In particular, the research seeks to address the absence of an integrative, phase-sensitive analytical model capable of capturing the interconnected operation of internal and external factors of long-term viability.

By addressing this gap, the research aims to move beyond fragmented explanations and to provide a structured conceptual and empirical foundation for understanding how business systems sustain themselves under varying environmental and developmental conditions.

3 Methodology

The presented methodological framework represents the design and preparation of the research, while the empirical phase has not yet been conducted. The research is designed as a proposed mixed-methods study that will combine qualitative and

quantitative approaches to provide a comprehensive treatment of the research problem and to integrate theoretical foundations with empirical evidence. The methodological framework will follow the principles of Design Science Research, within which a research artefact will be developed and analytically evaluated through iterative interaction between conceptual development and empirical inquiry. The central objective of the methodological design will be to identify, classify, and analyse internal and external factors influencing the long-term viability of business systems and, on this basis, to develop and empirically ground a phase-sensitive reference model and accompanying analytical factor cards.

The research design will connect theoretical synthesis, empirical data collection, and analytical integration within a structured comparative framework. The research will be cross-sectional and will not follow individual business systems longitudinally. Instead, it will analyse business systems at different life-cycle phases at a single point in time. The life-cycle will be used as an analytical typology. Its purpose is to enable structured comparison of factor configurations across developmental contexts.

The theoretical phase will consist of a systematic review of scientific literature covering organizational theory, business performance, internal and external factors, life-cycle models, organizational resilience, and long-term viability. Sources will be drawn from ProQuest, ScienceDirect, Scopus, and Web of Science. A comparative analysis of existing models will provide the conceptual foundation for identifying candidate factors and for constructing the initial version of the reference model and the analytical factor cards. The literature review will serve as a structural basis for artefact development.

The empirical phase will focus primarily on service-based business systems and will employ purposive sampling. The objective will not be statistical representativeness in the classical sense, but the identification of information-rich cases that allow the examination of complex, context-dependent relationships among internal and external factors. The sample will include business systems of varying sizes, ages, and life-cycle positions. Service systems will be selected due to their structural reliance on intangible resources such as knowledge, culture, relational capital, digital integration, and adaptive capacity, which makes them analytically suitable for examining long-term viability dynamics. Sampling will be conducted through professional and research networks as well as publicly available business databases.

Heterogeneity of the sample will be treated as an analytical advantage, enabling comparative analysis across different configurations. If insufficient cases are identified within a specific life-cycle phase, the sample will be iteratively supplemented to ensure phase comparability.

Data collection will consist of two complementary instruments. First, a structured survey questionnaire will collect quantitative data regarding the perceived importance and influence of internal and external factors. Measurement scales will be adapted from established constructs in organizational resilience, culture, and strategic management research. The questionnaire will be distributed via an online platform, and a minimum of 100 valid responses will be targeted to ensure basic statistical robustness for quantitative analysis. Second, semi-structured interviews will be conducted with representatives of at least 10 business systems to capture in-depth perspectives on factor dynamics, strategic decisions, and contextual influences. The number of interviews will be guided by the principle of theoretical saturation.

Quantitative data will be analysed using descriptive statistical methods in SPSS or JASP. Where appropriate, comparative analyses will be conducted across business systems that differ in size, age, or life-cycle phase. Qualitative data from interviews will be analysed using thematic analysis supported by Atlas.ti, enabling identification of recurring patterns, contextual mechanisms, and variations in factor influence. Integration of qualitative and quantitative findings supports triangulation and will strengthens the internal consistency of the analytical conclusions.

The reference model and analytical factor cards will constitute the central research artefact. Artefact development will follow an iterative refinement process. It will begin with conceptual construction based on the literature. It will then proceed with empirical calibration using collected data and will conclude with analytical integration. Evaluation of the artefact will be conducted as an analytically grounded assessment of conceptual coherence, phase sensitivity, internal consistency, explanatory capacity, and practical applicability. The evaluation will not take the form of experimental testing, but rather structured analytical validation supported by empirical findings.

The research is limited by purposive sampling, a cross-sectional design, and contextual variability of factors. The proposed model remains conceptual and requires further validation across additional sectors and institutional environments.

4 Expected results

The original scientific contribution of this research is expected to lie in developing a phase-sensitive reference model that systematically integrates internal and external factors influencing the long-term viability of business systems. The model advances existing theoretical approaches by moving beyond fragmented treatments of performance, resilience, sustainability, and life-cycle models in isolation. Instead, it provides a structured analytical framework that connects these domains within a unified conceptual architecture.

The research is expected to contribute to clearer differentiation between short-term success and long-term viability. While short-term success is reflected in results-oriented performance indicators, long-term viability is a systemic, capability-based phenomenon. The research is expected to demonstrate how these two dimensions overlap, diverge, and in certain configurations condition one another. In doing so, it is expected to address the persistent ambiguity in existing literature, where longevity is often equated with financial performance or survival duration without deeper structural analysis. This differentiation is expected to clarify under which conditions performance-enhancing mechanisms contribute to long-term resilience and under which conditions they may undermine systemic adaptability.

Another key expected result is the classification of the identified factors by their relative importance across different life-cycle phases. By applying the five-phase analytical typology, the research is expected to demonstrate how the relevance and interactions of factors change across developmental contexts. This phase-sensitive perspective represents a significant theoretical extension beyond static models of performance or resilience.

The final research artefact is expected to be a reference model of long-term viability factors, providing a structured representation of their interrelationships across life-cycle phases and long-term viability. The analytical factor cards will operationalize the model by enabling systematic monitoring, classification, and comparative

evaluation of factor influence across contexts. Together, they will constitute a theoretically and empirically informed analytical instrument.

The originality of the contribution is grounded in evidence of a research gap in existing literature. Searches in the Web of Science database reveal extensive research on individual concepts such as "life-cycle," "key success factors," and "long-term viability," yet minimal integration among them. Combined searches yield only a very limited number of publications, and closer examination indicates that none address the interconnected operation of internal and external factors influencing long-term viability across life-cycle phases.

Beyond conceptual advancement, the research is expected to generate empirically grounded insights into how business systems maintain stability while remaining adaptable amid digital transformation, globalization, and recurring systemic disruptions. The findings are expected to contribute to organizational sciences, strategic management, systems approaches, and organizational resilience by offering a structured explanation of how long-term viability emerges from interconnected factor dynamics rather than isolated strategic decisions.

5 Future development

Although the research is expected to develop an empirically grounded reference model of long-term viability, it also opens several avenues for further theoretical refinement, empirical expansion, and methodological advancement.

First, future research may extend the cross-sectional design to a longitudinal design, which would allow direct observation of how factor configurations evolve and interact as systems transition between phases. Such an approach would provide deeper insight into causal mechanisms and temporal sequences underlying long-term viability.

Second, additional empirical validation across diverse institutional and sectoral contexts would strengthen the robustness and generalizability of the proposed model. The current research primarily focuses on service-based business systems. Replication in other sectors could reveal sector-specific configurations of factors and refine the model's contextual sensitivity. Comparative international research may

further illuminate the moderating role of regulatory, cultural, and macroeconomic conditions.

Third, the structural relationships embedded in the reference model may be tested using advanced quantitative techniques. Methods such as structural equation modeling, multilevel modeling, or system dynamics simulation could examine indirect, mediating, or feedback effects among internal and external factors. Such methodological expansion would enhance explanatory depth and allow more precise operationalization of constructs such as resilience, adaptability, and strategic stability.

Fourth, future research may further develop operational and measurement approaches for analysing long-term viability. While this research establishes the conceptual distinction between short-term success and long-term viability, additional work may refine empirical indicators capturing long-term viability as a systemic capability. Such developments would strengthen cumulative theory building and facilitate cross-research comparability.

Fifth, the transformation phase of a business system's life-cycle represents a particularly fertile area for deeper investigation. Future research may analyse triggers of transformation, decision-making dynamics during renewal processes, and the conditions under which transformation strengthens rather than destabilizes long-term viability. Further examination of specific transformation contexts, including digital transformation and innovation ecosystems, could enhance the model's contemporary relevance.

Sixth, the practical application of the research artefact may evolve through the development of a digital diagnostic tool based on the reference model and analytical factor cards. Such a tool could support structured self-assessment, benchmarking, and scenario analysis within business systems. Iterative feedback from practice would simultaneously inform further theoretical refinement.

Overall, future development of this research should aim to deepen temporal, contextual, and structural understanding of long-term viability while progressively strengthening empirical validation of the model. Through iterative refinement and cross-contextual testing, the proposed framework may evolve into a transferable

analytical instrument for studying and supporting the long-term viability of business systems in complex and uncertain environments.

References

Barasa, E., Mbau, R., & Gilson, L. (2018). What Is Resilience and How Can It Be Nurtured? A Systematic Review of Empirical Literature on Organizational Resilience. International Journal of Health Policy and Management, 7, 491 - 503. https://doi.org/10.15171/ijhpm.2018.06.

Barney, J. B., & Hesterly, W. S. (2015). Strategic management and competitive advantage: Concepts and cases (5th ed.). Pearson Education Limited. https://students.aiu.edu/submissions/profiles/resources/onlineBook/P7p5f5_Strategic%20Management%20and%20Competitive%20Advantage.pdf

Blasini, J., & Leist, S. (2013). Success factors in process performance management. Business Process Management Journal, 19(3), 477–495. https://doi.org/10.1108/14637151311319914

Burns, T., & Stalker, G. M. (1961). The management of innovation. Tavistock.

Chandler, A. D. (1962). Strategy and structure: Chapters in the history of the industrial enterprise. MIT Press.

Christensen, C. M. (1997). The innovator's dilemma: When new technologies cause great firms to fail. Harvard Business Review Press.

Collins, J. C., & Porras, J. I. (1994). Built to last: Successful habits of visionary companies. HarperCollins.

Dickson, R. K. (2025). Organizational resilience as the springboard for organizational success in a turbulent business environment. European Journal of Management, Economics and Business, 2(2), 3–24. https://doi.org/10.59324/ejmeb.2025.2(2).01

Duchek, S. (2019). Organizational resilience: A capability-based conceptualization. Business Research, 13(1), 215–246. https://doi.org/10.1007/s40685-019-0085-7

Elkington, J. (1997). Cannibals with forks: The triple bottom line of 21st century business. Capstone.

Florez-Jimenez, M. P., Lleo, A., Danvila-del-Valle, I., & Sánchez-Marín, G. (2024). Corporate sustainability, organizational resilience and corporate purpose: A triple concept for achieving long-term prosperity. Management Decision, 62(7), 2189–2213. https://doi.org/10.1108/MD-06-2023-0938

Georgescu, I., Bocean, C. G., Vărzaru, A. A., Rotea, C. C., Mangra, M. G., & Mangra, G. I. (2024). Enhancing Organizational Resilience: The Transformative Influence of Strategic Human Resource Management Practices and Organizational Culture. Sustainability, 16(10), 4315. https://doi.org/10.3390/su16104315

Hillmann, J., Bergmann, A., & Guenther, E. (2022). Balancing efficiency and resilience: The risk of neglecting organizational adaptability. Strategic Management Review, 33(1), 45–66.

Hitt, M. A., Ireland, R. D., & Sirmon, D. G. (2015). Strategic entrepreneurship: Creating value for individuals, organizations, and society. Academy of Management Perspectives, 29(2), 239–255.

Kaplan, R. S. (2012). The balanced scorecard: Comments on balanced scorecard commentaries. Journal of Accounting & Organizational Change, 8(4), 539–545. https://doi.org/10.1108/18325911211273527

Kayani, U. N., Gan, C., Rabbani, M. R., & Trichilli, Y. (2023). Is short-term firm performance an indicator of a sustainable financial performance? Empirical evidence. Studies in Economics and Finance, 41(3), 619–637. https://doi.org/10.1108/SEF-03-2023-0136

Kornelakis, A., & Petrakaki, D. (2024). Technological innovation, industry platforms or financialization? A comparative institutional perspective on Nokia, Apple, and Samsung. Business History, 67, 1983 - 2008. https://doi.org/10.1080/00076791.2024.2377688.

Koronis, E., & Ponis, S. (2018). Better than before: the resilient organization in crisis mode. Journal of Business Strategy, 39, 32-42. https://doi.org/10.1108/jbs-10-2016-0124.

Lucas, H. C., & Goh, J. M. (2009). Disruptive technology: How Kodak missed the digital photography revolution. The Journal of Strategic Information Systems, 18(1), 46–55. https://doi.org/10.1016/j.jsis.2009.01.002

Makkonen, H., Pohjola, M., Olkkonen, R., & Koponen, A. (2013). Dynamic capabilities and firm performance in a financial crisis. Journal of Business Research, 67, 2707-2719. https://doi.org/10.1016/j.jbusres.2013.03.020.

McKinsey & Company. (2023, October 10). Organizational health is still the key to long-term performance. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/organizational-health-is-still-the-key-to-long-term-performance

McKinsey & Company. (2024). Organizational health is (still) the key to long-term performance. McKinsey & Company.

Miceli, A., Hagen, B., Riccardi, M., Sotti, F., & Settembre-Blundo, D. (2021). Thriving, Not Just Surviving in Changing Times: How Sustainability, Agility and Digitalization Intertwine with Organizational Resilience. Sustainability. https://doi.org/10.3390/su13042052.

Mintzberg, H. (1979). The structuring of organizations: A synthesis of the research. Prentice-Hall.

Mora-López, J., López-López, D., & Rivera-Hernaez, O. (2025). Unveiling the Generative AI boom: what hype metrics reveal for digital business and E-commerce. Electronic Commerce Research. https://doi.org/10.1007/s10660-025-09984-0.

Mui, C. (2012, January 18). How Kodak failed. Forbes. https://www.forbes.com/sites/chunkamui/2012/01/18/how-kodak-failed/

Napolitano, M., Marino, V., & Ojala, J. (2015). In search of an integrated framework of business longevity. Business History, 57, 955–969. https://doi.org/10.1080/00076791.2014.993613

Ortiz-de-Mandojana, N., & Bansal, P. (2016). The long-term benefits of organizational resilience through sustainable business practices. Strategic Management Journal, 37(9), 1615–1631. https://doi.org/10.1002/smj.2410

Padovano, A., & Ivanov, D. (2025). Towards resilient and viable supply chains: a multidimensional model and empirical analysis. International Journal of Production Research, 63, 6252 - 6290. https://doi.org/10.1080/00207543.2025.2470350.

Shi, Y., & Herniman, J. (2022). The role of expectation in innovation evolution: Exploring hype cycles. Technovation. https://doi.org/10.1016/j.technovation.2022.102459.

Sun, Y., Chen, Y., & Wu, W. (2024). Short-term success and long-term failure? The case of GDP growth targets and corporate innovation. Pacific-Basin Finance Journal, 84, 102308. https://doi.org/10.1016/j.pacfin.2024.102308

Tomišić, B., & Inić, S. (2024). Pharmaceutical company PLIVA: from Nobel laureate Vladimir Prelog to global blockbuster - antibiotic azithromycin.. Die Pharmazie, 79 10, 246-257 . https://doi.org/10.1691/ph.2024.4047.

Vuori, T., & Huy, Q. (2016). Distributed Attention and Shared Emotions in the Innovation Process. Administrative Science Quarterly, 61, 51 - 9. https://doi.org/10.1177/0001839215606951.

Williams, A., Whiteman, G., & Kennedy, S. (2019). Cross-scale systemic resilience: Implications for organization studies. Business & Society, 60, 95–124. https://doi.org/10.1177/0007650319825870

About the authors

Sara Kremsar, M.Sc., is the founder of Optimod, a company supporting fast-growing organizations in building clarity, structure, and alignment. She holds a master's degree in Business Systems Engineering from the Faculty of Organizational Sciences, specializing in organizational structures, processes, and system optimization. Since 2016, she has worked with companies across sectors to improve efficiency and support sustainable growth. Through initiatives such as the Improving Entrepreneurial Journey Initiative (IEJI) and the Founders' Off-the-Grid Retreat, Optimod promotes the exchange of practical experience and academic insight. Sara Kremsar's work bridges theory and practice in organizational

science and entrepreneurship. Her vision is to create resilient, human-centered ecosystems where collaboration, learning, and innovation help individuals and organizations reach their full potential.

Tomaž Kern, Ph.D., is a Full Professor of Organization and Informatics at the University of Maribor, Faculty of Organizational Sciences. He holds degrees in mechanical engineering, informatics, and organizational sciences. Since 1990, he has been actively involved in teaching and research at the Faculty of Organizational Sciences, where he lectures in business process management and multiproject management. He is the head of the Enterprise Engineering study program and leads the Laboratory of Enterprise Engineering (LIPPS). During his academic career, he has held several leadership positions at the university, including Vice-Rector, Vice-Dean, and Head of Institute, and currently serves as a member of the University of Maribor Senate. Professor Kern has authored or co-authored more than 500 scientific publications and has led numerous research and applied projects.

Eva Krhač Andrašec, Ph.D, is an Assistant Professor in the Department of Enterprise Engineering. She received her master's degree in organization and management of business and working systems at the University of Maribor, Faculty of Organizational Sciences, in 2016. In 2023, she successfully defended her doctoral dissertation on business process improvement methods and techniques and their impact on the efficiency of organizational systems. Her main research interests are business process management and quality management. She actively participates in international scientific conferences and research projects and is the co-author of several scientific articles on business process management.

DOCTORAL CONSORTIUM

GRAPH NEURAL NETWORKS AND DEEP REINFORCEMENT LEARNING FOR ADAPTIVE WAREHOUSE OPTIMIZATION

NEJC ČELIK, ANDREJ ŠKRABA

University of Maribor, Faculty of Organizational Sciences, Cybernetics & Decision Support Systems Laboratory, Kranj, Slovenia
nejc.celik1@um.si, andrej.skraba@um.si

Warehouse operations play a critical role in modern supply chains, with order picking representing one of the most resource-intensive processes in both manual and automated warehouses. Improving the efficiency of picking operations has therefore been a long-standing research focus in logistics and operations research. Traditionally, order-picking routing problems have been modeled as variants of the Traveling Salesman Problem (TSP) and addressed using heuristics and metaheuristics. Recent advances in machine learning, particularly deep reinforcement learning (DRL), offer promising opportunities for developing adaptive decision-making systems capable of operating in dynamic environments. At the same time, warehouse systems naturally exhibit graph structures that can be used by graph neural networks (GNNs). This doctoral research proposes an integrated framework combining GNNs and DRL to optimize warehouse processes.

DOI
https://doi.org/10.18690/um.fov.4.2026.77

ISBN
978-961-299-152-4

Keywords:
order picking, deep reinforcement learning, graph neural networks, warehouse routing optimization, joint optimization

University of Maribor Press

1 Introduction

Warehouse efficiency has become increasingly important with the growth of e-commerce and the rising demand for rapid order fulfilment. Among warehouse operations, order picking typically accounts for the largest share of operational costs and labor time. As a result, improving the efficiency of order-picking processes has been a major topic in logistics and operations research. (Masae et al., 2020)

The order-picking routing problem was initially formulated as a variant of the Traveling Salesman Problem (TSP) (Ratliff & Rosenthal, 1983), focusing solely on minimizing travel distance between storage locations. More recent approaches extend this formulation to jointly address additional operational problems such as order batching, congestion management, and coordination among multiple pickers (Masae et al., 2020). A wide range of heuristics and optimization algorithms have been developed for this problem (Masae et al., 2020) and similar problems like vehicle routing problem (Psaraftis et al., 2016).

Advances in machine learning have enabled novel approaches to solving complex combinatorial optimization problems. Among these, deep reinforcement learning (DRL) provides a framework in which agents learn decision policies through interaction with an environment, making it particularly suitable for sequential decision-making problems such as routing (Mazyavkina et al., 2021). At the same time, warehouse systems can naturally be represented as graphs, where nodes correspond to entities and edges represent relationships between them. Graph neural networks (GNNs) have emerged as powerful models for learning from such graph-structured data. By aggregating information from neighboring nodes, GNNs can capture structural relationships within complex networks (Ju et al., 2024; Wu et al., 2021). In the context of warehouse systems, entities such as storage locations, pickers, items and orders can be represented as nodes and connected with edges in a graph, enabling the application of graph-based machine learning methods.

Despite these promising developments, the integration of GNNs and reinforcement learning in warehouse systems is still limited in existing research. This research therefore aims to investigate how GNNs combined with reinforcement learning can be used to develop adaptive decision-support systems for warehouse operations.

2 Research Problem

The central research problem of this doctoral work can therefore be summarized as follows:

How can warehouse systems be modeled as graphs and optimized using machine learning methods that combine graph neural networks and reinforcement learning?

The primary objective of this research is to develop a methodological framework for modeling warehouse systems as graphs and using this representation within a reinforcement learning framework to optimize order-picking operations.

More specifically, the research seeks to answer the following research questions:

RQ1: To what extent does the integration of graph neural networks and deep reinforcement learning improve routing decisions compared to traditional heuristic methods?

RQ2: To what extent can learned policies generalize or adapt to new problems and configurations in warehouse environments?

Addressing these questions will help determine whether graph-based machine learning approaches can provide a viable foundation for adaptive warehouse optimization systems.

3 Research Methodology

The proposed research will combine graph-based modeling, simulation, and machine learning methods to investigate adaptive optimization of warehouse processes.

First, warehouse layouts will be represented as graphs derived from relational data structures commonly used in warehouse management systems (WMS). In this representation, nodes will correspond to storage locations, items, orders, or other relevant entities, while edges will represent connections between them.

Second, graph neural networks will be used to process the graph representation of the warehouse and generate node embeddings that capture structural information about the environment. These embeddings will serve as inputs for the reinforcement learning agent, enabling the model to make decisions based on relational information within the graph.

Third, the order-picking problem, in combination with other warehouse operational problems, will be formulated as a sequential decision-making task suitable for reinforcement learning. In this framework, the agent represents a warehouse worker or robotic system performing warehouse operations. The objective of the reinforcement learning policy is to minimize the cost associated with operational actions while maximizing overall warehouse throughput.

Experiments will be conducted using simulated warehouse environments of varying sizes and configurations. The performance of the proposed approach will be evaluated by comparing learned policies with classical heuristics. Additional analysis will investigate the interpretability of learned embeddings and the ability of models to generalize across different warehouse layouts.

4 Expected Contributions

This research is expected to contribute to both logistics research and machine learning applications in operations management.

The research will propose a holistic GNN–DRL methodology for adaptive warehouse optimization. Additionally, this research will investigate the interpretability of learned graph representations and their potential use for decision support in warehouse management systems. Building on our work (Čelik & Škraba, 2026), this research will extend the application of GNN–DRL to more complex and realistic warehouse scenarios.

More broadly, the research aims to contribute to the development of adaptive, data-driven decision-support systems for internal logistics, bridging the gap between traditional optimization methods and modern machine learning approaches.

Acknowledgment

This research was funded by Slovenian Research and Innovation Agency (research core funding No. P5-0018), and the Ministry of Higher Education, Science, and Innovation of the Republic of Slovenia as part of the Next Generation EU National Recovery and Resilience Plan (Grant No. 3330-22-3515; NOO No: C3330-22-953012),

References

Čelik, N., & Škraba, A. (2026). Graph Neural Networks and Deep Reinforcement Learning for Warehouse Order Picking and Representation Learning. *Preprints*. https://doi.org/10.20944/preprints202602.0896.v1

Ju, W., Fang, Z., Gu, Y., Liu, Z., Long, Q., Qiao, Z., Qin, Y., Shen, J., Sun, F., Xiao, Z., Yang, J., Yuan, J., Zhao, Y., Wang, Y., Luo, X., & Zhang, M. (2024). A Comprehensive Survey on Deep Graph Representation Learning. *Neural Networks*, *173*, 106207. https://doi.org/10.1016/j.neunet.2024.106207

Masae, M., Glock, C. H., & Grosse, E. H. (2020). Order picker routing in warehouses: A systematic literature review. *International Journal of Production Economics*, *224*, 107564. https://doi.org/10.1016/j.ijpe.2019.107564

Mazyavkina, N., Sviridov, S., Ivanov, S., & Burnaev, E. (2021). Reinforcement learning for combinatorial optimization: A survey. *Computers & Operations Research*, *134*, 105400. https://doi.org/10.1016/j.cor.2021.105400

Psaraftis, H. N., Wen, M., & Kontovas, C. A. (2016). Dynamic vehicle routing problems: Three decades and counting. *Networks*, *67*(1), 3–31. https://doi.org/10.1002/net.21628

Ratliff, H. D., & Rosenthal, A. S. (1983). Order-Picking in a Rectangular Warehouse: A Solvable Case of the Traveling Salesman Problem. *Operations Research*, *31*(3), 507–521. https://doi.org/10.1287/opre.31.3.507

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(1), 4–24. https://doi.org/10.1109/TNNLS.2020.2978386

DOCTORAL CONSORTIUM

WHO WILL KEEP US SAFE? RETHINKING CYBERSECURITY COMPETENCIES FOR AI-TRANSFORMED WORK IN SMES AND SMALL STATES

ANDREJA MARKUN

University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
andreja.markun@student.um.si

DOI
https://doi.org/10.18690/um.fov.4.2026.78

ISBN
978-961-299-152-4

Cybersecurity is ultimately a human challenge. The ability of organisations to defend themselves depends on the competencies of their people -- yet the competency frameworks designed to support workforce development have not kept pace with how artificial intelligence is transforming what cybersecurity work actually demands. This doctoral research investigates this transformation, focusing on two underserved contexts: small and medium-sized enterprises (SMEs), where one person may be responsible for all security functions simultaneously, and small states, where constrained labour markets and full EU regulatory obligations create a structural mismatch that generic frameworks cannot address. Building on prior empirical research among 235 Slovenian organisations that documented widespread competency gaps and multi-dimensional barriers to their resolution, the research will develop and empirically validate a new competency model designed for AI-augmented cybersecurity environments in SMEs and small states. This paper presents the research plan and explicitly invites feedback on its design and open methodological questions.

Keywords:
cybersecurity competencies,
artificial intelligence,
SMEs,
small states,
workforce development

University of Maribor Press

1 Introduction

Behind many cybersecurity breaches lies a human gap: between what a person knows and what the situation demands, between the competencies an organisation has developed and those it actually needs. The integration of AI into security operations is widening this gap by reshaping what cybersecurity professionals do, how they work, and which knowledge and skills are critical, while competency frameworks have evolved only slowly.

A security professional in a small Slovenian company may patch systems, manage access, respond to incidents, train colleagues, and navigate regulatory obligations under NIS2 and the new ZInfV-1 law. AI-powered tools automate routine tasks, generate alerts, and suggest responses, yet no existing framework tells this practitioner how to govern such tools, when to trust AI output, or how to explain AI-driven decisions to non-technical managers. The frameworks that shape their development were built for a different organisational reality.

Across Europe, AI adoption in cybersecurity is creating a new category of competency gap: not only a shortage of staff, but a mismatch between existing knowledge and evolving role demands. In Slovenia, 70.6% of organisations report shortages of qualified cybersecurity staff and face barriers extending beyond headcount, including insufficient resources, high turnover, weak development culture, and regulatory demands that grow faster than internal capacity (Markun & Brezavscek, 2025). This points to a structural problem that cannot be solved by hiring alone.

The challenge is particularly acute for SMEs and small states. SMEs – over 99% of EU businesses – often rely on a single security generalist, with no formal team or significant training budget, yet face similar threats and regulatory requirements as large organisations. In small states like Slovenia, these organisational constraints intersect with a labour market that cannot supply enough specialised cybersecurity professionals for a fully differentiated workforce (Blazic, 2021).

This doctoral research starts from these realities. It investigates how AI is transforming cybersecurity work and examines whether existing competency frameworks can support this transformation, particularly for SMEs and small states

that they were never designed to serve. Building on prior empirical work in Slovenia, the research aims to produce a validated competency model that reflects resource-constrained, AI-augmented environments and invites critical engagement from Doctoral Consortium advisors and peers.

2 Problem Definition

2.1 What Existing Frameworks Assume - and What They Miss

The cybersecurity competency landscape is shaped by four key frameworks. NICE maps the workforce into 33 specialty areas with detailed knowledge, skill, and ability statements. ENISA's European Cybersecurity Skills Framework (ECSF) defines 12 professional profiles for the European context (ENISA, 2022). CyBOK organises foundational knowledge across 19 domains, mainly for education (Rashid, 2021). ISC2 certification schemes define professional standards through credentials such as CISSP (ISC2, 2025).

All four frameworks assume large organisations with dedicated security teams, clear role boundaries, specialist career paths, and institutional resources for continuous professional development. This assumption is embedded in role profiles, competency statements, and learning pathways. Analyses of EU-based frameworks highlight limited flexibility for resource-constrained contexts and inadequate coverage of competencies needed in high-stakes environments with limited capacity (Almeida, 2025; Chowdhury & Gkioulos, 2021).

Three gaps are central to this research. First, current frameworks provide little guidance for multi-role environments where one person is responsible for all security functions – a common reality in SMEs (Rawindaran et al., 2025). Second, they overlook labour-market constraints and regulatory implementation challenges in small states (Blazic, 2021). Third, they insufficiently address how AI is restructuring cybersecurity work.

2.2 AI and the Changing Nature of Cybersecurity Work

AI is changing the character of security work – which tasks exist, what human judgement is needed, and what counts as competence. A review of 232 empirical studies shows AI-driven change across threat detection, incident response, vulnerability management, and security monitoring (Kaur et al., 2023). Machine-learning techniques are spreading across security operations, creating demand for professionals who can evaluate, govern, and critically assess AI systems, not just operate tools (Sarker et al., 2021).

This transformation has three dimensions. Some competencies are being replaced: log correlation, signature-based detection, and repetitive monitoring are increasingly automated, shifting focus from manual analysis to interpreting and governing AI output. Other competencies are being transformed: threat hunting now involves directing AI tools, understanding their assumptions and blind spots, and intervening when automated reasoning fails. New competencies are emerging with no precedent in current frameworks, including AI governance, explaining AI-driven decisions to non-technical stakeholders, detecting bias in models, and judging when human override is necessary (Zhang et al., 2022; Abidin et al., 2025).

For an SME security professional, this is particularly demanding. Without a team of specialists, a single practitioner must manage the transition from performing security tasks to overseeing AI systems that do so, while still handling other responsibilities as the only security person. No existing framework has been designed with this situation in mind.

2.3 The Human Reality of Cybersecurity in SMEs and Small States

The SME competency challenge is structurally different from the enterprise challenge. Studies of Welsh SMEs show that existing frameworks offer little actionable guidance for professionals covering multiple roles in organisations without formal security processes, dedicated budgets, or specialist support (Rawindaran et al., 2025). The problem is a misfit between tools and workplace realities, not a lack of motivation.

Research on cybersecurity competencies in SMEs remains limited. Most evidence on workforce development comes from large-organisation contexts, and frameworks often assume conditions SMEs cannot meet (Chaudhary et al., 2023). There is, however, clear demand for cost-effective, scalable, and context-appropriate frameworks that work within SME constraints (Jayathilaka & Wijayanayake, 2025).
In small states, organisational constraints intersect with structural labour-market limitations. Slovenia has a small pool of specialists, professionals often work across multiple institutional settings, and the state must implement NIS2, GDPR, ZInfV-1 and DORA with far less governance capacity than larger member states (Blazic, 2021; ENISA, 2025). Existing frameworks do not start from this context.

2.4 The Research Gap and Central Question

AI-driven change, SME constraints, and small-state conditions together create a misalignment between what cybersecurity work demands and what competency frameworks specify. Studies on frameworks seldom address AI or SME contexts (Almeida, 2025; Chowdhury & Gkioulos, 2021). Research on AI in cybersecurity rarely translates technical findings into competency implications (Kaur et al., 2023). SME-focused studies often omit AI-related competencies (Rawindaran et al., 2025). The small-state dimension remains under-researched (Ruoslahti et al., 2021; Nkongolo et al., 2023). This doctoral research addresses their intersection.
.
The central research question is:

How does artificial intelligence reshape cybersecurity competency requirements, and to what extent do existing competency frameworks reflect these changes in the context of SMEs and small states?

Six sub-questions guide the investigation:

- RQ1: How are AI-related competencies currently represented in major cybersecurity competency frameworks?
- RQ2: How do AI-related competency demands appear in real-world labour market requirements?
- RQ3: What are the key misalignments between framework-defined competencies and actual practice in SMEs?

- RQ4: How do competency requirements differ for multi-role cybersecurity professionals in SME contexts?
- RQ5: How do organisational size, resource constraints, and national context shape competency requirements?
- RQ6: How can a revised, empirically validated competency model be developed for AI-augmented SME environments in small states?

3 Research Contribution

The doctoral research will make three distinct and interconnected contributions to knowledge.

First, it offers a theoretical contribution by proposing a conceptual framework that treats cybersecurity competencies as dynamic functions of AI adoption and organisational context rather than static role attributes. Existing frameworks assume relatively stable profiles that can be mapped and periodically updated; in an AI-driven environment, that assumption no longer holds. The model introduces a categorisation of AI-driven competency change – replaced, transformed, newly required – integrated with a multi-level architecture that explicitly incorporates SME constraints and small-state context.

Second, it contributes empirically by building on the Slovenian survey (Markun & Brezavscek, 2025) to produce the first systematic evidence base on AI-related cybersecurity competency gaps in SMEs within a small EU member state. The doctoral research extends the earlier work by isolating AI-specific and SME-specific dimensions and comparing them with the general competency gap picture.

Third, it offers a practical contribution in the form of an operationally usable competency model tailored to organisations that cannot afford role specialisation and must manage AI-augmented security under resource constraints. Rather than simplifying NICE or ECSF, it aims to provide a tool for the organisational reality of the approximately 25 million SMEs across the EU and the people who work in them (ENISA, 2025), supporting SME managers, HR practitioners, training providers, and policymakers in small member states.

4 Methodology

The research adopts a mixed-methods approach structured across five sequential phases (Table 1). Qualitative and quantitative methods are combined to provide both deep theoretical grounding for model design and empirical validation in real organisational contexts..

Table 1: Research design overview

Phase	Purpose	Methods / Tools	Output
1	Systematic Literature Review	PRISMA; Scopus, WoS, ScienceDirect, IEEE Xplore (2020-2026)	Synthesis of knowledge; research gap map
2	Comparative Framework Analysis: coverage and gaps in NICE, ECSF, CyBOK, ISC2	Structured content analysis; gap matrix coding scheme	Gap matrix: missing dimensions per framework
3	Conceptual Model Development: integrate AI transformation, SME constraints, small-state context	Theory synthesis; AI competency categorisation	Draft conceptual model ready for validation
4	Empirical Validation: test model in Slovenian SMEs; method combination to be finalised (see Table 2)	Mixed methods — options under consideration	Empirically refined and validated model
5	Model Finalisation and Publication	Expert review panel; iterative revision	Final validated competency framework

Source: Own

4.1 Phases 1 and 2: Foundations

Phase 1 will conduct a PRISMA-guided review across four domains: existing competency frameworks and their limitations; AI-driven transformation of cybersecurity tasks; workforce challenges in SMEs; and small-state dynamics in cybersecurity workforce development. Searches in Scopus, Web of Science, ScienceDirect, and IEEE Xplore (2020–2026) will build on initial scoping, which already shows rich AI-focused work but few SME-specific studies.

Phase 2 will compare NICE, ENISA ECSF, CyBOK, and ISC2 using a coding scheme that captures (1) competency coverage and role-specialisation assumptions, (2) organisational context assumptions, including multi-role and resource-constrained environments, and (3) AI-related content such as AI governance, explainable AI, and human–machine teaming. The resulting gap matrix will directly inform the conceptual model.

4.2 Phase 3: Building the Conceptual Model

The conceptual model will integrate the Phase-2 gap matrix with insights from the literature review. It will operate at macro, educational, and organisational levels: accounting for the regulatory environment (NIS2, ZInfV-1, GDPR, DORA) and small-state labour-market dynamics; mapping gaps between existing training and SME needs; and reflecting the reality of multi-role professionals with limited time, budgets, and specialist support.

At its analytical core, the model will distinguish competencies replaced by AI automation, competencies transformed by AI augmentation, and new competencies created by AI adoption, including AI governance, explainable AI, AI ethics, and human–machine collaboration. This structure provides a basis for deciding which existing competencies to deprioritise, reframe, or newly specify. Phase 3 thus produces a testable model.

4.3 Phase 4: Empirical Validation

Phase 4 will test and refine the model in Slovenian SMEs. Several methodological options are under consideration.

A Delphi study with 15–20 cybersecurity experts, HR practitioners, and educators could build normative consensus around the model but is constrained by the size of Slovenia's expert pool. Semi-structured interviews with practitioners in SMEs would provide rich qualitative insight into how AI is experienced and what competency needs it creates. A quantitative survey extending the validated instrument from the earlier Slovenian study would allow statistical assessment across a broader organisational sample. A combined approach offers the most robust validation but

requires careful sequencing and substantial resources. The final design will be informed by a pilot phase and practical considerations around participant availability.

Table 2: Methodological options for Phase 4 empirical validation

Criterion	Delphi Method	Semi-structured Interviews	Quantitative Survey	Combined Approach
Builds normative expert consensus	Yes	Partial	No	Yes
Captures practitioner context and nuance	Partial	Yes	No	Yes
Enables statistical generalisation	No	No	Yes	Yes
Feasibility in small-state expert pool	Medium	High	Medium	Medium
Time and resource requirements	High	Medium	Low-Medium	High
Extends prior survey instrument	No	Partial	Yes	Yes
Suitability for normative model development	High	Medium	Low	High
Key limitation	Small expert pool; time-intensive	Not generalisable	Cannot build normative consensus	High complexity; careful sequencing needed

Source: Own

4.4 Phase 5: Finalisation

Findings from Phase 4 will be integrated with the conceptual model and subjected to structured expert review before publication. The publication strategy will target both academic and practitioner audiences, reflecting the model's intended application in SMEs and in national workforce-development policy.

5 Preliminary Results

5.1 What We Already Know: Evidence from Slovenian Organisations

A survey of 235 Slovenian organisations in May–June 2025 provides the empirical baseline (Markun & Brezavscek, 2025). The sample covered all size categories, ownership types, and sectors; about half were NIS2 or ZInfV-1 obligated entities.

Cybersecurity workforce availability was rated significantly below the neutral midpoint on a five-point scale ($M = 2.74$, $SD = 0.84$; $t(234) = -4.69$, $p < 0.001$, Cohen's $d = -0.31$). A total of 70.6% of organisations reported at least partial shortages of qualified staff, significantly above 50% (binomial test, $p < 0.001$). Shortages therefore appear as a general feature of the landscape.

Public-sector organisations reported higher perceived shortages ($M = 2.45$) than private ($M = 1.79$) or mixed-ownership organisations ($M = 1.73$), with a moderately large effect ($F(2, 232) = 23.82$, $p < 0.001$, $\eta^2 = 0.17$). Structural constraints in public-sector recruitment thus translate directly into higher experienced shortages, which a usable competency framework must acknowledge.

Organisations that treat cybersecurity as a strategic priority are more likely to recognise specific competency shortfalls ($\chi^2 = 59.42$, $df = 8$, $p < 0.001$, Cramer's $V = 0.36$), indicating that perceived gaps reflect both objective shortages and organisational awareness. A framework must therefore support organisations at different maturity levels.

Key barriers to competency development included shortages of qualified staff in the labour market (64.4%), insufficient resources for hiring or outsourcing (50.6%), and high staff turnover (47.8%). Factor analysis identified four underlying dimensions: organisational culture and leadership support, external resource and cost constraints, workforce-market instability, and weak internal motivation for professional development. These factors interact and reinforce one another.

Regulation has so far mainly driven training. NIS2 and ZInfV-1 obligated entities showed increased demand for additional training of existing staff ($\chi^2 = 5.25$, $p = 0.022$, Cramer's $V = 0.22$), but not for broader changes such as new role creation or increased leadership involvement. The doctoral research will explore this gap between regulatory pressure and organisational capacity.

The 2025 study did not address AI-related competencies, did not treat SMEs as a separate category, and did not examine the fit between organisations' needs and existing frameworks. These omissions define the agenda for the doctoral research.

5.2 Scoping the Literature

A preliminary scoping search in Scopus, Web of Science, and IEEE Xplore (2020–2026) found more than 570 results for queries combining cybersecurity competency gaps and AI, but fewer than 25 for queries combining cybersecurity competencies and SMEs. This asymmetry mirrors practice: research has followed new technologies such as AI more than the SME environments where much cybersecurity work takes place (Ruoslahti et al., 2021). Phase 1 will address this through a full PRISMA review.

5.3 Open Questions for Doctoral Consortium Discussion

The following questions are genuine open issues in the research design, presented to invite critical feedback:

- Validation methodology (Phase 4): Which combination of Delphi, interviews, and survey best balances rigour and feasibility in Slovenia?
- Scope of empirical validation: Should the study focus solely on Slovenia as a representative small EU member state, or include additional small states to strengthen generalisability?
- Operationalising AI driven competency change: How should the boundary between "transformed" and "newly required" competencies be defined in empirical instruments, and is the tripartite categorisation optimal?
- Positioning relative to ENISA ECSF: Should the model be framed as a contextual extension of ECSF for SMEs and small states, or as an independent, complementary model with its own policy logic?

6 Future Development

Immediate priorities are the PRISMA systematic literature review and the comparative analysis of NICE, ECSF, CyBOK, and ISC2, which will underpin conceptual model development and later empirical work.

More broadly, the research is driven by the view that the competency challenge in cybersecurity is a human and organisational issue. SME practitioners across Europe operate in conditions that current frameworks rarely recognise: they adopt AI tools

without clear guidance, carry regulatory obligations without specialist support, and build competencies largely in isolation (Rawindaran et al., 2025; Jayathilaka & Wijayanayake, 2025). A validated competency model that starts from these conditions, rather than from large-organisation assumptions, could impact how organisations develop staff, how educators design training, and how policymakers support workforce development in small states.

The Doctoral Consortium is an early but important milestone: a chance to test and refine the research design before the most resource-intensive phases begin. The questions around methodology, scope, operationalisation, and positioning are genuine and will materially shape the project, so feedback on the open issues in Section 5.3 is explicitly sought.

References

Abidin, A., Ariffin, N., Nasruddin, Z., Habidin, N., & Yusoff, M. (2025). Evaluating and modelling artificial intelligence and emotional intelligence to improve cybersecurity employee ethical competence model. Journal of Advanced Research Design, 130(1), 13-25. https://doi.org/10.37934/ard.130.1.1325

Almeida, F. (2025). Comparative analysis of EU-based cybersecurity skills frameworks. Computers & Security. https://doi.org/10.1016/j.cose.2025.104329

Blazic, B. J. (2021). The cybersecurity labour shortage in Europe: Moving to a new concept for education and training. Technology in Society, 66, 101769. https://doi.org/10.1016/j.techsoc.2021.101769

Chaudhary, S., Gkioulos, V., & Katsikas, S. (2023). Developing cybersecurity education and awareness programmes. Computer Science Review, 50, 100592. https://doi.org/10.1016/j.cosrev.2023.100592

Chowdhury, N., & Gkioulos, V. (2021). Key competencies for critical infrastructure cyber-security. Information and Computer Security, 29(5), 697-723. https://doi.org/10.1108/ICS-07-2020-0121

ENISA. (2022). European Cybersecurity Skills Framework (ECSF) - Role profiles. https://www.enisa.europa.eu/publications/european-cybersecurity-skills-framework-role-profiles

ENISA. (2025). Cybersecurity roles and skills for NIS2 essential and important entities. https://www.enisa.europa.eu/publications/cybersecurity-roles-and-skills-for-nis2-essential-and-important-entities

ISC2. (2025). 2025 ISC2 Cybersecurity Workforce Study. https://www.isc2.org/Insights/2025/12/2025-ISC2-Cybersecurity-Workforce-Study

Jayathilaka, H. M. T. N., & Wijayanayake, J. (2025). Systematic literature review on developing an AI framework for SME cybersecurity. Journal of Desk Research Review and Analysis, 2(1), 15-30. https://doi.org/10.4038/jdrra.v2i2.53

Kaur, R., Gabrijelcic, D., & Klobucar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, 101804. https://doi.org/10.1016/j.inffus.2023.101804

Markun, A., & Brezavscek, A. (2025). Skills gaps and workforce challenges in information and cyber security: Insights from Slovenia. Uporabna informatika. https://doi.org/10.31449/upinf.260

Nkongolo, M., Mennega, N., & Van Zyl, I. (2023). Cybersecurity career requirements: A literature review. In Emerging Technologies in Computing. Springer. https://doi.org/10.1007/978-981-99-7962-2_7

Rashid, A. (Ed.). (2021). Cyber Security Body of Knowledge (CyBOK), version 1.1. University of Bristol. https://www.cybok.org/

Rawindaran, N., Jayal, A., & Prakash, E. (2025). Cybersecurity framework: Addressing resiliency in Welsh SMEs. Journal of Cybersecurity and Privacy, 5(2), 17. https://doi.org/10.3390/jcp5020017

Ruoslahti, H., Coburn, J., Trent, A., & Tikanmaki, I. (2021). Cyber skills gaps - A systematic review of the academic literature. Connections: The Quarterly Journal, 20(2), 43-58. https://doi.org/10.11610/connections.20.2.04

Sarker, I., Furhad, M., & Nowrozy, R. (2021). AI-driven cybersecurity: An overview, security intelligence modelling and research directions. SN Computer Science, 2. https://doi.org/10.1007/s42979-021-00557-0

Zhang, Z., Hamadi, H., Damiani, E., Yeun, C., & Taher, F. (2022). Explainable artificial intelligence applications in cyber security. IEEE Access, 10, 93104-93139. https://doi.org/10.1109/ACCESS.2022.3204051

About the author

Andreja Markun holds a master's degree in Information Systems and works in the field of information security. Her research focuses on systemic approaches to developing cybersecurity competencies, the impact of regulation on human-resources measures, and the long-term resilience of organisations in the digital environment.

DOCTORAL CONSORTIUM

DESIGNING STRATEGIC PLANNING FOR HIGHER EDUCATION TO COLLECTIVELY SHAPE DIGITAL TRANSFORMATION

GÜL AKCAOVA

HU University of Applied Sciences, Research Center Digital Business & Media, Utrecht, the Netherlands
gul.akcaova@hu.nl

Educational institutions are exposed to a rapid changing environment, including digital transformation (DT). As digital transformation is a complex societal problem this can't be navigated by a single institution. Often organisations collaborate in networks to address wicked problems. However, a better understanding is needed about mechanisms for coordinating and planning activities in collaborative networks, especially for the development of shared understanding regarding a wicked problem like DT. This research intends to prototype and test interventions for networks by conducting participatory action research with multiple-case studies.

DOI
https://doi.org/10.18690/um.fov.4.2026.79

ISBN
978-961-299-152-4

Keywords:
digital transformation, strategic planning, network administrative organisation (NAO), network management, organisational capability, collective decision-making

University of Maribor Press

1 Introduction

The rapid pace of social, technological, economic, ecological and political developments leads to describing the Zeitgeist as volatile, uncertain, complex and ambiguous (Taskan et al., 2022). A few events, for example, in the education sector are the COVID-19 pandemic that forced virtual classrooms, large language models (LLMs) that became accessible to students, and (inter)national politics that cut budgets (UNL, 2025). Such events have raised the question: could we have seen this coming? Is it possible to collectively foresee and prepare for a changing context in the digital transformation?

The DT is a distinct and multifaceted process with different phases (Brooks & McCormack, 2020). To be more resilient and adaptive, a handful of Dutch educational institutions strengthened their strategic organisational capabilities to develop outside-in perspectives and long-term (strategic) planning (e.g. Delft University of Technology). However, DT isn't a straightforward process that can be addressed by a single institution. To navigate the complexity of IT and DT, there are collaborations between knowledge institutions, e.g. vocational education and training, higher education and university medical centres. To manage such collaborations, networks are established in the Netherlands, such as SURF (a collaborative IT organisation for education and research) and Npuls (a national initiative by the education sector).

Asking the question 'what is digital transformation (DT)?' results in divergent interpretations and inconsistencies among vision, policy and implementation in the absence of a well-defined digital vision (Singun, 2025). DT is an ongoing journey that drives increasingly significant organisational changes over time (Singun, 2025). So, the complexity of DT, along with the absence of a shared vision on digital transformation (or digitalisation), causes the risk of being an ineffective collaborative network (Brunetti et al., 2020). To understand what makes network management efficient and effective, further research is necessary. The know-how on how to collaborate is needed, and especially a better understanding of mechanisms and interventions in the network (Verhoest et al., 2024, p.241).

Therefore, this study proposes to investigate the following question: **How can strategic planning for digital transformation be facilitated by a collaborative network in an educational ecosystem?**

In the following sections, key concepts and perspectives are explored and scoped based on an initial literature review. Hereafter, a research design and methodology are presented for this research, including steps in the forthcoming years.

2 Literature review

In this literature review, an initial scoping is provided of the key concepts as articulated in the research question: digital transformation, (collaborative) network management, and strategic planning process. After the key concepts, a synthesis is provided, followed by a conceptual framework.

For the initial scoping and for future research, the following search strategy is applied to the literature review, as listed in Table 1.

Table 1: Search strategy literature review

Keywords	Network (manager OR governance) OR orchestration) AND ("strategic*" OR plan* OR vision*") AND (Digital (trans* OR innovation))
Database	EBSCO, ACM digital library, AIS eLibrary, Web of Science, Google Scholar, Wiley, Sage, IEEE
Search criteria	English/articles/full text/peer reviewed journals/open source
Period	2022 - 2026 & seminal works
Inclusion	All keywords mentioned in abstract, # citations
Exclusion	DT limited to single organisations; short-term innovation; commercial organisations;

2.1 Concept: Digital Transformation

This section does not provide an all-encompassing investigation of what Digital Transformation (DT) entails but provides some streams of thinking relevant for this research. A clear trend in the number of article publications in Elsevier journals covering DT can be observed across time (see Figure 1). Based on a set of definitions, we define a working definition suitable for this research.

Digital Transformation is referred to in various ways across studies. For an initial scoping, four definitions are identified. First, Gong and Ribiere (2021, p. 12) defined DT as "a fundamental change process, enabled by the innovative use of digital technologies accompanied by the strategic leverage of key resources and capabilities, aiming to radically improve an entity and redefine its value proposition for its stakeholders". A second definition, according to Verhoef et al. (2021, p. 890), refers to DT as a multidisciplinary character that is inherently linked to changes in strategy, organisation, etc., due to the implementation of a digital technology. A third definition by Dąbrowska et al. (2022) refers to DT as "a socioeconomic change across individuals, organisations, ecosystems, and societies that is shaped by the adoption and utilisation of digital technologies". These three definitions emphasise the complexity of DT and highlight the interplay between digital technologies and the adoption rate (resource and capability) by an entity (individual, organisation and/or ecosystem) to change or improve. Another perspective, fourth, is also provided by Verhoest et al. (2024, p. 242), who refer to DT as "a complex process of mutual shaping between different actors and technologies, institutions, and economic, political, and socio-cultural factors".

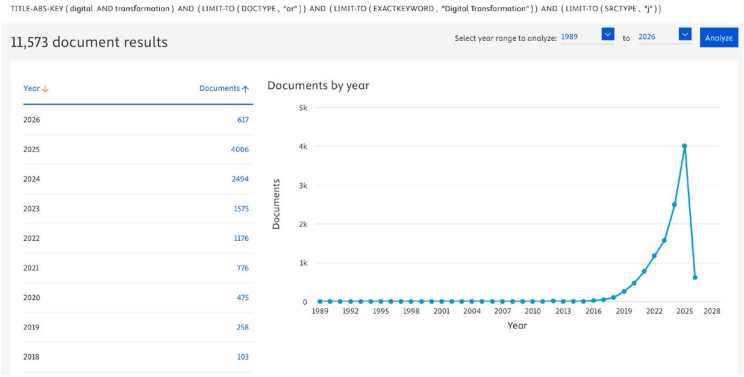

Figure 1: # article publications in journals with the keyword 'digital transformation'
Source: Scopus.com

Although this is a small overview of perspectives for defining digital transformation, we can observe a common denominator for DT as a socioeconomic change driven by the adoption and use of digital technologies. However, we can also identify two perceptions: 1) DT as a change in society due to the use of technology, and 2) DT that is shaped mutually by actors and factors, resulting in a socio-economic change.

As the complexity around DT is derived from the socioeconomic change and the interplay of multiple (f)actors, it is not the technology that makes this process complex, but rather the human in the loop and how culture, attitude and behaviour shift with digital technology. The definitions by Dąbrowska et al. (2022) and Verhoest et al. (2024) imply an active role for actors (entities) to influence how digital technology is adopted and used. Therefore, for this research, **DT will be referred to as a socioeconomic change driven by the adoption and use of digital technologies**.

In the context of higher education, DT can be distinguished into two challenges. First, education needs to rethink the teaching and learning processes that are aligned with the socioeconomic change to cater for the digital natives through knowledge and technical skills, and being digital, through the incorporation of skills, mindset, and digital attitudes (Oliveira & De Souza, 2021) as well as the economy that requires new c.q. different skills. Second, education as an organisation also requires a significant shift in culture, workforce dynamics, and technology that facilitate the implementation of innovative educational and operational frameworks, influencing an institution's operations, strategic objectives, and value proposition (Singun, 2025). This transformation encompasses a comprehensive rethinking of processes, pedagogies, and institutional cultures, therefore confirming the complexity of DT. To navigate the complexities of DT, higher education must adopt a holistic approach and create a shared vision for digital initiatives with stakeholders (Singun, 2025).

To summarise, DT in higher education (HE) is a complex topic as it touches on the business model: rethinking the teaching and learning process, the value of knowledge and operational processes like IT.

2.2 Concept: (Collaborative) network management

In this section, we provide an initial scoping of the concept to identify a working definition. First, a collaborative network refers to a network that comprises autonomous yet interdependent organisations collaborating for a common goal and/or a complex problem (Klijn et al., 2025). According to Provan & Kenis (2007), in large networks, often more than eight organisations, it is most likely that a network administrative organisation (NAO) fulfils the network management role. A collaboration of organisations benefits from better orchestration and the

identification of the (shared) value proposition (Pouru-Mikkola et al., 2023). Therefore, network management plays an important role in how a network is steered (governance). Beyond orchestration, network management also influences decision-making with mechanisms and interventions (Cristofoli et al., 2015). As an example of influential network management, sometimes the collective decision-making is steered to discuss non-controversial items first and push the controversial ones to the end. This is done on purpose to stimulate deliberations early, to build trust and cooperation by the time more controversial topics are addressed (Doberstein, 2016).

As a network has a purpose, we briefly explore network effectiveness to clarify what is expected from network management before defining 'network management'. Turrini et al. (2010) identified effectiveness on the network-level as: 1) ability to achieve stated goals; 2) sustainability and viability; and 3) ability to innovate and change. However, Bayne et al. (2017) identifies network effectiveness as the achievement of outcomes/goals and value gains. As effectiveness can mean different things depending on the maturity of the network and/or per (organisational) actor. Based on Bayne et al (2017), here network effectiveness relates to value gains as effectiveness is a matter of perceived value per actor.

For this initial scoping, a few definitions of network management are identified. According to Klijn et al., (2025), network management is "to deliberate strategies aimed at facilitating and guiding interactions and/or changing the features of the network with the intention to collaborate further in network governance". The motivation per actor can vary, and therefore, the perception of an effective network. Another perspective on network management relates to the design of mechanisms for active, balanced, and continuous interaction or the achievement of efficient coordination (Verhoest, Hans Klijn, et al., 2024). Reducing costs through the coordination of activities and resources is considered a main advantage (Bayne et al., 2017). Therefore, network management needs to be efficient (Verhoest, Hans Klijn, et al., 2024). To expand on the definition by Klijn et al. (2025), the suggested definition of network management is the **facilitation and guidance of interactions through interventions and mechanisms to create value in the network.**

2.3 Concept: Strategic planning

The last key concept for the initial scoping is strategic planning. In the literature, multiple signals have been identified regarding the importance of achieving goals and value gains. According to Turrini et al. (2010), network management should focus on establishing clear mission statements of the network. The importance of this process is also highlighted by Bayne et al. (2017). According to Bayne et al. (2017), visioning brings the different organisational actors' views towards a common trajectory. This convergence of views fosters a shared understanding, vision, or goal. Having a vision means having an outlook on what the network aims to achieve in the future and is also known as a strategic intent. To achieve the vision, a strategy that describes the course of action is formulated (Wit, 2018). Foreseeing outcomes will be impossible, but gaining insights for better sense-making of what is happening will prevent a short-sighted exploitation of digitalisation (Brunetti et al., 2020).

This phenomenon does more than produce a vision. Having a shared vision is suggested to create reciprocity in a collaboration (Ansell and Gash, 2007, p. 560), and, therefore, makes organisational actors more comfortable to adapt to different and multiple goals as long as it is consistent with the perception of the vision (Bayne et al., 2017). Ansell & Gash (2007) refer to the process of establishing a clear mission, vision or problem as '**shared understanding**' in their collaborative governance model. They identify various descriptions of this cyclical process in the literature[1]. Shared understanding implies an agreement on a definition of the problem or the relevant knowledge necessary for addressing a problem. Another perspective is provided by Bayne et al. (2017). They refer to it as '**collective sensemaking**' on the individual level rather than the network level. Common understandings, as defined by Ansell & Gash (2007), are relative as each actor will hold different perceptions of their own and others' aims. Therefore, Bayne et al. (2017) consider shared understanding and the "collective mental processes" that lead to similar, overlapping knowledge among individuals to constitute a group.

To compare, Ansell & Gash (2007) and Bayne et al. (2017) approach the process on a relational level. In contrast to Ansell & Gash (2007), Bayne et al. (2017) emphasise individuals' engagement and the development of a mindset to work as a group. Here,

[1] "common mission", "common ground", "common purpose", "common aims", "common objectives", "shared vision", "shared ideology", "clear goals", "clear and strategic direction" and "alignment of core values"

Bayne et al. (2017) bring in the perspective that individual actors, representatives of organisations, hold their own perspectives and interests that could lead to conflict. Bayne et al. (2017) suggest further research on the relationship between collective sensemaking and network performance. A third perspective in literature is provided by Turrini et al. (2010). Turrini et al. (2010) refer to the same phenomenon as the **steering of network processes** and identify these processes as part of network management. In addition to the process, three abilities for network management are identified as critical: 1) decision-making processes and styles, 2) activating actors, gaining trust and building consensus, and 3) the ability to patch and reposition the network objectives according to changes in the external environment.

Thus, based on the three perspectives, there are variations in strategic planning processes by network management. A common understanding, based on the three perspectives, is that strategic planning in the context of network management requires steering of the collaborative process for a shared understanding through collective sensemaking. For this research, **strategic planning is referred to as a strategic capability on a network level for developing shared visions that involve decision-making**. Since visioning and imagining the future state lead to reciprocity, better integration and cohesion in the network, this is a critical mechanism for network management that requires abilities as suggested by Turrini et al. (2010).

2.4 Conceptual Framework

Digital transformation is an environmental change to which a network needs to adapt through strategic planning processes that depend on the ability of the network management of a NAO, which is responsible for fostering collaboration in the network to create value gain. For this study, the concepts of a network and network management are separated.

From the literature and the suggested definition, a few relationships between the concepts can be identified for this research:

1. Digital transformation is identified as the contextual c.q. environmental change to which a network needs to respond through strategic planning processes that are steered by network management.

2. Steering strategic planning processes requires certain abilities from the network management to develop a shared vision regarding digital transformation.
3. Network management benefits from a shared vision on digital transformation to foster collaboration within a network

These suggested relationships lead to the following initial conceptual framework to guide the study (see Figure 2).

Figure 2: Conceptual framework

3 Methodology

To conduct the research, certain design choices are applicable given the research context. In the following sections, the approach and research strategy are explained. This also includes how research integrity and ethics are addressed.

3.1 Approach

This study aims to conduct the research through a professional doctorate (PD) trajectory. A PD[2] is an equivalent of a PhD with a focus on applied science. During the PD, the expectation is to combine research and innovation activities to design interventions for professionals to improve a practice. Given the research question and the knowledge gap, this trajectory provides a foundation to explore, design and assess interventions for network management, network effectiveness, and the

[2] currently a pilot by the Dutch Ministry of Education, Research and Culture

strategic planning to navigate digital transformation. Given the aim of developing interventions for and with professionals, a (participatory) action research strategy will be deployed as it is an emergent and iterative process that focuses on resolving real organisational issues (Saunders et al., 2016).

In the pre-research phase, actors in the context of networks have been consulted in the identification of the problem statement. The same actors (and their organisations) have expressed their interest in participating in this research trajectory. As the aim is to resolve real organisational problems with and for professionals, multiple organisations c.q. professionals will be part of the research to analyse and develop actionable and transferable results. Therefore, networks focusing on DT in the educational sectors will be treated as multiple case studies.

3.2 Research strategy

The research strategy entails a multiple-case study design with an iterative approach to facilitate the development and testing of an intervention. The research contains a few phases: 1) selection, 2) collection & analysis, 3) prototype & testing, and 4) conclusion & implementation.

In the first phase (selection), attention will be paid to the selection criteria of case studies (see Table 2), assessment of individual cases, and the onboarding of the participants. For the multiple case studies, NAOs will be assessed and selected based on the selection criteria: network management, digital transformation and strategic planning (see Table 1). Our first criterion is a NAO fulfilling network management activities. A second criterion requires that a network should focus on digital technologies in/for education. And thirdly, strategic planning activities that are already deployed, or the desire to do more, by the network management.

Table 2: Case selection criteria

Key concepts	Criteria
Network management	A network administrative organisation (NAO) managing a network comprising more than eight organisations
Digital transformation	Digital technologies in/for education
Strategic planning	Activities related to vision building, shared understanding of mission, problem, etc.

Based on these criteria, a few NAOs in the Netherlands, Germany and the United Kingdom have already expressed their interest in participation. Next to the further definition of the case selection, a data collection protocol will be developed to inform all participants and the reliability of the research. In addition, the participants (organisation or individual) will have the option to remain anonymous.

In the second phase (collection & analysis), data will be collected via interviews, documentation and, where possible, by observation. To collect data, employees from NAOs will be approached, who potentially can provide archival records that are not available online. In addition, per case, actors who are part of the network will be invited for an interview. The interviews and documentation provided by participants and interviewees will give insight into the current situation regarding strategic planning and DT, and the perceived value of current activities. Observations provide additional insights in a real-world setting (Yin, 2018). The reports of interviews and observations will be shared with the involved actors for review and validation. Moreover, in this phase, attention will be paid to the chain of evidence and the case study database. Part of this phase is also the analysis of the collected data. Case-based visualisations of current processes related to the deployment of strategic planning activities and DT will be made, e.g. input, throughput, output and outcome. This includes the analysis of the governance structure and a stakeholder mapping per case. A second analysis technique is a synthesis c.q. comparison of the cases to identify patterns across cases (Yin, 2018) and insights for the transferability and sustainability of the to-be-prototyped interventions by the professionals. For the analysis of the qualitative data, coding for the key concepts will be defined based on a literature review.

In the third phase (prototype & testing), together with the professionals, the findings will be discussed to collectively identify and develop prototypes of intended interventions (tools, mechanisms, processes, etc.). The developed prototypes will be tested and evaluated within the networks. Where possible, an observation will be done. The involved professionals and actors will also be interviewed again to gain insights regarding the evaluation. This is an iterative process of multiple prototypes and testing of interventions. As this phase also includes data collection, similar techniques will be applied to analyse the data. As each network might have a slightly different approach, this analysis (visualisation of processes) might lead to (new) insight. The analysis will (again) be used to discuss, co-design, develop and assess

mechanisms with professionals. Per case and cross-case, a comparison of the situation 'before' and 'after' the intervention is to be made to identify the impact.

In the fourth and last phase (conclusion & implementation), the results of the research activities will be concluded. The conclusions will be reported c.q. presented to each case. There will also be an overarching report prepared that captures the deployed research activities, the outcomes and overall insights gained from this study. Additionally, the co-designed interventions that contribute to strategic planning in a NAO will be handed over to the professionals. Part of this phase is also to identify how the interventions (deliverables) can be embedded in the NAOs as part of change (improvement).

3.3 Research integrity

To ensure the integrity of the research and the researcher, a few measures will be taken regarding validity, reliability and ethics (see Table 2). As mentioned in the research strategy, several measures are taken to ensure validity: multiple sources of evidence, review of drafts by participants and co-designing with the participants. The involvement of the participants in the data analysis minimises the risks regarding the validity of the research. For the reliability of the research, a protocol, case study database, and chain of evidence will be established. Ethics are also taken into account by providing options for anonymity, defining a code of conduct (e.g. Chatham House rules), review of external (science) communication and attention will be paid to inclusion and equity of participants in the co-design and co-creation activities.

Table 2: Validity, reliability and ethics

	Measures
Construct validity	Multiple sources of evidence, review of drafts by participants, co-design with participants
Internal validity	Co-designing the visualisation of workflows with participants
External validity	Patterns via cross-case synthesis
Reliability	Case study protocol, development of case study database, chain of evidence
Ethics	Anonymity options, review of science communication, code of conduct

Source: (Yin, 2018)

4 Future development

Over the next year(s), the literature review and research strategy will be further expanded. Activities will be deployed to kick off the multiple-case study research in combination with the co-design process with professionals. Whilst collecting data, the analysis and interventions will be done iteratively and in parallel.

The expected research outputs are peer-reviewed articles, a data management plan, a set of designed and assessed interventions for professionals, and a portfolio of deliverables (reflections on research activities).

References

Bayne, L., Schepis, D., & Purchase, S. (2017). A framework for understanding strategic network performance: Exploring efficiency and effectiveness at the network level. *Industrial Marketing Management*, *67*(June), 134–147. https://doi.org/10.1016/j.indmarman.2017.07.015

Brooks, D. C., & McCormack, M. (2020, June). *Defining Digital Transformation*. Educause. Driving Digital Transformation in Higher Education. https://www.educause.edu/ecar/research-publications/driving-digital-transformation-in-higher-education/2020/defining-digital-transformation

Brunetti, F., Matt, D. T., Bonfanti, A., De Longhi, A., Pedrini, G., & Orzes, G. (2020). Digital transformation challenges: Strategies emerging from a multi-stakeholder approach. *The TQM Journal*, *32*(4), 697–724. https://doi.org/10.1108/TQM-12-2019-0309

Cristofoli, D., Macciò, L., & Pedrazzi, L. (2015). Structure, Mechanisms, and Managers in Successful Networks. *Public Management Review*, *17*(4), 489–516. https://doi.org/10.1080/14719037.2013.798025

Dąbrowska, J., Almpanopoulou, A., Brem, A., Chesbrough, H., Cucino, V., Di Minin, A., Giones, F., Hakala, H., Marullo, C., Mention, A. L., Mortara, L., Nørskov, S., Nylund, P. A., Oddo, C. M., Radziwon, A., & Ritala, P. (2022). Digital transformation, for better or worse: A critical multi-level research agenda. *R&D Management*, *52*(5), 930–954. https://doi.org/10.1111/RADM.12531

Doberstein, C. (2016). Designing Collaborative Governance Decision-Making in Search of a 'Collaborative Advantage'. *Public Management Review*, *18*(6), 819–841. https://doi.org/10.1080/14719037.2015.1045019

Klijn, E. H., Koppenjan, J., Spekkink, W., & Warsen, R. (2025). *Governance Networks in the Public Sector* (2nd edn). Routledge. https://doi.org/10.4324/9781003402138

Oliveira, K. K. D. S., & De Souza, R. A. C. (2021). Digital Transformation towards Education 4.0. *Informatics in Education*. https://doi.org/10.15388/infedu.2022.13

Pouru-Mikkola, L., Minkkinen, M., Malho, M., & Neuvonen, A. (2023). Exploring knowledge creation, capabilities, and relations in a distributed policy foresight system: Case Finland. *Technological Forecasting and Social Change*, *186*, 122190. https://doi.org/10.1016/j.techfore.2022.122190

Provan, K. G., & Kenis, P. (2007). Modes of Network Governance: Structure, Management, and Effectiveness. *Journal of Public Administration Research and Theory*, *18*(2), 229–252. https://doi.org/10.1093/jopart/mum015

Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (Seventh edition). Pearson.

Singun, A. Jr. (2025). Unveiling the barriers to digital transformation in higher education institutions: A systematic literature review. *Discover Education*, *4*(1), 37. https://doi.org/10.1007/s44217-025-00430-9

Turrini, A., Cristofoli, D., Frosini, F., & Nasi, G. (2010a). Networking literature about determinants of network effectiveness. *Public Administration*, *88*(2), 528–550. https://doi.org/10.1111/j.1467-9299.2009.01791.x

Turrini, A., Cristofoli, D., Frosini, F., & Nasi, G. (2010b). Networking literature about determinants of network effectiveness. *Public Administration*, *88*(2), 528–550. https://doi.org/10.1111/j.1467-9299.2009.01791.x

UNL. (2025, April 10). *Education cuts passed – universities go to court.* Universiteiten van Nederland. https://www.universiteitenvannederland.nl/en/current/news/education-cuts-passed-universities-go-to-court

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Qi Dong, J., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research*, *122*, 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022

Verhoest, K., Hammerschmid, G., Rykkja, L. H., & Klijn, E.-H. (Eds). (2024). *Collaborating for digital transformation: How internal and external collaboration can contribute to innovate public service delivery*. Edward Elgar Publishing.

Wit, B. de. (2018). *Strategy synthesis: For leaders* (Fifth edition). Cengage Learning.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (Sixth edition). SAGE.

DOCTORAL CONSORTIUM

CO-CREATING HUMAN-CENTRED AND RESPONSIBLE AI GOVERNANCE IN CONSULTING SMES

LUDMILA JIŘÍČKOVÁ

Prague University of Economics and Business, Faculty of Informatics and Statistics, Prague, Czechia
ludmila.malinova@vse.cz

Small and medium-sized enterprises (SMEs) in consulting increasingly rely on AI tools in client work, yet often lack proportionate and human-centred governance approaches that translate high-level "responsible AI" principles into everyday organisational practice. This doctoral research proposes AIROF – an adaptive AI Risk & Opportunity Framework – grounded in fourth-order cybernetics to enable co-created governance through explicit roles, feedback loops, and continuous contextual adaptation. Using a Design Science Research approach, the study synthesises theoretical foundations (AI governance, risk management, Balanced Scorecard logic, fourth-order cybernetics) and triangulates them with empirical insights in consulting SMEs to design and validate AIROF. The planned instantiation (AIROF Scorecard) will support decision-making without requiring disruptive process redesign. This paper contributes a research agenda and evaluation plan for co-creating human-centred, responsible digital futures in resource-constrained knowledge organisations.

DOI
https://doi.org/10.18690/um.fov.4.2026.80

ISBN
978-961-299-152-4

Keywords:
responsible digital futures, human-centred design, SMEs, cybernetics, risk and opportunity management, design science research

University of Maribor Press

1 Introduction

The 2026 conference theme "Co-Creating Human-Centred and Responsible Digital Futures" highlights a core challenge of organisational AI adoption: digital futures become "responsible" not only through technology choices, but through governance arrangements that keep humans meaningfully in the loop and enable participation, accountability, and learning over time (*European Commission, 2021; UNESCO, 2021*).

In consulting SMEs, AI tools (including generative AI) can accelerate knowledge work and augment service delivery, but they also introduce risks that are tightly coupled with client trust, confidentiality, and professional judgement (*Davenport & Mittal, 2023; Kolbjørnsrud, 2024*). The tension is intensified by rapid AI capability evolution, which compresses governance adaptation cycles from years to months (*Russell & Norvig, 2021; Haenlein & Kaplan, 2019*).

This doctoral research addresses the problem by proposing **AIROF (Artificial Intelligence Risk & Opportunity Framework)**: an adaptive, systemic meta-framework for governing AI use through co-creation (roles, feedback loops, shared sense-making), grounded in fourth-order cybernetics as a theory of context-aware self-governance and continuous learning.

1.1 Research aim and positioning

The dissertation's aim is to identify theoretical foundations for an adaptive framework to manage AI risks and opportunities in consulting SMEs (*Nonaka & Takeuchi, 1995; Davenport & Mittal, 2023*), and to design a practical instrument (AIROF Scorecard) to support adoption and evaluation in at least 15 SME organisations.

This Doctoral Consortium paper focuses on the **conceptual and methodological** contribution (AIROF framing, research questions, and evaluation plan) aligned to the conference theme.

2 Problem definition

2.1 Practical problem

SMEs face disproportionate constraints in AI adoption: limited governance capacity, limited specialist roles, and increased dependency on external AI providers (*Sánchez et al., 2025; Tominc et al., 2024)*. Yet they operate in environments where responsible AI practice requires clear accountability, risk controls, and transparency. Recent OECD evidence shows persistent adoption gaps between SMEs and large firms, and highlights enabling factors such as skills, inputs, and finance implying that governance instruments must be lightweight, implementable, and adapted to SME realities (*OECD, 2024; OECD, 2019*).

In parallel, formal governance frameworks and regulations are expanding. The **EU AI Act** establishes harmonised rules for AI, reinforcing the need for structured governance and risk management across the AI lifecycle (*European Commission, 2024*).

The **NIST AI RMF** provides a voluntary framework to manage AI risks to individuals, organisations, and society *(NIST, 2023)*.

ISO/IEC 42001 introduces an AI management system standard intended to structure organisational governance of AI risks and opportunities *(ISO/IEC 42001, 2023)*.

However, SMEs – especially in knowledge-intensive services–still require a **contextual, human-centred operational bridge** from these high-level expectations to everyday decision-making and learning cycles.

2.2 Research gap and theoretical framing

The dissertation hypothesises that **fourth-order cybernetics** provides a suitable meta-framework for adaptive AI governance because it emphasises context, systemic integration, and the human actor's reflexive role in steering socio-technical systems (*Umpleby et al., 2019; Medvedeva & Umpleby, 2024*).

This perspective naturally aligns with co-creation: governance is not a one-time policy artefact but an evolving set of relationships, roles, and feedback loops that shape responsible digital futures over time (*Božičnik & Mulej, 2011; Chiolerio, 2020*).

2.3 Research questions

The dissertation defines one main and four sub-questions:

RQ0 (main): How can a conceptual framework AIROF be designed and empirically validated for the adaptive management of risks and opportunities of the use of AI tools in consulting SMEs, using fourth-order cybernetics as a meta-framework?

RQ1: Which concepts and principles from fourth-order cybernetics, AI governance, risk management, and the Balanced Scorecard are relevant and sufficiently robust for the design of an adaptive framework for managing risks and opportunities of AI in consulting SMEs?

RQ2: What are the key needs, expectations, and problem areas associated with the use of AI tools in consulting services SMEs?

RQ3: How can a practical tool (AIROF Scorecard) be designed, based on theoretical foundations and empirical findings, using the logic of the Balanced Scorecard, in order to support the management of risks and opportunities of AI in consulting SMEs?

RQ4: What value does the implementation of the AIROF Scorecard bring to consulting organisations in the management of risks and opportunities of AI, in terms of reducing the impact of risks, increasing the utilisation of opportunities, improving transparency, and supporting decision-making?

2.4 Contribution statement

This doctoral research aims to contribute mainly two artefacts:
A conceptual contribution: AIROF as an adaptive, cybernetically grounded governance meta-framework that treats responsible AI as a co-created

socio-technical capability (roles, relationships, and feedback), rather than a static control checklist.

A methodological contribution: a Design Science Research process and evaluation plan tailored to consulting SMEs, connecting theory synthesis with empirical validation and a lightweight artefact instantiation (*Hevner et al., 2004; Peffers et al., 2007*).

A practical contribution: a proportionate governance approach that can complement established frameworks (EU AI Act, NIST AI RMF, ISO/IEC 42001) while remaining implementable in resource-constrained organisations as an operationalised form of AIROF through the AIROF Scorecard tool.

3 Methodology

3.1 Research approach

The dissertation is positioned within a Design Science Research logic, connecting theory, empirical insights, and artefact design through iterative cycles. This is consistent with the Doctoral Consortium's focus on the "research journey" and on sharpening a proposal through feedback and planned validation.

3.2 Design Science Research process and phases

The dissertation structures the work into phases that progress from diagnosis to conceptual design, artefact creation, implementation, and reflexive evaluation.

Table 1: Research phases, methods, and outputs (planned)

Phase	Purpose	Main methods	Key outputs
Diagnostic	Define AI governance and risk–opportunity problem in consulting SMEs	Literature synthesis; initial SME probing	Problem definition; requirements for a lightweight governance approach
Conceptual design	Construct AIROF framework	Theory integration (4th-order cybernetics; AI governance; risk management; Balanced Scorecard); empirical input	AIROF constructs; co-creation logic and role design
Artefact (instantiation)	Create AIROF Scorecard concept/prototype	Scorecard modelling; tool prototyping (Excel/VBA)	Initial tool concept for structured risk/opportunity monitoring
Implementation	Deploy in SMEs	Onboarding + training; guided adoption	Tool use in real organisational context
Reflexive evaluation	Evaluate contribution and adjust	Post-use surveys; follow-up interviews; triangulation	Evidence on perceived value; refinements; design knowledge

Source: Author's own work

3.3 Empirical strategy and evaluation design

The dissertation proposes triangulated validation methods, including semi-structured interviews with management and key staff, the analysis of organisational documents where available (e.g., internal policies), and post-implementation questionnaires after approximately one month of tool usage. It also proposes evaluation across 15 SME size categories (five micro, five small, five medium) within the consulting domain.

This evaluation plan aligns with a human-centred and responsible digital futures frame in two ways:

First, it treats governance effectiveness as **experienced** by human actors (usability, decision support, transparency). Second, it explicitly anticipates adaptation based on feedback, reflecting cybernetic learning rather than one-off compliance exercises (*Argyris & Schön, 1978*).

4 Preliminary and expected results

4.1 AIROF as a co-creation governance loop

The preliminary outcome is the conceptual structure of AIROF: an adaptive framework integrating AI governance principles, risk and opportunity management, and a scorecard logic inspired by Balanced Scorecard (*Kaplan & Norton, 1996; Kaplan & Norton, 2001*) to keep governance actionable and multi-perspective.

To maintain proportionality for SMEs, AIROF is defined to support decision-making without forcing immediate redesign of core business processes.

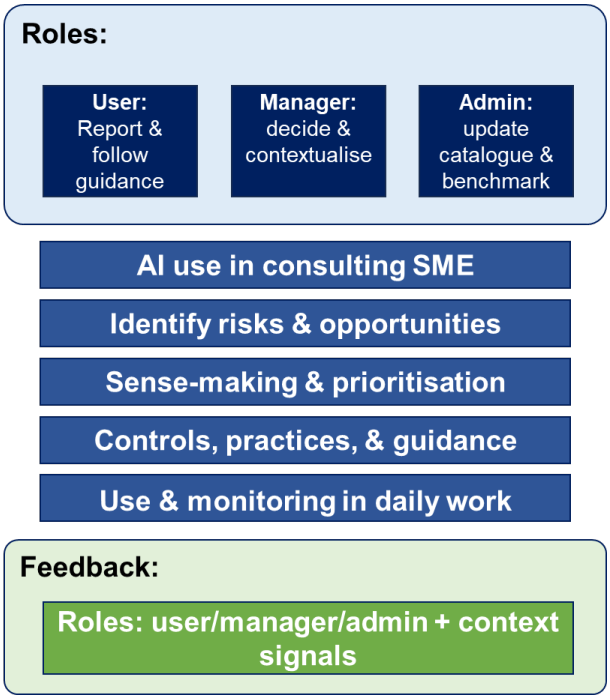

Figure 1: AIROF co-creation loop (conceptual view)
Source: Author's own work.

This representation directly reflects the dissertation's emphasis on adaptive governance cycles, explicit roles (user/manager/admin), and the possibility of cross-organisational learning via benchmarking.

4.2 AIROF Scorecard positioning

The dissertation introduces AIROF Scorecard as the practical instrument to operationalise AIROF through four perspectives and three access/interaction roles, implemented as a lightweight tool (prototype in spreadsheet form).

In this Bled Doctoral Consortium paper, the Scorecard is treated as the planned instantiation and evaluation vehicle; the detailed operational design (metrics, UX, automation, and governance workflows) is under preparation as the next steps of the dissertation.

4.3 Expected evaluation outcomes and validation criteria

The final evaluation will test whether AIROF (and its instantiation) improves working with AI tools in organisational practice. The focus is on consulting SMEs:

- **Transparency**: clearer shared understanding of AI risks/opportunities and responsibilities.
- **Decision support:** better prioritisation and management attention in AI adoption choices.
- **Risk impact reduction and opportunity realisation:** evidence via post-use feedback and benchmarking.
- **Adaptability:** the ability to update the governance catalogue and practices as AI capabilities and regulatory expectations evolve (e.g., in response to the EU AI Act and organisational learning).

5 Future development

5.1 PhD work plan and timeline

Figure 2 outlines the overall research design of the dissertation. The theoretical background – grounded in fourth-order cybernetics, AI governance, risk management, and the Balanced Scorecard – has been completed. The subsequent phases, including empirical data collection among consulting SMEs, the design of the AIROF framework and its scorecard tool, as well as validation through case studies, are currently in progress.

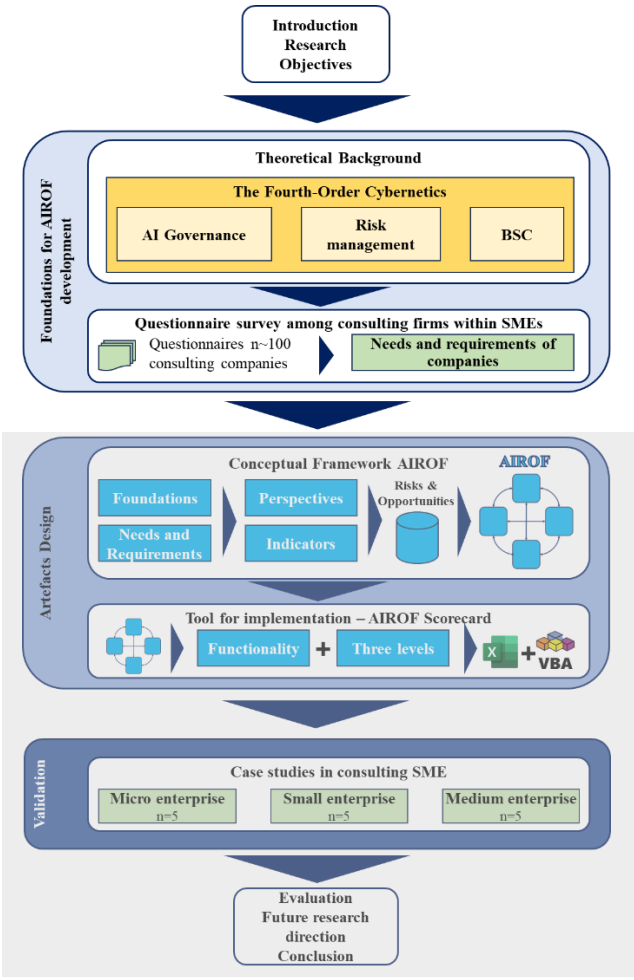

Figure 2: Remaining PhD work (high-level timeline)
Source: Author's own work

5.2 Ethical considerations and responsible research practice

The research involves human participants in SMEs (interviews, surveys) and therefore requires informed consent, careful handling of potentially sensitive organisational information, and clear separation of client data from research data. The research also concerns AI governance, which intersects with risk to individuals and society – consistent with the emphasis in NIST AI RMF on managing AI risks to individuals, organisations, and society.

AI governance in Europe is also a moving regulatory target; aligning AIROF's constructs with the EU AI Act's lifecycle and risk-based logic is part of ensuring that the designed governance approach supports responsible digital futures rather than encouraging superficial compliance (*Floridi & Cowls, 2021; Hagendorff, 2020*).

Similarly, the OECD's values-based principles for trustworthy AI provide a relevant normative baseline for "human-centred" co-creation.

5.3 Planned dissemination

This Doctoral Consortium paper will be used to refine the research design through feedback. The next phase will focus on the AIROF Scorecard operationalisation and implementation details (measurement model, tool features, and evaluation outcomes), thereby separating conceptual framing from artefact-level contribution.

Disclosure of AI use

In the preparation of this manuscript, the author used ChatGPT-5, a generative AI tool developed by OpenAI, for translation and language correction when translating text from Czech to English. The tool was used solely to improve linguistic clarity and ensure accurate translation (with author's corrections afterwards), without generating or altering content. All intellectual and conceptual work remains the responsibility of the author.

References

Argyris, C., & Schön, D. (1978). Organizational learning: A theory of action perspective. Addison-Wesley.

Božičnik, S., & Mulej, M. (2011a). A new–fourth order cybernetics and sustainable future. *Kybernetes*, 40(5/6), 670–684. https://doi.org/10.1108/03684921111142232

Davenport, T. H., & Mittal, N. (2023). All-in on AI: How smart companies win big with artificial intelligence. *Harvard Business Review Press*.

European Commission. (2021). Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L 1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Floridi, L. (2021). Ethics, governance, and policies in artificial intelligence. Springer.

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5–14. https://doi.org/10.1177/0008125619864925

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105. https://www.jstor.org/stable/25148625

International Organization for Standardization, & International Electrotechnical Commission. (2023). ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system. ISO/IEC.
Kaplan, R. S., & Norton, D. P. (1996). The balanced scorecard. Harvard Business School Press.
Kaplan, R. S., & Norton, D. P. (2001). Strategy-focused organization. Harvard Business School Press.
Kolbjørnsrud, V. (2024). Designing the intelligent organization: Six principles for human–AI collaboration. *California Management Review*, 66(2), 44–64. https://doi.org/10.1177/00081256231211020
Medvedeva, T. A., & Umpleby, S. A. (2024b). Cybernetics of large social systems: The example of economics. *Systems Research and Behavioral Science*, 41(3), 524–533. https://doi.org/10.1002/sres.2991
National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. https://www.nist.gov/itl/ai-risk-management-framework
Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford University Press.
Organisation for Economic Co-operation and Development. (2019). OECD principles on artificial intelligence. OECD Publishing. https://www.oecd.org/going-digital/ai/principles/
Organisation for Economic Co-operation and Development. (2024). Update on the implementation of the OECD AI principles. OECD Publishing. https://oecd.ai/en/ai-principles
Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302
Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.
Tominc, P., Oreški, D., Čančer, V., & Rožman, M. (2024). Statistically significant differences in AI support levels for project management between SMEs and large enterprises. AI, 5(1), 136–157. https://doi.org/10.3390/ai5010008
Umpleby, S. A., Medvedeva, T. A., & Lepskiy, V. E. (2019). Recent developments in cybernetics, from cognition to social systems. Cybernetics and Systems, 50(4), 367–382. https://doi.org/10.1080/01969722.2019.1574326

About the author

Ludmila Jiříčková is a doctoral researcher at the Department of Systems Analysis, Faculty of Informatics and Statistics, Prague University of Economics and Business, Czech Republic. Her research focuses on AI governance, systemic approaches to managing risks and opportunities of artificial intelligence, and the application of fourth-order cybernetics in organisational contexts. She combines academic research with practical experience in consulting, digital business, and organisational development. Her work emphasises human-centred and adaptive governance frameworks for SMEs, particularly in knowledge-intensive services. She is currently developing the AIROF framework and its practical implementation tool, the AIROF Scorecard, within a Design Science Research paradigm.

Summary

This paper addresses the challenge of translating high-level AI governance principles into practical, human-centred approaches suitable for consulting SMEs. It introduces AIROF (Artificial Intelligence Risk & Opportunity Framework), an adaptive meta-framework grounded in fourth-order cybernetics, enabling co-created governance through feedback loops, explicit roles, and continuous learning. The research adopts a Design Science Research methodology, integrating theoretical foundations from AI governance, risk management, and Balanced Scorecard logic with empirical insights from SMEs. The planned AIROF Scorecard serves as a lightweight implementation tool supporting decision-making without requiring disruptive process changes. The study contributes a conceptual framework, a

methodological approach, and a practical governance solution aligned with responsible digital futures in resource-constrained organisations.

DOCTORAL CONSORTIUM

AN INTEGRATED DECISION SUPPORT FRAMEWORK FOR MULTI-PHASE FLOOD EVENT MANAGEMENT

VESNA VIDMAR, MIRJANA KLJAJIĆ BORŠTNAR
University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
vesna.vidmar@student.um.si, mirjana.kljajic@um.si

Effective flood event management requires coordinated decision-making across preparedness, response, and recovery. Current decision support systems (DSS) handle these phases separately, leaving operational managers without a unified informational picture when it is most needed. This doctoral research develops an integrated DSS framework that unifies heterogeneous data sources - from hydrological models and IoT sensors to satellite imagery and social media - following a Design Science Research (DSR) approach. In this paper, we present the results of a preliminary literature review aimed at identifying the research gap. The review confirms that no published work integrates data across the full risk management cycle - from early warning through response to recovery. Building on this gap, the paper outlines the research design and expected contributions: (1) a DSS framework addressing cross-phase decision support, and (2) a methodology for heterogeneous data integration in flood management. The research addresses both technical and institutional dimensions of fragmentation.

DOI
https://doi.org/10.18690/um.fov.4.2026.81

ISBN
978-961-299-152-4

Keywords:
decision support system, flood management, heterogeneous data integration, interoperability, disaster response, design science research

University of Maribor Press

1 Introduction

Floods are one of the most common and destructive natural disasters in the world, requiring rapid, informed decisions across multiple organizations. Effective flood event management comprises three phases: preparation (before the event), response (during the event), and recovery (post-event). Each phase involves different actors, data requirements, and decisions.

These decisions are currently supported by a disjointed set of tools. Meteorological forecasts, satellite imagery, IoT sensor data, hydrological models, and logistical data are all maintained in distinct systems that were not designed for interoperability. Under time constraints, operational managers are required to integrate data from multiple heterogeneous sources without access to a fully coherent situational overview.

Slovenia's August 2023 floods made the problem visible. They ranked among the most damaging in the country's recent history, sweeping through buildings, infrastructure, and farmland across several river basins (Bezak et al., 2023). Existing tools could not handle an event of that magnitude. As recovery demands accumulated, fragmented systems fell behind, and decision-makers operated without a unified picture of conditions on the ground.

The research question motivating this doctoral project is*, "How can heterogeneous data sources be integrated into a unified DSS architecture that supports decision-making across all phases of a flood event?"*

This proposal presents the research design, expected contributions, and current progress toward answering that question.

2 Problem Definition

2.1 Organizational and Institutional Dimensions

The fragmentation is not solely technical. Sensor systems often fail to integrate because of how agencies handle the data – not because of the technology itself. One agency may treat a reading as reliable, another may discount it, and a third may interpret the same value under its own protocol (Hong & Shi, 2023).

The main obstacle to next-generation operational flood DSS is the absence of cross-agency collaboration on shared frameworks (Schwanenberg et al., 2018). Recent empirical work identifies three structural barriers behind this gap. First, flood management is institutionally fragmented, split across several sector mandates. Second, incompatible datasets prevent agencies from forming a common operational picture. Third, capability injustice persists: access to advanced tools is unequal, leaving local responders with limited operational information.

This research addresses both the technical and organizational dimensions of integration.

2.2 Research questions

RQ1: What are the documented gaps in existing DSS approaches for flood management, across data sources, supported phases, and architectural design?

RQ2: What data integration approach can consolidate heterogeneous flood management data sources into an interoperable format for operational decision-making?

RQ3: What DSS framework supports decision-making across all three phases of a flood event, and how does it address the identified gaps?

3 Methodology

The research follows a Design Science Research (DSR) approach, grounded in the foundational framework for information systems research by Hevner et al. (2004) and operationalized through the process model of Peffers et al. (2007). This DSR approach is selected because it combines methodological rigor with practical utility — it demands both theoretical grounding and a demonstrable artifact validated against the needs of operational decision-makers. Steps are visually explained in Figure 1.

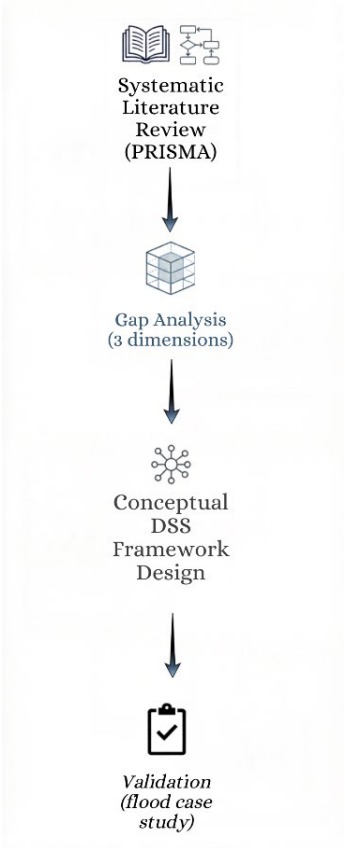

Figure 1: Research Design Following a DSR approach
Source: Own

3.1 Phase 1: Systematic Literature Review

The review performs a structured search across Web of Science, SCOPUS, and ProQuest, using keyword combinations covering DSS, flood management, data integration, interoperability, real-time operation, and heterogeneous data sources. The search strategy is supplemented by snowball sampling following key references (Wohlin, 2014; Snyder, 2019).

Results will be organized in a comparative matrix showing, for each reviewed work, data sources used, flood management phases supported, architectural approach, and documented limitations. This matrix is the primary input for gap analysis.

3.2 Phase 2: Gap Analysis

The gap analysis proceeds across three dimensions:

1. Data sources — which sources are systematically absent from integrated approaches;
2. Phase coverage — which flood management phases lack adequate decision support;
3. Architectural gaps — which interoperability and semantic alignment challenges remain unresolved.

The output is a gap taxonomy serving as the design specification for Phase 3.

3.3 Phase 3: Conceptual DSS Framework Design

The conceptual DSS framework will specify:

- Which data sources to integrate (hydrological models, IoT sensors / Internet of Things, satellite imagery, meteorological forecasts, social media data, logistics data);
- What architecture supports consistent data exchange and semantic interoperability across these sources?
- How the system structure supports different decision types in each flood management phase.

The framework is the primary scientific artifact of the research.

3.4 Phase 4: Validation

The framework will be validated through a case study of the August 2023 Slovenia floods (Bezak et al., 2023), with semi-structured interviews with domain experts providing operational perspective. Two questions guide validation: Would the framework have given decision-makers a more integrated picture than tools available in 2023, and does it address the specific decision points where information was missing or delayed?

4 Preliminary and Expected Results

4.1 Systematic Literature Review

A systematic keyword search across Web of Science, SCOPUS, and ProQuest Dissertations & Theses confirms that existing DSS research for flood management is extensive but partial. The closest matching search string - *"decision support tool AND flood AND web-based AND forecast AND operational AND real-time AND river"* - returns zero results in Web of Science, confirming that no published work addresses the full problem scope. The literature clusters around five distinct areas.

Early warning systems. The strongest concentration of research addresses flood early warning. Multiple studies confirm that IoT sensor networks, machine learning algorithms, and web-based forecasting interfaces can reliably detect flood risk and alert decision-makers (Ilukkumbure et al., 2021; Lin et al., 2022). The technical performance of these systems has improved substantially over the past decade. Yet a consistent operational gap persists: early warning tools function as standalone applications. They generate alerts but do not feed those alerts into the logistics and coordination systems that emergency managers need to act on them. The practice, the warning reaches the right channels, but converting it into coordinated action still depends on manual handoffs between disconnected systems (Hong & Shi, 2023; Pantalona et al., 2021).

Hydrological forecasting models. Flood forecasting has matured technically. Hybrid approaches now blend numerical weather prediction with remote sensing and in-situ data, producing compound flood impact estimates that go beyond simple river-stage prediction (Láng-Ritter et al., 2022; Refadah, 2025). The challenge that remains is operational forecasts are probabilistic outputs, but emergency managers must usually act on a single, readable risk picture. Even an accurate model loses practical value when its confidence cannot be delivered alongside the forecast (Yu et al., 2025; Adler et al., 2025).

Social media and crowdsourcing. Citizen-generated data is gaining traction in flood monitoring research. Geo-referenced social media posts capture real-time flood conditions at a level of detail that fixed sensor networks cannot reach (Dujardin et al., 2025). Pairing crowdsourced observations with remote sensing

improves damage assessment (Akhtar et al., 2024), and feeding them into hydrological models yields measurable gains in forecast performance (Songchon et al., 2023). The gap is structural rather than technical: these data are still treated as research inputs and rarely make it into the real-time tools that emergency managers actually use during an event. Their value sits mostly in post-event analysis.

Coupled agent-institution models. A smaller but theoretically significant body of research moves beyond technical components to examine how institutional factors shape flood response. The CLAIM framework (Abebe et al., 2019) couples hydrological simulation with agent behavior and institutional rules, modeling how organizational protocols and inter-agency mandates influence response effectiveness. This line of work is valuable because it names what purely technical DSS approaches tend to assume away: that institutions will coordinate smoothly if given good data. In practice, coordination depends on shared procedures, compatible authority structures, and agreed-upon definitions of critical concepts. Despite its relevance, this modeling tradition has not been integrated into operational tools.

Recovery-phase systems. Recovery represents the most significant gap in the DSS literature. Existing systems are designed to support the active event phase and are not built to carry information forward into reconstruction decisions. When a flood event ends, the data generated during response — damage locations, affected populations, and resource allocations — is rarely aggregated in a form that supports structured recovery planning. Socioeconomic dimensions are consistently underrepresented in this literature. Community-level vulnerabilities shape the speed and fairness of recovery (Park et al., 2024; Mendis et al., 2024), yet DSS tools rarely capture either dimension. The build-back-better principle, although prominent in policy texts, sees little concrete implementation (Birkmann et al., 2023).

The pattern across these five areas is consistent. Each domain has produced technically capable, well-validated tools that address one component of flood management effectively. What the literature has not produced is an artifact that connects these components — carrying information from early warning through active response into recovery, across the institutional boundaries that separate the agencies involved. Most existing DSSs handle single decision points and stop short of supporting the full risk management cycle (Pantalona et al., 2021). The only

doctoral study addressing all three flood management phases does so exclusively in a developing-country context, making direct transfer to institutional systems such as Slovenia's unreliable (Thomas, 2017). A summary is provided in Table 1.

Table 1: Summary of DSS Research Areas — Key Capabilities and Documented Gaps

Research Area	Key Capability	Documented Gap	Source
Early Warning Systems	Real-time detection, IoT alerts	No integration with response logistics	Ilukkumbure et al. (2021); Lin et al. (2022); Hong & Shi (2023); Pantalona et al. (2021)
Hydrological forecasting	Compound flood models	Uncertainty not communicated operationally	Láng-Ritter et al. (2022); Refadah (2025); Yu et al. (2025); Adler et al. (2025)
Social Media & Crowdsourcing	Granular real-time data	Not integrated in operational tools	Dujardin et al. (2025); Akhtar et al. (2024); Songchon et al. (2023)
Coupled agent-institution models	Institutional behaviour modelling	Not integrated into operational DSS	Abebe et al. (2019)
Recovery-phase systems	-	No cross-phase data continuity	Park et al. (2024); Mendis et al. (2024); Birkmann et al. (2023)

Source: Own

4.2 Expected Scientific Contribution

Contribution 1 - Methodology for heterogeneous data integration in flood management. The methodology details a procedural sequence for unifying five heterogeneous source types – hydrological models, IoT sensors, satellite imagery, meteorological forecasts, and social media data – into a single interoperability layer that operational decision-makers can directly use. The methodology will address both technical interoperability (data standards, APIs, update frequencies) and semantic interoperability (ensuring shared meaning of key operational concepts such as "risk level" across agencies).

Contribution 2 - DSS framework for integrated flood management. A specification covering all three flood management phases, with defined components, data flows, and adaptation mechanisms for different stages of event development - from early warning through active crisis management to recovery coordination.

Together, these contributions address a documented gap at the intersection of information systems design and disaster management practice. The research extends the DSR tradition in information systems (Hevner et al., 2004) into the domain of multi-phase, multi-agency crisis management.

4.3 Expected Practical Contributions

The practical value of the research is direct. Emergency management organisations - civil protection services, hydrological agencies, and local authorities - currently coordinate across multiple disconnected systems. A DSS framework would:

- Enable a Common Operational Picture (COP) accessible to all involved agencies in real time;
- Reduce the capability injustice between national and local responders by providing a shared informational baseline;
- Support recovery planning by making response-phase data (damage extent, affected areas, priority needs) continuously available, without requiring manual re-aggregation.

Data uncertainty remains a poorly addressed problem in flood decision-making (Yu et al., 2025). The framework tackles this by attaching reliability and provenance markers to each integrated data stream. Decision makers can then judge both what the data shows and how much to trust it.

5 Future Development

Phases 1 and 2 – the systematic literature review and gap analysis – are scheduled for completion by early 2027.

Following the validation phase, the research opens several directions for future development:

- Empirical validation and potential implementation of the conceptual artifact in collaboration with operational flood management organizations, such as the Administration for Civil Protection and Disaster Relief (URSZR), as a direction for follow-on research beyond this doctoral project;
- Extension to other disaster types where similar data fragmentation problems exist (wildfires, earthquakes, industrial accidents), testing the generalisability of the integration methodology;
- Development of open data standards for interoperability in emergency management systems, informed by the gap taxonomy;
- Deeper integration of socioeconomic vulnerability data into recovery-phase decision support, addressing the finding that marginalized communities are consistently underrepresented in post-flood reconstruction decisions (Park et al., 2024; Mendis et al., 2024).

The research is positioned to contribute to both the academic literature on information systems design and practical improvements in how flood events are managed — reducing the informational gaps that slow response and increase damage.

AI Usage Statement

The authors used artificial intelligence tools during the preparation of this manuscript. Specifically, Claude (Anthropic) was used to support drafting, language refinement, and paraphrasing of the text. Scribbr AI Proofreader was used for grammar checking. The authors verified all AI-generated content against primary sources, conducted independent literature searches, formulated the research questions and methodology, selected and evaluated the cited works, and reviewed and revised all AI-assisted text to ensure accuracy and academic integrity. The intellectual contributions presented in this paper — including the identification of the research gap, the proposed DSS framework concept, and the research design — represent the authors' original work. The authors take full responsibility for the content of this submission.

References

Abebe, Y. A., Ghorbani, A., Nikolic, I., Vojinovic, Z., & Sanchez, A. (2019). A coupled flood-agent-institution modelling (CLAIM) framework for urban flood risk management. *Environmental Modelling & Software, 111*, 483–492. https://doi.org/10.1016/j.envsoft.2018.10.015

Adler, R., Gu, G., Zhou, N., Wu, H., Matgen, P., Galantowicz, J., Schumann, G., & Yilmaz, K. K. (2025). The role of satellite information in forecasting, modeling, and mapping the 2019 Mozambique flood. *Journal of Flood Risk Management, 18*(1), e12843. https://doi.org/10.1111/jfr3.12843

Akhtar, Z., Qazi, U., El-Sakka, A., Sadiq, R., Ofli, F., & Imran, M. (2024). Flood Insights: Integrating Remote and Social Sensing Data for Flood Exposure, Damage, and Urgent Needs Mapping. *Proceedings of the AAAI Conference on Artificial Intelligence, 38*(21), 22716–22724. https://doi.org/10.1609/aaai.v38i21.30305

Bezak, N., Panagos, P., Liakos, L., & Mikoš, M. (2023). Brief communication: A first hydrological investigation of extreme August 2023 floods in Slovenia, Europe. *Natural Hazards and Earth System Sciences, 23*(12), 3885–3893. https://doi.org/10.5194/nhess-23-3885-2023

Birkmann, J., Schüttrumpf, H., Handmer, J., Thieken, A., Kuhlicke, C., Truedinger, A., Sauter, H., Klopries, E.-M., Greiving, S., Jamshed, A., Merz, B., Solecki, W., & Kirschbauer, L. (2023). Strengthening resilience in reconstruction after extreme events — Insights from flood affected communities in Germany. *International Journal of Disaster Risk Reduction, 96*, 103965. https://doi.org/10.1016/j.ijdrr.2023.103965

Dujardin, S., Arifi, D., Schmidt, S., Linard, C., & Resch, B. (2025). Tracing online flood conversations across borders: A watershed-level analysis of geo-social media topics during the 2021 European flood. *Natural Hazards and Earth System Sciences, 25*(7), 2351–2369. https://doi.org/10.5194/nhess-25-2351-2025

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105.

Hong, J.-H., & Shi, Y.-T. (2023). Integration of Heterogeneous Sensor Systems for Disaster Responses in Smart Cities. *ISPRS International Journal of Geo-Information, 12*(7), 279. https://doi.org/10.3390/ijgi12070279

Ilukkumbure, S. P. M. K. W., Samarasiri, V. Y., Mohamed, M. F., Selvaratnam, V., & Rajapaksha, U. U. S. (2021). Early Warning for Pre- and Post Flood Risk Management by Using IoT and Machine Learning. *2021 3rd International Conference on Advancements in Computing (ICAC)*, 252–257. https://doi.org/10.1109/ICAC54203.2021.9671141

Láng-Ritter, J., Berenguer, M., Dottori, F., Kalas, M., & Sempere-Torres, D. (2022). Compound flood impact forecasting: Integrating fluvial and flash flood impact assessments into a unified system. *Hydrology and Earth System Sciences, 26*(3), 689–709. https://doi.org/10.5194/hess-26-689-2022

Lin, Q., Lin, B., Zhang, D., & Wu, J. (2022). Web-based prototype system for flood simulation and forecasting based on the HEC-HMS model. *Environmental Modelling and Software, 158*, 105541. https://doi.org/10.1016/j.envsoft.2022.105541

Mendis, K., Thayaparan, M., Kaluarachchi, Y., & Ingirige, B. (2024). Effective Stakeholder Management for Inclusive Post-Flood Management: Sri Lanka as a Case Study. *Water, 16*(10), 1429. https://doi.org/10.3390/w16101429

Pantalona, G., Tsalakanidou, F., Nikolopoulos, S., Kompatsiaris, I., Lombardo, F., Norbiato, D., Ferri, M., Kovats, L., & Haberstock, H. (2021). Decision Support System for Flood Risk Reduction Policies: The case of a flood protection measure in the area of Vicenza. *Data & Policy, 3*, e26. https://doi.org/10.1017/dap.2021.23

Park, H. J., Song, S. M., Kim, D. H., & Lee, S. O. (2024). Socioeconomic Impact on Urban Resilience against Flood Damage. *Applied Sciences, 14*(17), 7882. https://doi.org/10.3390/app14177882

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems, 24*(3), 45–77.

Refadah, S. S. (2025). Development in flood forecasting: A comprehensive review of complex and machine learning models. *Physics and Chemistry of the Earth, Parts A/B/C, 139*, 103975. https://doi.org/10.1016/j.pce.2025.103975

Schwanenberg, D., Natschke, M., Todini, E., & Reggiani, P. (2018). Scientific, technical, and institutional challenges towards next-generation operational flood risk management decision support systems. *International Journal of River Basin Management, 16*(3), 345–352.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Songchon, C., Wright, G., & Beevers, L. (2023). The use of crowdsourced social media data to improve flood forecasting. *Journal of Hydrology, 622*, 129703. https://doi.org/10.1016/j.jhydrol.2023.129703

Thomas, E. I. (2017). *Application of Open-Access and 3rd Party Geospatial Technology for Integrated Flood Risk Management* [Doctoral dissertation]. ProQuest.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, Article 38. https://doi.org/10.1145/2601248.2601268

Yu, J., Li, Y., Huang, X., & Ye, X. (2025). Data quality and uncertainty issues in flood prediction: A systematic review. *International Journal of Digital Earth, 18*(1), 2495738. https://doi.org/10.1080/17538947.2025.2495738

DOCTORAL CONSORTIUM

CLOSING THE LOOP IN AI-DRIVEN ATHLETIC TRAINING PRESCRIPTION

LUKA LAH, MIRJANA KLJAJIĆ BORŠTNAR

University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
luka.lah5@um.si, mirjana.kljajic@um.si

Training planning and load management in competitive athletics remain largely unsystematic, with most coaches and athletes relying on static tools that generate plans without any mechanism for ongoing adaptation. This open-loop limitation fails to account for the athlete's continuously evolving physiological state, increasing the risk of overtraining and injury, particularly among athletes without access to expensive professional infrastructure. This paper proposes a closed-loop decision-support architecture that integrates annual training plan generation with daily adaptation using consumer-grade wearable data and subjective athlete input. The proposed system could combine knowledge graph plan generation, Bayesian adaptation, longitudinal individual learning, and generative adversarial refinement into unified architecture. Using Design Science Research as the overarching methodological strategy, complemented by a PRISMA-ScR scoping review as its empirical foundation, the research aims to evaluate whether such a system could produce prescription quality comparable to that of experienced human coaches.

DOI
https://doi.org/10.18690/um.fov.4.2026.82

ISBN
978-961-299-152-4

Keywords:
closed-loop training system, artificial intelligence in sports, dynamic training plan adaptation, wearable device data, athlete load management

University of Maribor Press

1 Introduction

The digitalization of athletic training and the widespread availability of wearable monitoring devices have fundamentally transformed the conditions for planning and executing training processes over the past decade. Continuously measuring physiological parameters such as heart rate, heart rate variability, sleep quality, and physical activity, these devices generate extensive data streams that potentially enable highly individualized and evidence-based load management for competitive athletes. In parallel, the rapid commercial growth of training platforms based on artificial intelligence has produced broadly accessible solutions for automated training plan generation (Mateus et al., 2024; Zhou et al., 2025).

Despite this progress, training planning and load management in competitive athletics remain largely unsystematic in practice. Coaches and athletes most commonly rely on paper-based records or general-purpose productivity tools, where digital solutions are adopted, these typically consist of static template-based applications that generate a training plan from a predefined dataset without any mechanism for ongoing adaptation. This approach is insufficient because it fails to account for the athlete's evolving functional state, including fatigue, sleep quality, and acute load responses, thereby increasing the risk of overtraining, suboptimal training stimuli, and delayed responses to changes in the athlete's condition (Puce et al., 2025). The consequences of this gap are not negligible: injury rates in competitive and professional sport remain persistently high, with a substantial proportion of injuries attributable to inadequate load monitoring and insufficient recovery management (Mateus et al., 2024; Zhou et al., 2025).

While real-time feedback-driven systems do exist, they are predominantly built on Internet of Things (IoT) architectures that require expensive sensor infrastructure, dedicated communication networks, and continuous technical maintenance. Such systems are feasible within elite professional environments but remain inaccessible to most athletes who train under budget constraints, including amateur competitors, youth athletes, and recreational sport participants. This creates a structural inequality: the athletes who could benefit most from systematic load management, precisely because they lack dedicated coaching staff, are also those for whom current technology-driven solutions are practically out of reach.

The root cause of this insufficiency lies in the fragmented data landscape surrounding the athlete. Wearable device data (heart rate, sleep metrics, training load), subjective athlete reports (fatigue, pain, mood), and long-term training goals exist in isolated silos, trapped within proprietary ecosystems without standardization or integration mechanisms. As a result, coaches and athletes who interact with AI-based tools must still manually enter textual descriptions rather than connecting real-time data feeds, effectively preserving the open-loop limitation of current systems (Puce et al., 2025; Qi et al., 2024).

This practical problem is compounded by a scientific gap: the empirical literature remains methodologically heterogeneous, comparative studies are scarce, and no validated framework yet exists for dynamic, feedback-driven training plan adaptation (Qi et al., 2024; Zhao et al., 2025). Rigorous evaluation of AI-based training platforms has not kept pace with their commercial expansion, leaving their effectiveness relative to classical periodization largely untested (Mateus et al., 2024; Zhou et al., 2025). Issues of data standardization, model explainability, and governance of autonomous systems in competitive sport also remain inadequately addressed in the literature (Qi et al., 2024; Zhao et al., 2025).

This problem is theoretically grounded in cybernetic systems theory, which holds that the effectiveness of a regulatory system increases with the quality and speed of feedback (Wiener, 1948). Within athletic periodization, this aligns with the individualized training model (Bompa & Haff, 2009), which requires adjusting training loads according to the athlete's current functional status. Puce et al. (2025) explicitly identify the development of hybrid LLM and wearable approaches for dynamic adaptation as a priority for future research, directly framing the gap this paper addresses.

The proposed research therefore responds to a concrete, practical problem: the inability to reliably and continuously adapt training prescriptions based on real-time athlete data in a way that is accessible beyond high-budget professional environments. It investigates whether and how a closed-loop architecture integrating Large Language Models (LLMs) with consumer-grade wearable device data and subjective athlete feedback can serve as a viable and accessible decision support concept. The absence of such validated architecture is not the problem itself, but the proposed direction for its resolution.

This paper presents a research proposition for the design, development, and evaluation of AI closed-loop training system. The following research questions guide the investigation:

- **RQ1:** Which artificial intelligence and machine learning systems have been developed for generating athletic training plans, and what are their documented capabilities within sports science contexts?
- **RQ2:** How does the quality of training plans generated by AI systems compare to those created by experienced human coaches, and to what extent do model type, level of input detail, and athletic context influence this quality?
- **RQ3:** How effective are machine learning systems that incorporate wearable device data and subjective athlete wellness measures in predicting fatigue and optimizing load management among competitive athletes?
- **RQ4:** Does an AI system that generates an annual training plan and adapts it daily based on wearable device data and athlete feedback produce higher quality training prescriptions and more effective load management than classical non-adaptive periodization methods?

AI systems in the sports domain rest on the assumption that artificial intelligence can generate high-quality personalized training plans and effectively manage athlete loads. Within the existing scientific literature, this assumption has not been sufficiently tested (Mateus et al., 2024; Puce et al., 2025).

Qi et al., (2024); Zhao et al. (2025) point out that questions of data standardization, model explainability, and the governance of autonomous systems in competitive sports contexts receive inadequate attention in literature. Puce et al. (2025) explicitly argues that future research should prioritize reinforcement learning models or hybrid approaches combining large language models with wearable device data for dynamic training plan adaptation. Without a synthesis of available evidence, developers of decision-support solutions are left without an empirical foundation on which to base architectural decisions.

1.1 Prior research

The existing body of prior research can be organized into three mutually complementary clusters: systems for generating and adapting training plans within a closed-loop architecture, the evaluation of Large Language Model (LLM) quality in training plan generation, and machine learning systems for athlete load management.

Three prior studies directly address closed-loop architectures that combine training plan generation with dynamic sensor-driven adaptation. Using deep learning and Internet of Things (IoT) sensors, Wang et al. (2026) achieved an 18 to 23 percent improvement in athletic performance with the CoachXNet system compared to systems without individualization. Razo-Ballon et al. (2025) proposed architecture integrating IoT infrastructure with the Llama-2-7b large language model to produce daily adapted training plans based on physiological and environmental data collected from smartwatches, representing the closest existing technological design to the problem addressed in this research. By combining reinforcement learning with an inertial measurement unit (IMU), electromyography, and heart rate monitoring. Li (2025) developed a system for competitive aerobics athletes that demonstrated approximately 16 percent average improvement in training performance and athlete satisfaction exceeding 90 percent.

Several comparative studies have evaluated training plan quality produced by large language models. Genç et al. (2025) found statistically significant quality differences across ChatGPT-3.5, ChatGPT-4o, and ChatGPT-4.1, with scores increasing progressively across versions. Havers et al. (2025) had domain experts rate muscle hypertrophy and strength plans, with Google Gemini receiving the highest ratings, followed by Copilot and GPT-3.5. J. Yang et al. (2026) evaluated six large language models on soccer training plan quality, with ChatGPT-o1 scoring highest among expert evaluators. Washif et al. (2024) benchmarked three ChatGPT versions against established resistance training guidelines, concluding that generated plans are realistic but frequently require adjustment and are best treated as coach templates. Düking et al. (2024) showed that input detail significantly influences output quality in runner-targeted plans, with the least detailed input receiving significantly lower expert ratings on 15 of 22 criteria compared to the most detailed. Z. Yang et al. (2026) corroborated this finding, demonstrating that detailed input reduces variability and raises overall quality across dimensions of personal adaptability,

safety, and feasibility. Tan & Chen (2025) reported that a GAN-based approach achieved a 22 percent reduction in MSE, and 45 percent faster generation compared to classical machine learning, with higher subjective ratings from athletes and coaches. Masagca (2024) confirmed in a quasi-experimental study that a five-week AI-generated calisthenics program produced significant improvements in flexibility and muscular endurance, though cardiovascular outcomes remained inferior to those achieved through professionally designed programs.

Empirical studies on machine learning-supported load management and physiological monitoring form a closely related body of evidence. Simonelli et al. (2026) achieved subjective fatigue prediction accuracy of 70 to 82 percent in professional soccer players using four machine learning models combined with wearable sensor data. Carvalho et al. (2025) demonstrated through explainable artificial intelligence that heart rate variability effectively tracks daily load changes in high-performance swimmers, with two predictive models performing well at a 27 percent increase in training load and a 20 percent increase in perceived effort. Teixeira et al. (2024) found that maturation offset was a statistically significant but weak predictor of high-intensity actions in youth soccer, with all tested algorithms showing poor performance when used as the sole predictor. Mandorino et al. (2024), in a two-season longitudinal study with 106 professional soccer players, showed that large weekly load decreases improved locomotor efficiency in less fit athletes, while both large decreases and large increases were detrimental in well-conditioned players. Hao & Qian (2024) proposed a support vector machine-based performance evaluation model designed to provide multi-dimensional decision support for training planning.

2 Proposed methodology and results

To achieve the objectives of the proposed research, we will adopt Design Science Research (DSR) as the overarching methodological strategy (figure 2). Hevner et al. (2004) and Peffers et al. (2007) recommend DSR as the primary strategy for developing and evaluating novel information artifacts within the context of information systems research. The framework is methodologically appropriate here because the proposed research does not merely describe or explain an existing phenomenon. Rather, it seeks to design, construct, and evaluate a new technological artifact that addresses a clearly identified gap in scientific literature.

Several conceptual directions could plausibly address the identified problem of personalized sport-specific training prescription. One direction could involve rule-based expert systems, in which domain knowledge from sports scientists and coaches is encoded as explicit logical rules. Another direction could rely on collaborative filtering, recommending training content based on patterns observed across large populations of athletes with similar profiles. A third direction could center on retrieval-augmented generation, pairing data mining approaches with structured databases of validated training plans. Each of these approaches carries distinct trade-offs in terms of explainability, data requirements, and adaptability to individual athletes. The model proposed in this research draws from the strengths of each direction while addressing their individual limitations through a multi-phase architecture that combines semantic embedding, knowledge graph construction, Bayesian adaptation, and generative refinement.

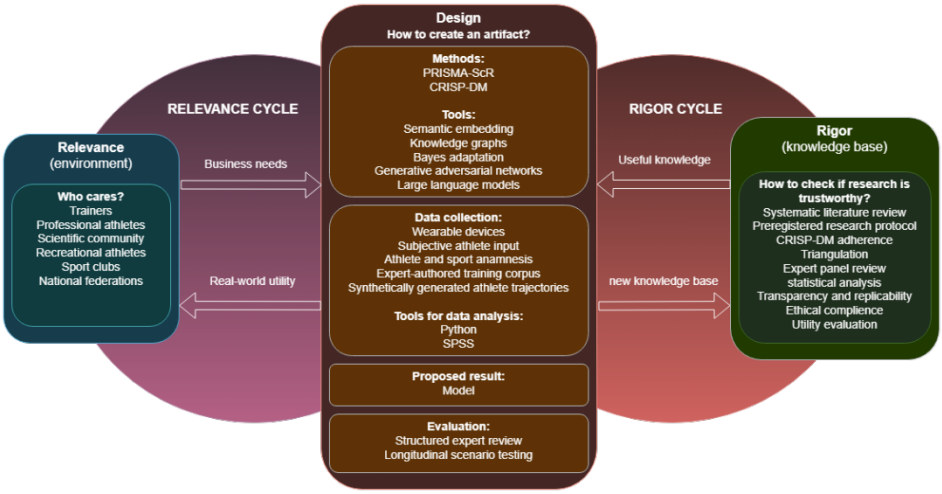

Figure 2: Design Science Research Framework applied to the closed-loop AI training system (adapted from Hevner et al., 2004)

Source: Own

The DSR framework is complemented by a systematic scoping review conducted in accordance with the PRISMA-ScR protocol (Tricco et al., 2018). This review serves as the foundational step of the research, ensuring that the artifact is grounded in the best available evidence regarding existing methodologies, data sources, and algorithmic approaches in the field. Within the Design cycle, the development of all

machine learning components will follow the CRISP-DM process model. This means the development proceeds through the stages of business understanding, data understanding, data preparation, modeling, evaluation, and deployment in an iterative manner. DSR therefore functions as overarching research logic, while CRISP-DM provides the structured procedural scaffold governing each machine learning development iteration within that broader framework.

Expected contributions and evaluation

Personalized conditioning training presents a multi-dimensional optimization problem that resists straightforward algorithmic solutions. The core difficulty lies in simultaneously representing the complexity of individual athlete profiles and encoding the structured demands of diverse sports in a form that is both computationally tractable and scientifically grounded. Existing approaches each address a part of this challenge: rule-based expert systems struggle to scale and capture individual variability, collaborative filtering is limited by cold-start problems and cross sport generalization failure, and retrieval-augmented generation raises concerns about prescriptive consistency and traceability (Tan & Chen, 2025; Washif et al., 2024). No existing system has addressed these limitations as an integrated architecture.

The proposed model could respond to this gap through a closed-loop architecture combining four computational mechanisms: knowledge-graph-based plan generation, Bayesian closed-loop adaptation, longitudinal individual learning, and generative adversarial refinement. The system would address exclusively the conditioning dimension of athletic preparation, deliberately excluding technique-oriented content so that technical coaching authority remains with human practitioners.

A foundational challenge concerns the computational representation of athlete identity and sport structure. An athlete profile could encode medical history, injury records, prior performance assessments, training age, competition level, and contextual constraints such as equipment access. A sport profile could encode dominant movement patterns, energy system demands, and competitive calendar structure. Both representations could be independently embedded into high-dimensional vectors and blended into a unified athlete-sport fingerprint, whose

weighting between sport-general and athlete-specific dimensions could function as an adjustable parameter reflecting athlete experience level. The model could then construct a knowledge graph in which nodes represent exercises and athlete and sport characteristics, with edges encoding biomechanical similarity and prescriptive relationships derived from a corpus of expert-authored plans. This formulation would enable relevance-scored exercise selection and structured plan assembly according to established periodization principles (Bompa & Haff, 2009).

A second challenge concerns temporal validity. Any plan generated at a single point in time will progressively diverge from an athlete's actual physiological state, as fatigue, recovery, and adaptation continuously alter the appropriate training stimulus (Carvalho et al., 2025; Simonelli et al., 2026). The proposed adaptation mechanism could model this as a Bayesian inference problem, in which the system maintains and continuously updates a probabilistic belief over prescription parameters. At each observation interval, this belief could be updated using objective wearable signals such as heart rate variability, sleep quality, and activity load, together with a subjective wellness rating from the athlete. Adjustable parameters could include training volume, intensity, exercise selection, and recovery duration, all constrained within the periodization structure established during plan generation.

A third challenge is that population-level models cannot account for pronounced inter-individual variability in physiological response. Two athletes with identical intake profiles may respond very differently to the same prescription, and this divergence compounds over time (Mandorino et al., 2024). The proposed longitudinal learning component could address this by progressively shifting the model's generative assumptions away from corpus-derived population priors and toward the specific response history of the individual. In early use, the model would rely primarily on population patterns; as personal response data accumulates, it could assign progressively greater weight to that individual's own trajectory. The rate and conditions under which individual data should override population priors represent an open methodological challenge that would constitute an original contribution of this research. Initial development and stress-testing of this component could rely on synthetically generated athlete trajectories produced by a generative adversarial network conditioned on corpus-level patterns, before iterative retraining on real-world observations.

A fourth challenge concerns the structural coherence of generated output. Even when individual exercise parameters are appropriate, a computationally assembled plan may lack the contextual flow and compositional logic characteristic of expert-authored programs, a limitation that Bayesian parameter optimization alone cannot resolve (Tan & Chen, 2025). A generative adversarial component introduced as the final architectural layer could address this: a generator would combine the Bayesian adaptation output with the athlete-sport embedding to produce candidate plans, while a discriminator would evaluate whether each candidate is structurally indistinguishable from expert-authored programs. Iterative adversarial training could progressively refine the generator's outputs toward expert quality, with the Wasserstein formulation with gradient penalty applied to improve training stability and output diversity.

Evaluating a decision-support model of this nature presents challenges not fully addressed by conventional software testing, since prescription quality is not reducible to a single metric and expert judgment remains an indispensable reference point. A two-stage evaluation strategy is proposed within the DSR evaluation cycle (Hevner et al., 2004). In the first stage, domain experts could assess model outputs against structured criteria derived from sports science literature, covering progressive overload adherence, periodization integrity, sport-specific appropriateness, and practical feasibility. In the second stage, the dynamic behavior of the adaptation mechanism could be examined across representative longitudinal scenarios defined by distinct combinations of wearable biometric trajectories and subjective wellness ratings, testing whether the Bayesian update mechanism produces responses that are both directionally correct and proportionate over time. Semi-structured interviews with coaches and sports scientists could provide qualitative triangulation alongside quantitative performance metrics.

The primary anticipated contribution of this research is a functioning decision-support artifact integrating the described mechanisms into a unified closed-loop architecture (Hevner et al., 2004; Peffers et al., 2007). The scientific contribution lies not only in the artifact itself but in the knowledge design generated through its construction and evaluation, as the proposed combination of knowledge graph generation, Bayesian adaptation, longitudinal learning, and generative refinement has not been examined as an integrated system in the existing literature. The research could yield a meaningful contribution regardless of empirical outcome. If expert

evaluators judge generated prescriptions as comparable to coach-authored plans, this would constitute evidence that architecture is a viable computational approach to a problem previously considered to require exclusively human expertise. If generated plans prove inferior, the evaluation process could identify the specific architectural or data-related limitations responsible, providing a theoretically grounded foundation for future development.

References

Bompa, T. O., & Haff, G. (2009). *Periodization: Theory and Methodology of Training*. Human Kinetics.

Carvalho, D. D., Goethel, M. F., Erblang, M., Vilas-Boas, J. P., Pyne, D. B., Fernandes, R. J., & Lopes, P. (2025). Impact of an Overload Period on Heart Rate Variability, Sleep Quality, Motivation, and Performance in High-level Swimmers: Use of Explainable Artificial Intelligence (XAI) to Assess Training Load Variations. *Sports Medicine*. https://doi.org/10.1007/s40279-025-02369-1

Düking, P., Sperlich, B., Voigt, L., Hooren, B. V., Zanini, M., & Zinner, C. (2024). ChatGPT Generated Training Plans for Runners are not Rated Optimal by Coaching Experts, but Increase in Quality with Additional Input Information. *Journal of Sports Science and Medicine*, 56–72. https://doi.org/10.52082/jssm.2024.56

Genç, A., Aydın, G. R., Kasap, M., Özkan, A., Rakhymzhanov, A., Gürcan, H. H., Güllü, S., & Demirhan, B. (2025). Comparative evaluation of ChatGPT versions in training program design: Scientific approach, accuracy, and practical applicability. *BMC Sports Science, Medicine and Rehabilitation*, *18*(1), 19. https://doi.org/10.1186/s13102-025-01409-7

Hao, P., & Qian, K. (2024). The Integration of Personalized Training Program Design and Information Technology for Athletes. *Scalable Computing: Practice and Experience*, *25*(5), 4351–4359. https://doi.org/10.12694/scpe.v25i5.3083

Havers, T., Jelonnek, C., Masur, L., Isenmann, E., Sperlich, B., Geisler, S., & Düking, P. (2025). A professional assessment of training plans for muscle hypertrophy and maximal strength developed by generative artificial intelligence. *Biology of Sport*, *42*(4), 353–380. https://doi.org/10.5114/biolsport.2026.152350

Hevner, A., R, A., March, S., T, S., Park, Park, J., Ram, & Sudha. (2004). Design Science in Information Systems Research. *Management Information Systems Quarterly*, *28*, 75.

Li, H. (2025). Intelligent IoT-based training decision and evaluation system for competitive aerobics athletes. *Discover Computing*, *28*(1), 182. https://doi.org/10.1007/s10791-025-09709-9

Mandorino, M., Tessitore, A., & Lacome, M. (2024). Loading or Unloading? This Is the Question! A Multi-Season Study in Professional Football Players. *Sports*, *12*(6), 148. https://doi.org/10.3390/sports12060148

Masagca, R. C. (2024). The AI coach: A 5-week AI-generated calisthenics training program on health-related physical fitness components of untrained collegiate students. *Journal of Human Sport and Exercise*, *20*, 39–56. https://doi.org/10.55860/13v7e679

Mateus, N., Abade, E., Coutinho, D., Gómez, M.-Á., Peñas, C. L., & Sampaio, J. (2024). Empowering the Sports Scientist with Artificial Intelligence in Training, Performance, and Health Management. *Sensors*, *25*(1), 139. https://doi.org/10.3390/s25010139

Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*, 45–77.

Puce, L., Bragazzi, N. L., Currà, A., & Trompetto, C. (2025). Harnessing Generative Artificial Intelligence for Exercise and Training Prescription: Applications and Implications in Sports

and Physical Activity—A Systematic Literature Review. *Applied Sciences*, *15*(7), 3497. https://doi.org/10.3390/app15073497

Qi, Y., Sajadi, S. M., Baghaei, S., Rezaei, R., & Li, W. (2024). Digital technologies in sports: Opportunities, challenges, and strategies for safeguarding athlete wellbeing and competitive integrity in the digital era. *Technology in Society*, *77*, 102496. https://doi.org/10.1016/j.techsoc.2024.102496

Razo-Ballon, H., Ticona-Esquivel, R., & Montalvo, P. (2025). Design of IoT Architecture and LLM Model for Personalized Training Recommendations for Athletes. In T. Guarda, F. Portela, & G. Gatica (Eds), *Advanced Research in Technologies, Information, Innovation and Sustainability* (pp. 306–320). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-83210-9_23

Simonelli, C., Trecroci, A., Formenti, D., & Rossi, A. (2026). Prediction of Subjective Fatigue in Professional Soccer Players: A Data-Driven Method to Optimize Training Approach to the Match. *Research Quarterly for Exercise and Sport*, *97*(1), 137–145. https://doi.org/10.1080/02701367.2025.2560562

Tan, J., & Chen, J. (2025). Generating context-specific sports training plans by combining generative adversarial networks. *PLOS ONE*, *20*(1), e0318321. https://doi.org/10.1371/journal.pone.0318321

Teixeira, J. E., Encarnação, S., Branquinho, L., Morgans, R., Afonso, P., Rocha, J., Graça, F., Barbosa, T. M., Monteiro, A. M., Ferraz, R., & Forte, P. (2024). Data Mining Paths for Standard Weekly Training Load in Sub-Elite Young Football Players: A Machine Learning Approach. *Journal of Functional Morphology and Kinesiology*, *9*(3), 114. https://doi.org/10.3390/jfmk9030114

Tricco, A., Lillie, E., Zarin, W., O'Brien, K., Colquhoun, H., Levac, D., Moher, D., Peters, M., Weeks, L., Hempel, S., Akl, E., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M., Garritty, C., & Straus, S. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*. https://doi.org/10.7326/M18-0850

Wang, Y., Feng, X., Wu, Y., Zhang, X., & Wu, J. (2026). CoachXNet: An Artificial Intelligence and Internet of Things Integrated Platform for Personalized Training and Feedback in Digital Sports. *International Journal of Computational Intelligence Systems*, *19*(1), 49. https://doi.org/10.1007/s44196-025-01146-2

Washif, J., Pagaduan, J., James, C., Dergaa, I., & Beaven, C. (2024). Artificial intelligence in sport: Exploring the potential of usingChatGPT in resistance training prescription. *Biology of Sport*, *41*(2), 209–220. https://doi.org/10.5114/biolsport.2024.132987

Wiener, N. (1948). Cybernetics. *Scientific American*, *179*(5), 14–19.

Yang, J., Qin, S., & Ren, D. (2026). Artificial intelligence coaches on the sidelines: Evaluating readability and quality of soccer training plans from six generative models. *International Journal of Sports Science & Coaching*, *21*(1), 75–81. https://doi.org/10.1177/17479541251369593

Yang, Z., Zhang, X., Li, H., & Ye, J. (2026). More details, less variability? A crossover design study on the impact of information granularity on ChatGPT's training program stability. *Biology of Sport*. https://doi.org/10.5114/biolsport.2026.154148

Zhao, X., Zheng, H., & Zhang, Q. (2025). Modern Advances in Artificial Intelligence Across the Athletic Domain. *Concurrency and Computation: Practice and Experience*, *37*. https://doi.org/10.1002/cpe.70068

Zhou, D., Keogh, J. W. L., Ma, Y., Tong, R. K. Y., Khan, A. R., & Jennings, N. R. (2025). Artificial intelligence in sport: A narrative review of applications, challenges and future trends. *Journal of Sports Sciences*, 1–16. https://doi.org/10.1080/02640414.2025.2518694

UNI
BIZ
TECH
GOV

UNI
BIZ
TECH
GOV

39TH BLED ECONFERENCE CO-CREATING HUMAN-CENTRED AND RESPONSIBLE DIGITAL FUTURES

DOROTEJA VIDMAR (ED. ET AL.)
University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
doroteja.vidmar@um.si

The Bled eConference, organized by the University of Maribor's Faculty of Organizational Sciences, has been shaping the development of electronic, now digital business and society since 1988. The theme of the 39th conference is "Co-creating Human-Centred and Responsible Digital Futures." We emphasize the importance to include diverse groups in digital development, whose creators understand and accept responsibility in shaping a future for broader society based on ethics, inclusivity, and sustainability. Digital technologies, as a means of transforming systems and solutions, should be focused on long-term well-being of people. We encourage innovation, resilience, social and economic progress. The papers of this conference address various topics, including opportunities and challenges of twin transition, new digital technologies, artificial intelligence and data science, decision analytics for business and social change, digital business models, jobs of the future, digital health, digital ethics, digital education, smart and sustainable cities, digital consumers, and digital transformation of the public sector.

DOI
https://doi.org/10.18690/um.fov.4.2026

ISBN
978-961-299-152-4

Keywords:
resilience, digital transformation, digital innovation, twin transition, sustainability, digital transition, green transition, designing digital future

University of Maribor Press

University of Maribor

Faculty of Organizational Sciences

UNI BIZ TECH GOV