



38TH BLED eCONFERENCE

EMPOWERING TRANSFORMATION: SHAPING DIGITAL FUTURES FOR ALL

**JUNE 8 - 11, 2025, BLED, SLOVENIA
CONFERENCE PROCEEDINGS**

**Andreja PUCIHAR
Mirjana KLJAJIČ BORŠTNAR
Staša BLATNIK
Marjeta MAROLT**

**BONS Roger W. H.
SMIT Koen
GLOWATZ Matthias**

EDITORS



University of Maribor Press





University of Maribor

Faculty of Organizational Sciences

38th Bled eConference

Empowering Transformation: Shaping Digital Futures for All

Conference Proceedings

Editors

Andreja Pucihar

Mirjana Kljajić Borštnar

Staša Blatnik

Marjeta Marolt

Roger W. H. Bons

Koen Smit

Matthias Glowatz

June 2025

Title	38th Bled eConference
Subtitle	Empowering Transformation: Shaping Digital Futures for All: Conference Proceedings
Editors	Andreja Pucihar (University of Maribor, Faculty of Organizational Sciences, Slovenia) Mirjana Kljajić Borštnar (University of Maribor, Faculty of Organizational Sciences, Slovenia) Staša Blatnik (University of Maribor, Faculty of Organizational Sciences, Slovenia) Marjeta Marolt (University of Maribor, Faculty of Organizational Sciences, Slovenia) Roger W. H. Bons (Open University; Heerlen; the Netherlands) Koen Smit (HU University of Applied Sciences Utrecht, Utrecht, the Netherlands) Matthias Glowatz (University College Dublin, Ireland)
Review	Alenka Baggia (University of Maribor, Slovenia), Samaneh Bagheri (Open University of the Netherlands, the Netherlands), Gulet Barre (Open University of the Netherlands, the Netherlands), Matthijs Berkhout (HU University of Applied Sciences Utrecht, the Netherlands), Roland Bijvank (HU University of Applied Sciences Utrecht, the Netherlands), Staša Blatnik (University of Maribor, Slovenia), Josien Boetje (HU University of Applied Sciences Utrecht, the Netherlands), Florian Bongard (Frankfurt School of Finance & Management, Germany), Martin Brehmer (Augsburg University, Germany), Serhat Çaliskan (Hogeschool Utrecht, the Netherlands), Mehmet Cem Bölen (Ataturk University, Turkey), Helen Cripps (Edith Cowan University, Australia), Nejc Čelik (University of Maribor, Slovenia), Thu Ha Dang (La Trobe University, Australia), Marja Exalto-Sijbrands (HU University of Applied Sciences, the Netherlands), Matt Glowatz (University College Dublin, Ireland), Mirjam de Haas (HU University of Applied Sciences Utrecht, the Netherlands), Nerko Hadziarapovic (HU University of Applied Sciences Utrecht and Open University of the Netherlands, the Netherlands), Tobias Hassmann (Munich Business School, Germany), Biao He (Tartu university, Estonia), Marikka Heikkilä (University of Turku, Finland), Tim Huygh (Open University, the Netherlands), Tina Jukić (University of Ljubljana, Slovenia), Jürgen Karla (University of Applied Sciences, Germany), Ander de Keijzer (Avans Hogeschool, the Netherlands), Štefan Kohek (University of Maribor, Slovenia), Simon Kolmanič (University of Maribor, Slovenia), Robert-Jan Korteland (HU University of Applied Sciences Utrecht, the Netherlands), Sam Leewis (HU University of Applied Sciences Utrecht, the Netherlands), Pasi Luukka (LUT University, Finland), Marjeta Marolt (University of Maribor, Slovenia), Navin Sewberath Misser (HU University of Applied Sciences Utrecht, the Netherlands), Andrei Munteanu (University of Qubeck at Montreal, Canada), Ayesha Nilashani (La Trobe University, Australia), Guido Ongena (HU University of Applied Sciences Utrecht and Open University of the Netherlands, the Netherlands), Rob Peters (Province of Utrecht, Utrecht, the Netherlands), Andreja Pucihar (University of Maribor, Slovenia), Christopher Reichstein (Duale Hochschule Baden-Württemberg, Germany), Alexander Ruh (University of Cooperate Education, Germany), Zaid Sako (La Trobe University, Australia),

Nadeem Shuakat (La Trobe University, Australia), Bo Sichterman (HU University of Applied Sciences Utrecht, the Netherlands), Damjan Strnad (University of Maribor, Slovenia), Andrej Škraba (University of Maribor, Slovenia), Nevena Toporova (Technical University of Munich, Germany), Nalika Ulapane (La Trobe University, Australia), Nadia Bij de Vaate (Hogeschool Utrecht, the Netherlands), Astrid van Barneveld (Princess Máxima Center for Pediatric Oncology, the Netherlands), Suzan van Brussel (Avans Hogeschool, the Netherlands), Annemae van de Hoef (HU University of Applied Sciences Utrecht, the Netherlands), Christel van de Wal (Open University of the Netherlands, the Netherlands), Steijn van der Craats (HU University of Applied Sciences Utrecht, the Netherlands), Emanuel van Dongen (Utrecht University, the Netherlands), Stan van Ginkel (University of Applied Sciences Utrecht, the Netherlands), Jeroen van Grondelle (HU University of Applied Sciences Utrecht, the Netherlands), Thijs van Houwelingen (HU University of Applied Sciences Utrecht, the Netherlands), John van Meerten (HU University of Applied Sciences, the Netherlands), Johan Versendaal (HU University of Applied Sciences Utrecht and Open University of the Netherlands, the Netherlands), Doroteja Vidmar (RLS, Slovenia), Pirkko Walden (Åbo Akademi University, Finland), Thorsten Weber (FOM Hochschule, Germany), Simon Weisskopf (University of Trento, Switzerland), Luc van der Zandt (HU University of Applied Sciences Utrecht, the Netherlands)

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

Jan Perša
(University of Maribor, University Press)

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

Graphics material Sources are own unless otherwise noted.
Authors and Pucihar, Borštnar Kljajić, Blatnik, Marolt, Bons, Smit, Glowatz (editors), 2025

Conference 38th Bled eConference – Empowering Transformation: Shaping Digital Futures for All

Location and date Bled, Slovenia, June 8 – 11, 2025

Programme committee Art Alishani (Tartu University, Estonia), Matthias Baldauf (Eastern Switzerland University of Applied Sciences, Switzerland), Staša Blatnik (University of Maribor, Faculty of organizational sciences, Slovenia), Roger Bons (Open Universiteit, the Netherlands), Christer Carlsson (Institute for Advanced Management Systems Research, Åbo Akademi University, Finland), Helen Cripps (Edith cowan University, Australia), Matt Glowatz (University college Dublin, Ireland), Nerko Hadziarapovic (HU University of Applied Sciences Utrecht, the Netherlands), Tina Jukić (University of Ljubljana, Faculty of Public Administration, Slovenia), Mirjana Kljajić Borštnar (University of Maribor, Faculty of organizational sciences, Slovenia), Sam Leewis (HU University of Applied Sciences Utrecht, the Netherlands), Xander Lub (HU University of Applied Sciences Utrecht, the Netherlands), Venkata Marella (Open University of the Netherlands, the Netherlands), Marjeta Marolt (University of Maribor, Faculty of organizational sciences, Slovenia), Axel Minten (FOM University of Applied Sciences, Germany), Domen Mongus (UM-FERI, Slovenia), Guido Ongena (HU University of Applied Sciences Utrecht, the Netherlands), Marijn Plomp (Vrije Universiteit Amsterdam, the Netherlands), Andreja Pucihar (University of Maribor, Faculty of organizational sciences, Slovenia), Uroš Rajković (University of Maribor, Faculty of organizational sciences, Slovenia), Pascal Ravesteijn (Utrecht University of Applied Sciences, the Netherlands), Juergen Seitz (DHBW Heidenheim, Germany), Anand Sheombar (Manchester Metropolitan University Business School & Utrecht University of Applied Sciences., the Netherlands),

Koen Smit (Utrecht University, the Netherlands), Esther van der Stappen (Avans University of Applied Sciences, the Netherlands), Stan van Ginkel (Hogeschool Utrecht, the Netherlands), Johan Versendaal (HU University of Applied Sciences Utrecht, the Netherlands), Tara Vester (HU University of Applied Sciences Utrecht, the Netherlands), Doug Vogel (Harbin Institute of Technology, China), Nilmini Wickramasinghe (Swinburne University of Technology & Epworth HealthCare, Australia), Hans-Dieter Zimmermann (OST Eastern Switzerland University of University of Applied Sciences, St. Gallen, Switzerland)

Organizing committee	Staša Blatnik (University of Maribor, Faculty of Organizational Sciences), Petra Gorjanc (University of Maribor, Faculty of Organizational Sciences), Mirjana Kljajić Borštnar (University of Maribor, Faculty of Organizational Sciences), Klara Knific (University of Maribor, Faculty of Organizational Sciences), Maruša Luštrik (University of Maribor, Faculty of Organizational Sciences), Aljaž Murko (University of Maribor, Faculty of Organizational Sciences), Andreja Pucihar (University of Maribor, Faculty of Organizational Sciences)
Published by	University of Maribor, University Press Slomškov trg 15, 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	1 ST
Publication type	E-book
Available at	https://press.um.si/index.php/ump/catalog/book/947
Published	Maribor, Slovenia, June 2025



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

Text © The authors & Pucihar, Borštnar Kljajić, Blatnik, Marolt, Bons, Smit, Glowatz (editors), 2025

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 Empowering Transformation (38 ; 2025 ; Bled)
38th Bled eConference Empowering Transformation: Shaping Digital Futures for All [Elektronski vir] : conference proceedings : [Bled, Slovenia, June 8 - 11, 2025] / editors Andreja Pucihar ... [et al.]. - 1st ed. - E-zbornik. - Maribor : University of Maribor, University Press, 2025

Način dostopa (URL): <https://press.um.si/index.php/ump/catalog/book/947>

ISBN 978-961-286-998-4 (PDF)

doi: 10.18690/um.fov.4.2025

COBISS.SI-ID 238031107

ISBN 978-961-286-998-4 (pdf)

DOI <https://doi.org/10.18690/um.fov.4.2025>

Price Free copy

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

Attribution Pucihar, A., Borštnar Kljajić, M., Blatnik, S., Marolt, M., Bons, R. W. H., Smit, K., Glowatz, M.. (eds.) (2025). *38th Bled eConference Empowering Transformation: Shaping Digital Futures for All: Conference Proceedings*. University of Maribor, University Press. doi: 10.18690/um.fov.4.2025

Table of Contents

1	Towards a Taxonomy for Digital Assistant Technologies: Addressing the Jingle-Jangle Fallacies <i>Niklas Preiß, Markus Westner</i>	1
2	Bridging Cyber Resilience and IT Business Value: A Longitudinal Bibliometric Analysis <i>Ziggy Van Giel, Tim Huygh, Anant Joshi, Steven De Haes</i>	21
3	Digitalization as a Catalyst for Resilience in the Hospitality Sector <i>Alinda Kokkinou, Michel Altan, Abdon Dantas, Ondrej Mitas</i>	37
4	Framing Knowledge: Enhancing Communication Strategies in the Dutch Second-hand Clothing Industry <i>Mats Belgers, Marja Exalto-Sijbrands, Michèlle Bal, Pascal Ravesteijn</i>	53
5	Where is Emotion? Comparing Emotion Mapping Technologies <i>Ondrej Mitas, Tamara Surla, Ben Wielenga, Jelena Farkić, Peter Ward, Alinda Kokkinou, Sandy Thomas</i>	67
6	Business Benefits and Customer Perceptions in Customer Service Chatbots <i>Jason van Hamond, Denys Bespalko, Margot Neggers, Alinda Kokkinou</i>	85
7	Aligning Chatbot Features with Industry Needs: A Comparative Study of User Expectations in Utilitarian and Hedonic Sectors for SMEs <i>Denys Bespalko, Jason van Hamond, Margot Neggers, Alinda Kokkinou</i>	101
8	Using Video Segmentation for Unsupervised Industrial Optical Production Control – Case Study for the SmartFactory@OST <i>Valmir Bekiri, Stefan Stöckler</i>	115
9	Assessing the Impact of Chatbots on Customer Service Efficiency: A Comparative Study with Human Agents <i>Fatima Bilal, József Mezei, Bad re Alam</i>	129

10	Explainable Digital Twins of Patients: Towards Precision and Personalisation through Cohort Matching Nilmini Wickramasinghe, Nalika Ulapane	151
11	Technological Frames and Learning Experiences: Students' Interpretations of Rule-Based Chatbots in Higher Education Merja Alanko-Turunen, Olli-Jaakko Kupiainen, Mari Raaska	165
12	Value Co-creation Through Digitalisation: From a Micro-level to a Meso-level Approach Yiran Chen Zhou, Anastasia Tsvetkova	183
13	"Dear Applicant, the AI Will See you Now!" Job Applicants' Reactions to AI-Enhanced Selection Floris Verheijen, Xander Lub, Ran Zhang	197
14	Preventing Unauthorized Use of Personal Data in Ghostbots Post-Mortem: An Access and Authorization Model for Digital Wills Jasmin Bödecker, Jürgen Karla	213
15	Analyzing the Determinants of Healthcare Technology Adoption Using the Task-Technology Fit (TTF) Model: A Systematic Review and Meta-Analysis Ayesha Thantrige, Nilmini Wickramasinghe	229
16	Exploring the Maturity of Smart Public Governance: Evidence from Two Policy Sectors in Slovenia Petra Vujković, Tina Jukić	253
17	Advanced RAG-LLM Prototype AI on PubMed for Cardiac Health Luuk P.A. Simons, Pradeep K. Murukannaiah, Budi S. Han, Mark A. Neerincx	265
18	Shaping Participation: How Digital Platform Design Influences User Participation Gabriel März	281
19	The Construction of Soft Skills Scale Targeted Generation Z For IT Sector Orkun Yildiz, Tuğba Elif Toprak-Yildiz	305
20	Barriers to the Adoption of Digital Health Solutions Among Type 2 Diabetic Patients During the COVID-19 Pandemic Ayesha Thantrige, Thu Ha Dang, Nilmini Wickramasinghe	319

21	AI's Role in Collaborative Learning Settings in Higher Education – A Systematic Review Henna Pernu, Sanna Paananen, Satu Aksovaara, Minna Silvennoinen	335
22	“Rage Against the Machine?”: The Impact of Clinical Decision Support Systems on Hospital Nursing Decision-Making, Workflow Efficiency, and Patient Outcomes: A Rapid Review Matthijs Berkhout, Astrid van Barneveld, Koen Smit, Thijs van Houwelingen	351
23	AI-Based Business Model Innovation: A ChatGPT Integration Approach Florian Bongard, Jürgen Moormann	367
24	Virtual Care Centers and the Evolving Role of Nurses Henk Sonneveld, Apphia Tan, Jobbe P.L. Leenen, Thijs van Houwelingen	385
25	The Cost-Benefit Analysis of Digital Timestamping in Maritime Logistics: The Case of Aikatieto Jukka Heikkilä, Marikka Heikkilä, Manju Kc Deuja	399
26	Green(ing) of Transport Chains: Implications to Business Models Marikka Heikkilä, Kristel Edelman, Riikka Franzén	417
27	Digitalization Meets Circular Economy Hans-Dieter Zimmermann	437
28	Developing DigiWish: The Co-Creation of a Visual Strengths-Based Training Tool for Digital Skills Dian A. de Vries, Anders Kropveld, Yurrian Rozenberg, Jannie Guo	453
29	Students' Usage of GenAI in Universities of Applied Sciences: Experiences and Development Needs for Guidance and Support Minna Silvennoinen, Satu Aksovaara, Merja Alanko-Turunen	467
30	LLM Pipeline for Mapping Heterogeneous Data: A Case Study in Food Classification Kevin Nils Röhl, Rainer Alt, Jan Wirsam	483
31	Developing a Value Sensitive VR Learning Environment Fostering Professionals' Communication Skills Olivier van der Molen, Emanuel van Dongen, Stan van Ginkel	499
32	Energy Behavior and Smart Home Systems: A Bibliometric Analysis Xiao Peng, Xiuzhu Zang, Mengqiu Deng	517

33	Breaking the ICE? Examining the Cognitive, Social, and Psychological Determinants of User Resistance to Electric Non-Road Work Vehicle Adoption Henri Jalo, Marko Seppänen	533
34	Generational Research on Technology Acceptance and Use: What's Popular Isn't Always Right Anna Sell, Markus Makkonen, Pirkko Walden	549
35	Towards the Automatic Extraction of Decision Model & Notation from Dutch Legal Text Annemae van de Hoef, Sam Leewis, Koen Smit	567
36	Better Together: Conditions for Implementing Innovations to Improve Quality of Care in Health Delivery Organizations Pieter Kievit, Roger W.H. Bons, Nadine Roijackers	585
37	An Analysis of Scientific Requirements on Artificial Intelligence Governance Marius Goebel, Roger W. H. Bons, Rüdiger Buchkremer	601
38	Building Green Generative AI: An Ecosystem-Wide Approach to Environmental Sustainability Fabian Helms, Kay Hönemann, Manuel Wiesche	617
39	Causal Effects of Addressing and PSI on Compliance with a Social Media Influencer's Recommendations Jussi Nyrhinen, Eero Rantala, Jesse Tuominen, Markus Makkonen, Terhi-Anna Wilska, Lauri Frank	633
40	From Simple Sandbox Process to Regulatory Sandbox Framework: Serving the Dual Objectives of AI Regulation Jukka Heikkilä, Altti Lagstedt, Seppo Heikura, Eeva Kaakkurivaara, Anastasiia Dardykina, Alice Teilhard de Chardin	643
41	Data-Driven Ecosystem Business Models in Agriculture with Focus on Sustainability: A Systematic Literature Review Shahin Rezvanian	661
42	Advancements in News Recommendation Systems: The Role and Impact of Artificial Intelligence and Large Language Models Ceyhun Kazanç, Nazım Taşkın	677

RESEARCH IN PROGRESS

43	Exploring the Relationship Between Personal Values and Perceived Obsolescence of High-Tech Products Mehmet Cem Bölen, Burak Borulu, Fulya Acikgoz	705
44	A Systematic Approach to Define Digital Well-Being at Work Nadia Bij de Vaate, Dian de Vries, Machiel Bouwmans	717
45	Green IS – External Pressures and IS Executives Tobias Hassmann	727
46	Towards Understanding Cognitive Biases in Cybersecurity Governance Gulet Barre, Tim Huygh, Dinh Khoi Nguyen, Arno Nuijten	737
47	Exploring the Feasibility of Generative AI in Enhancing the Identification of Spatial and Regulatory Opportunities Using Urban Digital Twins Rob Peters, Fabian Kok, John van Meerten, Koen Smit	745
48	Using Digital Twin Technology for Developing a Healthy and Sustainable Living Environment: A Dutch Case Study Serhat Çaliskan, Sam Leewis, Koen Smit	755
49	Bridging the Digital Divide for Older Adults: An Investigation of Barriers and Facilitators Milena Head, Khaled Hassanein	765

DOCTORAL CONSORTIUM

50	Acceptance and Development of Beyond Banking Ecosystems Sebastian Voigt, Alexander Holland	773
51	Better Together: Conditions for Sustainable Innovation Aimed at Improving Quality of Care in Health Delivery Organizations Pieter J. Kievit	783
52	Safeguarding Public Values by Design: A Socio-Technical Approach to Digital Twins in Regional Spatial Planning Under the Environmental Planning Act John van Meerten	797
53	Enhancing AI Adoption in European SMEs: A Reference Model for Secure and Ethical Integration Marius Goebel	809

54	The Journey to Sustainable Digital Business Models: Embedding Twin Transformation for Long-Term Survival <i>Tara Vester</i>	821
55	Organizing for Cyber Resilience: Balancing Resilience Paradigms for Sustainable IT Business Value <i>Ziggy Van Giel</i>	841

TOWARDS A TAXONOMY FOR DIGITAL ASSISTANT TECHNOLOGIES: ADDRESSING THE JINGLE-JANGLE FALLACIES

NIKLAS PREIB,¹ MARKUS WESTNER²

¹ University of Sheffield Hallam, Sheffield, United Kingdom of Great Britain and Northern Ireland

niklas.j.preiss@student.shu.ac.uk

² OTH Regensburg, Faculty of Computer Science and Mathematics, Regensburg, Germany

markus.westner@oth-regensburg.de

This study proposes a unified taxonomy for Digital Assistant Technologies (DATs) to resolve terminological inconsistencies and eliminate »Jingle-Jangle fallacies.« By employing a systematic taxonomy development method on 137 papers, the framework categorizes DATs across four meta-characteristics: AI technology, context, intelligence, and interaction. This taxonomy facilitates the clear differentiation of three primary DAT concepts: assistant, chatbot, and agent. By providing a structured framework, the study enhances conceptual clarity, fosters more focused research, and ensures better alignment of DATs.

Keywords:

taxonomy,
digital assistant
technologies,
jingle-jangle
fallacies,
assistant,
chatbot,
agent



University of Maribor Press

DOI
[https://doi.org/
10.18690/um.fov.4.2025.1](https://doi.org/10.18690/um.fov.4.2025.1)

ISBN
978-961-286-998-4

1 Introduction

Recent advancements in Artificial Intelligence (AI), particularly of Generative AI (GAI), have intensified competition among organizations leveraging these innovations for productivity, efficiency, and strategic advantage (Khaokaew et al., 2022; Oldemeyer et al., 2024). Digital Assistant Technologies (DATs), including chatbots, intelligent personal assistants, and intelligent agents (Bowman et al., 2024; de Barcelos Silva et al., 2020; Nenni et al., 2024), drive digital transformation through customer service, task automation, and personalized support. A challenge in DAT research is terminology inconsistency, where terms are used interchangeably or identical terms describe distinct functionalities, creating a jingle-jangle fallacy (Block, 1995; Henry & Liu, 2024; Marsh et al., 2019). Studies show that various terms like Conversational Agents (CA), Virtual Agent (VAG), and Conversational AI Agents (CAIA) often describe overlapping functions (Ahmad et al., 2022; Grimes et al., 2021; Gupta & Dandapat, 2023), complicating research consolidation and technology classification. Broad labels like digital assistants further blur distinctions by unifying diverse DAT, such as voice assistants, AI chatbots, and chatbot assistants, under an overly generalized term (Gkinko & Elbanna, 2023; Sharma et al., 2024). This inconsistency complicates efforts to consolidate research, making it harder to build on existing knowledge and clearly identify the defining features and capabilities of these technologies. Our previous systematic literature review (SLR) of 137 academic articles (January 2013-May 2024) identified 39 distinct DAT terms across three categories: assistants, chatbots, and agents (Preiß & Westner, 2025). Among these 137 papers, only two taxonomy papers were found (Hanelt et al., 2015; Schmidt-Kraepelin et al., 2018), both narrowly focused on specific DAT applications. This reveals a significant gap in standardized classification, leading to our research questions (RQ):

- RQ1: What dimensions and characteristics are essential for constructing a rigorous and generally applicable taxonomy for DATs?
- RQ2: How can a rigorous and generally applicable DAT taxonomy enhance conceptual clarity and practical usability?

This paper addresses these questions by proposing a unified taxonomy that systematically organizes DATs based on shared characteristics. Using Nickerson et al.'s (2013) taxonomy development methodology, refined by Kundisch et al. (2022),

we aim to reduce terminological ambiguity and enhance classification coherence. Our study contributes to Information Systems (IS) research by providing a robust framework for categorizing emerging DATs (Nickerson et al., 2013; Schmidt-Kraepelin et al., 2018), advancing digital transformation literature understanding in this domain.

2 Background and related works

2.1 Research background

DATs drive digital transformation by automating interactions, enhancing customer engagement, and reshaping organizational processes (Bálan, 2023; Choudhary et al., 2024; Pais et al., 2015). From Eliza’s scripted responses in 1966 (Weizenbaum, 1966), to today’s natural language processing (NLP)- and machine learning (ML)-powered chatbots, they have evolved from rule-based systems to context-aware, adaptive assistants supporting diverse tasks, including customer service and legal guidance (Choudhary et al., 2024; Y. Li et al., 2023). This shift positions DATs as enablers of human-machine collaboration (Shneiderman, 2020), moving beyond command-driven tools to proactive, context-sensitive partners.

2.2 Related DAT research

In recent years, an extensive number of terms describing DAT have emerged (Preiß & Westner, 2025). Terms like Virtual Assistant (VAS), Intelligent Virtual Assistant (IVA), Voice Assistant (VA), Digital Assistant (DA), Intelligent Personal Assistant (IPA), Chatbot, AI Chatbot, CA, VAG, Intelligent Agent (IA) have been employed with considerable overlap, often without consistent distinctions. For instance, Beer et al. (2015, p. 2) observed that the label ‘agent’ is widely used, yet there is no agreed-upon definition” and assistant and agent are frequently synonymous (Lee et al., 2024). Terms like IVA and IPA are often associated with specialized functionalities, indicating capabilities such as personalized, voice-driven interactions and adaptive responses (Alimamy & Kuhail, 2023; Pais et al., 2015; Priya et al., 2023). However, as pointed out by Ammari et al. (2019), VA has become the »industry standard« term for speech-driven assistants, despite the existence of more specific terms like IVA and IPA. These terms are applied to tools like Siri, Alexa, and Cortana, highlighting the terminological overlap (Brachten et al., 2020; C. Li et al., 2023; Porra et al., 2020;

Ripa et al., 2023). VA and IVA presents varied interpretations: some sources consider VAS an umbrella term for voice-enabled systems, while others suggest that »intelligent virtual assistant« denotes systems with more advanced, context-aware capabilities with text-based input (C. Li et al., 2023; Richards & Bransky, 2014). Chatbot and IVA exemplify distinctions between task-specific chatbots (Chen et al., 2023; Koivunen et al., 2022) and broader-function technologies like Siri and Alexa (McKillop et al., 2021; Rapp et al., 2021). Similarly, while DAs are sometimes viewed as equivalent to CAs, CAs are occasionally considered simpler, rule-based systems (Khaokaew et al., 2022; Schuetzler et al., 2020). These overlapping terms frequently lead to different concepts being conflated under similar labels. While Chatbot and CA are often used interchangeably (McKillop et al., 2021; Schuetzler et al., 2020), CA typically suggests a broader range of conversational functionality (Ngai et al., 2021), with CAs being more adaptable across applications and chatbots serving narrower, task-focused roles (Grimes et al., 2021; Schuetzler et al., 2020). A central differentiator among these assistants is the level of anthropomorphism and autonomy. Terms like IVA and IPA frequently emphasize adaptive, user-centered feature (Alimamy & Kuhail, 2023; Pais et al., 2015), while simpler VAs and CAs typically operate within rule-based, pre-programmed parameters (Behera et al., 2024; Chen et al., 2023; Stieglitz et al., 2022). Labels like »intelligent« or »personal« denote a stronger emphasis on customization. The term's application to sophisticated systems like ChatGPT (Ding et al., 2024), demonstrates the persisting jingle-jangle fallacy within the latest generation of DAT encompassing GAI. These findings underscore the need for a coherent taxonomy, as inconsistent terminology creates ambiguity evident in technologies like Alexa and Siri, described by over 11 different terms (Preiß & Westner, 2025). To our knowledge, no taxonomy fully captures DAT concepts, dimensions and characteristics, highlighting the need for a structured framework to enhance clarity in research and practice.

3 Research method: Taxonomy development process

Phenomenon and motivation (I): Taxonomies organize and classify knowledge through hierarchical relationships (Šmite et al., 2014). We build on the iterative taxonomy design process outlined by Nickerson et al. (2013) and refined by Kundisch et al. (2022), ensuring rigor and relevance in our systematic framework. This section delineates the taxonomy development process, a fundamental methodological approach in IS research that structures complex domains into

organized categories (Šmite et al., 2014). According to Nickerson et al. (2013), a taxonomy is a formal system for classifying objects based on a set of characteristics or dimensions. Szopinski et al. (2019) emphasize that taxonomies must remain iterative to capture evolving IS domains effectively. The framework by Kundisch et al. (2022) integrates empirical and conceptual iterations through multiple cycles with defined termination conditions, ensuring completeness, clarity, and practical applicability. According to Nickerson, taxonomy development begins with identifying meta-characteristics, which establish the taxonomy's primary focus and direct the selection of essential dimensions. Kundisch et al. (2022) outline three preparatory steps: specifying the observed phenomenon, identifying target user group(s), and articulating intended purpose(s). The observed phenomenon is the rapid proliferation of DATs. The primary user group includes academics, researchers, and practitioners who seek a robust and systematic classification framework. The main objective and motivation is to standardize DAT terminology for clarity, consistency, and practical use. This taxonomy aims to integrate all three concepts identified in Preiß & Westner (2025), moving beyond prior taxonomy efforts that have typically focused on a single concept or isolated subsets of DATs within the scope of DAT research (Bahja & Lowry, 2021; Gkinko & Elbanna, 2023; Janssen et al., 2020; Nißen et al., 2022; Schmidt-Kraepelin et al., 2018). Meta-characteristics, introduced by Nickerson et al. (2013) define central categories and guide dimension selection. Alternative terms include perspectives (Janssen et al., 2020) or meta-dimensions (Möller et al., 2022; Rosian et al., 2021). We adopt the three-tier hierarchical structure from Nickerson et al. (2013) meta-characteristics, dimensions, and characteristics, using a multilevel index for precise organization (Saldana, 2021).

Objectives and ending conditions (II): The taxonomy development process requires (1) ending conditions (Nickerson et al., 2013) and (2) evaluation goals (Kundisch et al., 2022) before initiation. Ending conditions define when to conclude iterations based on conceptual (e.g., theoretical justification) and empirical (e.g., saturation) criteria. This study adopts Nickerson's conditions with an added meta-characteristics criterion (Szopinski et al., 2019). The evaluation goals selected for this study are 'analyzing' and 'clustering'. The analyzing goal is »investigating objects that represent a particular phenomenon, utilizing characteristics and dimensions to identify similarities and differences« (Kundisch et al., 2022, p. 432). The clustering goal focuses on grouping objects based on »identified commonalities, enabling the

classification of types [...]» (Kundisch et al., 2022, p. 432). After three iterations analyzing 106 papers (most present 5 DATs from each concept; from 137 identified in Preiß & Westner, 2025), all ending conditions were met, confirming the taxonomy's robustness (Nickerson et al., 2013). The second iteration introduced three dimensions, improving clarity. The final taxonomy classified 13 DATs, representing 33% of identified DATs but covering 77% of the analyzed literature.

Development approach – conceptual-to-empirical (III): Building on the steps outlined in the previous chapters, Nickerson et al. (2013) differentiate between two fundamental approaches to taxonomy development: conceptual-to-empirical and empirical-to-conceptual. In the conceptual-to-empirical approach, taxonomy development begins with theoretical constructs or literature, which are then validated through empirical data. Conversely, the empirical-to-conceptual approach starts with data collection (e.g., case studies or interviews) to derive characteristics, which are subsequently refined and grounded in theory. This iterative nature is supported by Hevner et al. (2004, p. 88), who describe »design as a search process« that enables continuous testing and validation cycles throughout taxonomy development. In our study, we applied the conceptual-to-empirical approach, building our taxonomy on existing DAT literature in IS as discussed in the SLR (Preiß & Westner, 2025).

4 Taxonomy for DATs

4.1 Taxonomy overview

This section provides a detailed description of the meta-characteristics (MC), dimensions (D), and characteristics (C) of the taxonomy, along with the final DAT taxonomy (see Figure 1). According to Nickerson et al. (2013), a taxonomy's characteristics must be mutually exclusive within each dimension, a principle upheld during the taxonomy's development. Non-self-explanatory characteristics were thoroughly described to ensure clarity and understanding. Dimensions were systematically ordered using three criteria: alphabetical order, prevalence (low to high or less to many), and evaluation of presence (present to absent) and individual characteristics were further organized alphabetically to ensure clarity.

Meta-Characteristic (MCn)	Dimension (Dnm)	Characteristics (Cnmk)							
AI technology	AI models	traditional AI models			GAI / DL models			none	
	application domain	academia	daily life	e-commerce	e-learning / education	healthcare	industry	service industry	mobile / work / business
context	beneficiary audience	corporate / organization		individual		both			
	motivation / goal	customer service	daily life assistance	healthcare service	socializing / cohesiveness	productivity / efficiency	other		
	scenario coverage	use case = 1			use case > 1				
	service intent	commercial intention			individual support			both	
intelligence	context awareness	low			extensive				
	degree of anthropomorphism	low			high				
	environmental adaptability	static			adaptive				
	learning ability	pre-defined knowledge			rudimentary			extensive	
interaction	appearance	no appearance			low appearance			high appearance	
	initiative	delegated			mixed			autonomous	
	input	multiple		text	visual / gesture		voice / speech		
	number of users addressed	users = 1			users > 1				
	user interface	app	device		multiple	proprietary software		website	
		Order coding:		alphabetical	less / low to many / high	present (alphabetical) to absent			

Figure 1: Thorough DAT taxonomy

Source: Own

AI technology (MC₁) includes the dimension AI models (D₁₁), which categorizes the types of AI models that power DAT. Traditional AI models (C₁₁₁) are rule-based systems, relying mostly on traditional machine learning techniques that perform fixed tasks by following patterns in historical data. These models are simpler, interpretable, and suited to structured, stable tasks. Examples include decision trees, Bayesian method, reinforcement learning and rule-based systems (Assaf et al., 2020; Groshev et al., 2021; Somers et al., 2019). GAI / DL models (C₁₁₂) use deep learning, especially transformer models, to generate dynamic responses (Okey et al., 2023; Vaswani et al., 2017). Unlike classical AI, they create new content by predicting sequences based on large datasets, allowing for flexible, context-aware interactions ideal for complex tasks like conversational AI (Rillig & Kasirzadeh, 2024). Some DATs may operate without AI, relying instead on simple, predefined logic and data sets to facilitate only basic interactions and are categorized as none (C₁₁₃) (Beldad et al., 2016). **Context** (MC₂) defines the environmental and situational framework in which DATs operate, encompassing implicit and explicit information about the intended user group, purpose, and domain (Abowd et al., 1999; K. Kim et al., 2018). This meta-characteristics clarifies where and why DATs operate, illustrating how they are designed or applied to fulfill specific user needs and achieve operational goals across diverse scenarios (Diederich et al., 2020; Gnewuch et al., 2017). MD₂ structures five key dimensions that encompass 22 characteristics (C₂₁₁-C₂₅₃). Application domain (D₂₁) specifies the primary application area for which the DAT is designed for. This dimension highlights the variety of contexts in which DATs operates in (Zumstein & Hundertmark, 2017). Ranging from academic (C₂₂₁) to work/business (C₂₂₆) Beneficiary audience (D₂₂) identifies the primary user group the DAT is intended to serve. This includes

corporates or organizations (C₂₂₁) for professional use, individual users (C₂₂₂) in personal settings, or both (C₂₂₃), catering to a combination of personal and organizational needs (Gkinko & Elbanna, 2021). Motivation/goal (D₂₃) identifies the primary purpose or objective behind the deployment of DATs, clarifying what the DAT is intended to achieve for its user within a specific domain (Bittner et al., 2019). This dimension captures the role the DAT plays in supporting users' objectives in various contexts (Knote et al., 2018; Ryan & Deci, 2000). It includes customer service (C₂₃₁), handling customer interactions and needs, daily life assistance (C₂₃₂), aimed at supporting routine personal tasks; healthcare service (C₂₃₃), designed to address health-related needs, socializing/caretakership (C₂₃₄), providing interpersonal support or companionship, productivity/efficiency (C₂₃₅), enhancing operational effectiveness and increasing productivity, and other (C₂₃₆), encompassing objectives outside the described characteristics. Scenario coverage (D₂₄) describes the DAT's design across different usage scenarios, indicating whether it is tailored for single-use case scenarios (C₂₄₁) or broader applications with multiple use cases (C₂₄₂). This dimension highlights the DAT's flexibility and adaptability across various contexts, showcasing its capacity to support diverse interaction needs and operational demands (Følstad et al., 2019). Service intent (D₂₅) differentiates between the primary purpose of the DAT. It includes commercial intention (C₂₅₁), where the DAT is designed to generate revenue directly or indirectly through its assistance, individual support (C₂₅₂), where the primary goal is to provide non-commercial help to users, and both (C₂₅₃), where the DAT combines commercial objectives with individual support (de Barcelos Silva et al., 2020; Knote et al., 2021). **Intelligence** (MC₃) in digital technology, is defined by a system's ability to perceive, learn from data, make autonomous decisions, and adapt to its environment (Boden, 2016; Brynjolfsson & McAfee, 2011; Turing, 1950). This encompasses essential qualities such as simulating human cognition including human-like emotions (anthropomorphism), gaining new knowledge through interaction or training (learning ability), maintaining contextual awareness of interactions and surroundings (context awareness), and modifying behavior in response to new information (adaptability) (Gunkel, 2012; Hengstler et al., 2016; Nilsson, 2009; Russell, 2016). These attributes distinguish intelligent systems from traditional technologies which are mostly rule-based and pre-defined (Gignac & Szodorai, 2024). Context awareness (D₃₁) defines a DAT's ability to incorporate contextual information into interactions. Systems with low context awareness (C₃₁₁)

operate strictly based on predefined parameters, without considering past interactions or real-time inputs. More advanced DATs (C₃₁₂) dynamically adapt to ongoing interactions, integrating past exchanges and situational cues to personalize responses (Abowd et al., 1999; Noonian et al., 2024). The degree of anthropomorphism (D₃₂) captures the extent to which DATs exhibit human-like qualities (Kim & Im, 2023). It ranges from low anthropomorphism (C₃₂₁), where the system lacks human-like features, to high anthropomorphism (C₃₂₂). According to (Epley, 2018), DATs are »humanized« through features such as physical traits (e.g., voice) along with personality and emotional expressions. Environmental adaptability (D₃₃) reflects a DAT's ability to recognize and respond to contextual changes such as user behavior, location, or ambient conditions. Less adaptable systems (C₃₃₁) rely on static, predefined responses, while adaptive systems (C₃₃₂) anticipate user needs and modify their functionality accordingly (Akata et al., 2020; Pais et al., 2015). Learning ability (D₃₄) evaluates a DAT's capacity to acquire, adapt, and apply knowledge. It spans from predefined knowledge (C₃₄₁) (Gignac & Szodorai, 2024), where the DAT relies entirely on static predefined knowledge or datasets, to rudimentary (C₃₄₂) (Weizenbaum, 1966), which involves basic adjustments, and extensive (C₃₄₃) (Noonian et al., 2024), characterized by advanced self-learning capabilities that enable the system to adapt, evolve, and enhance interactions over time (Nilsson, 2009; Turing, 1950). This ability is critical for improving knowledge and delivering increasingly effective and personalized interactions (Rashid & Kausik, 2024). Interaction (MC₄) explores how DATs engage with users, focusing on appearance, input modes, initiative, and the number of users addressed. This dimension highlights the methods and modalities that shape user visual experience and accessibility (Kim & Im, 2023; Noonian et al., 2024). Appearance (D₄₁) in DATs ranges from no appearance (C₄₁₁), where the assistant operates invisibly without a visual representation, to low appearance (C₄₁₂), characterized by robot-like features, limited gestures, and blank facial expressions. High appearance (C₄₁₃) at DATs include human-like avatars with smooth hand gestures and varied facial expressions, such as smiling and mimicking emotions (Alabed et al., 2022; Kim & Im, 2023). Initiative (D₄₂) in DATs describes the level of control over the interaction flow, ranging from delegated (C₄₂₁), where users initiate and direct all interactions, to mixed (C₄₂₂), where control is shared, with the DAT occasionally offering suggestions or taking action based on context (Angin et al., 2018; Kugele et al., 2021). At the highest level, autonomous (C₄₂₃), the DAT

independently initiates interactions, anticipates needs, and adapts dynamically to changes, showcasing proactive and context-aware capabilities (Jiang & Arkin, 2015). Input (D₄₃) in DATs defines how the system receives input from users or other entities, ranging from text-based (C₄₃₂) inputs, such as chat typing, to visual/gesture-based (C₄₃₂) commands involving gestures, emotional expressions or visual cues. It also covers voice or speech (C₄₃₃) inputs using spoken language, multiple (C₄₃₁) inputs combining various modalities for enhanced flexibility (Kiseleva et al., 2016; Rubio-Drosdov et al., 2017; Wellsandt et al., 2021; Zhu et al., 2014). Number of users addressed (D₄₄) in DATs defines the system's capacity to interact with one or multiple users simultaneously and how the experience differs for each user. It indicates whether the DAT operates in a 1:1 setup (C₄₄₁), focusing on individual interactions, or a 1-to-many setup (C₄₄₂), managing interactions across multiple users concurrently (Noonia et al., 2024). User interface (D₄₅) in DATs refers to the platform or medium enabling user interaction with the system. It includes app (C₄₅₁), where interaction occurs via a dedicated application; device (C₄₅₂), where the DAT is embedded in a physical device, such as a smart speaker; multiple (C₄₅₃), where the DAT operates seamlessly across various platforms or devices; proprietary software (C₄₅₄), which relies on custom-built interfaces tailored to its ecosystem; and website (C₄₅₅), where interaction is facilitated through an interface built into a website (Klopfenstein et al., 2017; Knote et al., 2019; Noonina et al., 2024; Valtolina et al., 2020).

4.2 Taxonomy evaluation

Our evaluation employs morphological field analysis (Ritchey, 2011; Rittelmeyer & Sandkuhl, 2023) to validate the taxonomy's effectiveness in classifying DATs and addressing terminology inconsistencies (see online Appendix¹). This approach is particularly relevant for DATs, which frequently recombine patterns in their application, perception, and implementation (Knote et al., 2019; Rittelmeyer & Sandkuhl, 2023). The assistant concept encompasses five DATs (VAS, VA, IVA, DA, and IPA), revealing significant terminology overlap. For instance, while IPA and IVA both target individual users, they lack distinctive characteristics even in areas their names suggest (e.g., »personal« applicability or advanced intelligence).

¹ <https://doi.org/10.6084/m9.figshare.28330274.v1>

One clear distinction emerges with VA, which is uniquely defined by exclusive voice input functionality, providing a legitimate taxonomic differentiation (Brachten et al., 2020; Ki et al., 2020; C. Li et al., 2023; Porra et al., 2020; Ripa et al., 2023). Chatbot (30% of our sample; Preiß & Westner, 2025). follow a simpler structure, mainly split between »chatbot« and »AI chatbot.« As AI has become standard, the »AI« modifier is now redundant. ChatGPT, though often labeled a chatbot, functions more like an assistant due to GAI and multimodal capabilities (Khennouche et al., 2024; Okey et al., 2023), distinguishing it from traditional customer-service chatbots (Ayers et al., 2023; Caccavale et al., 2024). The agent concept (29% of the sample, Preiß & Westner, 2025), shows excessive overlap, with IA, AI agents, and VAG largely indistinguishable in function. For example, CA and CAIA demonstrate identical morphological patterns, supporting their treatment as synonymous terms (Adamopoulou & Moussiades, 2020; Alabed et al., 2022; Ashfaq et al., 2020).

4.3 Synthesis and Integration

Validating the findings, we synthesized the three concepts into a unified visualization (Figure 2), which maps their manifestations within the taxonomy. This synthesis reveals that while distinct patterns emerge for each concept, significant overlap exists in certain characteristics and modifiers like »AI,« »intelligent,« and »personal« are often used interchangeably (Knote et al., 2019; Pais et al., 2015), confirming the need for more precise terminology. This empirical validation demonstrates our taxonomy's effectiveness in: Clarifying distinctions between core concepts, identifying redundant terminology, providing a framework for consistent classification, and supporting more precise naming conventions for future DAT development. The taxonomy also aids decision-making by helping practitioners align the terminology with the appropriate framework. For example, the process begins by identifying the underlying concept being targeted (assistant, chatbot, or agent), followed by determining the dominant differentiator—whether a technical aspect or a contextual element. Moving forward, the use of DAT terms will be more systematic, reducing variability and focusing on a single defining characteristic, such as »virtual« or »voice,« to ensure consistency and clarity in classification.

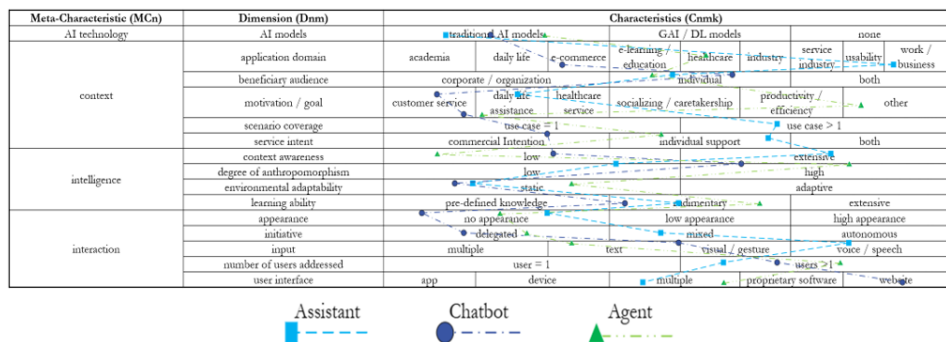


Figure 2: Visualization of the three main concepts within the taxonomy as a morphological field

Source: Own

5 Discussion & Conclusion

The findings from the morphological fields highlight jingle-jangle fallacies, where identical terms represent distinct concepts, and similar terms denote different functionalities (Marsh et al., 2019). To address this, the taxonomy is structured around four meta-characteristics: AI technology, context, intelligence, and interaction, further divided into 15 dimensions and 52 characteristics, ensuring clarity and consistency. Testing on chatbots, agents, and assistants confirmed the taxonomy's ability to clarify distinctions and address terminological inconsistencies from the morphological field analysis. Academically, the taxonomy establishes a unified framework that resolves overlapping and redundant terminologies, such as the conflation of IPA and VA or the overuse of »intelligent.« It supports comparative studies, theory-building, and systematic identification of research gaps by categorizing DATs into well-defined meta-characteristics. Practically, the taxonomy serves as a decision-making tool for aligning DAT functionalities with organizational goals and contexts. For RQ1, effective classification requires both technical (AI models, learning capabilities) and contextual (application domains, interaction modalities) elements. Regarding RQ2, the taxonomy enhances conceptual clarity and practical usability, bridging theoretical and practical gaps while addressing jingle-jangle fallacies (see Figure 1). Despite a rigorous methodology (Nickerson et al., 2013), general validity remains a challenge due to the rapid evolution of DATs. Subjective choices in defining dimensions and ending conditions (Nickerson et al., 2013) also influence its structure, and an empirical approach (e.g., developer

interviews) could offer alternative insights. The focus on GAI-based DATs may have narrowed findings, and excluding terms like bot and multi-agent systems (Preiß & Westner, 2025), could limit relevance to specific fields. Future research should broaden the technological scope and validate the taxonomy across varied sectors and emerging technologies, as current findings are primarily based on recent proprietary GAI-based DATs. Exploring interrelations among meta-characteristics and their impact on DAT development and adoption is essential. Harmonizing existing DAT classifications can also address terminology inconsistencies and enhance coherence. As technological innovations continue, the taxonomy must evolve, highlight the necessity of regular updates to the taxonomy as it »may be a moving target« (Nickerson et al., 2013, p. 341).

Data availability. Additional data in form of an online appendix is provided here: <https://doi.org/10.6084/m9.figshare.28330274.v1>

References

- Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., & Steggles, P. (1999). Towards a better understanding of context and context-awareness. In H.-W. Gellersen (Ed.), *Handheld and Ubiquitous Computing* (pp. 304–307). Springer. https://doi.org/10.1007/3-540-48157-5_29
- Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In I. Maglogiannis, L. Iliadis, & E. Pimenidis (Eds.), *Artificial Intelligence Applications and Innovations* (pp. 373–383). Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31
- Ahmad, R., Siemon, D., Gnewuch, U., & Robra-Bissantz, S. (2022). Designing personality-adaptive conversational agents for mental health care. *Information Systems Frontiers*, 24(24), 923–943. <https://doi.org/10.1007/s10796-022-10254-9>
- Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F., ... Welling, M. (2020). A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(8), 18–28. <https://doi.org/10.1109/MC.2020.2996587>
- Alabed, A., Javornik, A., & Gregory-Smith, D. (2022). AI anthropomorphism and its effect on users' self-congruence and self-AI integration: A theoretical framework and research agenda. *Technological Forecasting and Social Change*, 182, 121786. <https://doi.org/10.1016/j.techfore.2022.121786>
- Alimamy, S., & Kuhail, M. A. (2023). I will be with you Alexa! The impact of intelligent virtual assistant's authenticity and personalization on user reuse intentions. *Computers in Human Behavior*, 143. <https://doi.org/10.1016/j.chb.2023.107711>
- Ammari, T., Kaye, J., Tsai, J. Y., & Bentley, F. (2019). Music, search, and iot: How people (really) use voice assistants. *ACM Transactions on Computer-Human Interaction*, 26(3), 28. <https://doi.org/10.1145/3311956>

- Angin, P., Bhargava, B., & Ranchal, R. (2018). A self-protecting agents based model for high-performance mobile-cloud computing. *Computers and Security*, 77, 380–396. <https://doi.org/10.1016/j.cose.2018.04.011>
- 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>
- Assaf, A. G., Bu, R., & Tsionas, M. G. (2020). A bayesian approach to continuous type principal-agent problems. *European Journal of Operational Research*, 280(3), 1188–1192. <https://doi.org/10.1016/j.ejor.2019.07.058>
- 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., & 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), 589–596. <https://doi.org/10.1001/jamainternmed.2023.1838>
- Bahja, J., & Lowry, P. (2021). Toward a taxonomy of service-oriented chatbots. *SIGHCI 2021, SIGHCI 2021 Proceedings*. 15. <https://aisel.aisnet.org/sighci2021/15>
- Bălan, C. (2023). Chatbots and voice assistants: Digital transformers of the company. *Journal of Theoretical and Applied Electronic Commerce Research*, 18(2), 995–1019. <https://doi.org/10.3390/jtaer18020051>
- Beer, J. M., Smarr, C.-A., Fisk, A. D., & Rogers, W. A. (2015). Younger and older users' recognition of virtual agent facial expressions. *International Journal of Human Computer Studies*, 75, 1–20. <https://doi.org/10.1016/j.ijhcs.2014.11.005>
- Behera, R. K., Bala, P. K., & Ray, A. (2024). Cognitive chatbot for personalised contextual customer service: Behind the scene and beyond the hype. *Information Systems Frontiers*, 26(3), 899–919. <https://doi.org/10.1007/s10796-021-10168-y>
- Beldad, A., Hegner, S., & Hoppen, J. (2016). The effect of virtual sales agent (VSA) gender: Product gender congruence on product advice credibility, trust in VSA and online vendor, and purchase intention. *Computers in Human Behavior*, 60, 62–72. <https://doi.org/10.1016/j.chb.2016.02.046>
- Bittner, E., Oeste-Reiß, S., & Leimeister, J. M. (2019). *Where is the bot in our team? Toward a taxonomy of design option combinations for conversational agents in collaborative work*. Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2019.035>
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117(2), 187–215. <https://doi.org/10.1037/0033-2909.117.2.187>
- Boden, M. A. (2016). *AI: Its nature and future*. Oxford University Press.
- Bowman, R., Cooney, O., Newbold, J. W., Thieme, A., Clark, L., Doherty, G., & Cowan, B. (2024). Exploring how politeness impacts the user experience of chatbots for mental health support. *International Journal of Human Computer Studies*, 184. <https://doi.org/10.1016/j.ijhcs.2023.103181>
- Brachten, F., Brunker, F., Frick, N. R. J., Ross, B., & Stieglitz, S. (2020). On the ability of virtual agents to decrease cognitive load: An experimental study. *Information Systems and E-Business Management*, 18(2), 187–207. <https://doi.org/10.1007/s10257-020-00471-7>
- Brynjolfsson, E., & McAfee, A. (2011). Winning the race with ever-smarter machines. *MIT Sloan Management Review*, 53, 53–60.
- Caccavale, F., Gargalo, C. L., Gernaey, K. V., & Krühne, U. (2024). Towards education 4.0: The role of large language models as virtual tutors in chemical engineering. *Education for Chemical Engineers*, 49, 1–11. <https://doi.org/10.1016/j.ece.2024.07.002>
- Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers*, 25(1), 161–182. <https://doi.org/10.1007/s10796-022-10291-4>
- Choudhary, S., Kaushik, N., Sivathanu, B., & Rana, N. P. (2024). Assessing factors influencing customers' adoption of ai-based voice assistants. *Journal of Computer Information Systems*, 1–15. <https://doi.org/10.1080/08874417.2024.2312858>

- de Barcelos Silva, A., Gomes, M. M., da Costa, C. A., da Rosa Righi, R., Barbosa, J. L. V., Pessin, G., De Doncker, G., & Federizzi, G. (2020). Intelligent personal assistants: A systematic literature review. *Expert Systems with Applications*, 147. <https://doi.org/10.1016/j.eswa.2020.113193>
- Diederich, S., Brendel, A. B., & Kolbe, L. M. (2020). Designing anthropomorphic enterprise conversational agents. *Business and Information Systems Engineering*, 62(3), 193–209. <https://doi.org/10.1007/s12599-020-00639-y>
- Ding, H., Simmich, J., Vaezipour, A., Andrews, N., & Russell, T. (2024). Evaluation framework for conversational agents with artificial intelligence in health interventions: A systematic scoping review. *Journal of the American Medical Informatics Association*, 31(3), 746–761. <https://doi.org/10.1093/jamia/ocad222>
- Epley, N. (2018). A mind like mine: The exceptionally ordinary underpinnings of anthropomorphism. *Journal of the Association for Consumer Research*, 3(4), 591–598. <https://doi.org/10.1086/699516>
- Følstad, A., Skjuve, M., & Brandtzaeg, P. (2019). Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design. *INSCI 2018*, 145–156. https://doi.org/10.1007/978-3-030-17705-8_13
- Gignac, G. E., & Szodorai, E. T. (2024). Defining intelligence: Bridging the gap between human and artificial perspectives. *Intelligence*, 104, 101832. <https://doi.org/10.1016/j.intell.2024.101832>
- Gkinko, L., & Elbanna, A. (2021). Chatbots at work: A taxonomy of the use of chatbots in the workplace. In D. Dennehy, A. Griva, N. Pouloudi, Y. K. Dwivedi, I. Pappas, & M. Mäntymäki (Eds.), *Responsible AI and Analytics for an Ethical and Inclusive Digitized Society* (pp. 29–39). Springer International Publishing. https://doi.org/10.1007/978-3-030-85447-8_3
- Gkinko, L., & Elbanna, A. (2023). The appropriation of conversational AI in the workplace: A taxonomy of AI chatbot users. *International Journal of Information Management*, 69, 1–11. <https://doi.org/10.1016/j.ijinfomgt.2022.102568>
- Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. *ICIS 2017*. International Conference on Information Systems.
- Grimes, G. M., Schuetzler, R. M., & Giboney, J. S. (2021). Mental models and expectation violations in conversational AI interactions. *Decision Support Systems*, 144, 1–10. <https://doi.org/10.1016/j.dss.2021.113515>
- Groshev, M., Guimaraes, C., Martin-Perez, J., & de la Oliva, A. (2021). Toward intelligent cyber-physical systems: Digital twin meets artificial intelligence. *IEEE Communications Magazine*, 59(8), 14–20. <https://doi.org/10.1109/MCOM.001.2001237>
- Gunkel, D. (2012). Communication and artificial intelligence: Opportunities and challenges for the 21st century. *Communication +1*, 1, 1–25. <https://doi.org/10.7275/R5QJ7F7R>
- Gupta, S., & Dandapat, S. K. (2023). SEEC and CHASE: An emotion-cause pair-oriented approach and conversational dataset with heterogeneous emotions for empathetic response generation. *Knowledge-Based Systems*, 280(1), 1–15. <https://doi.org/10.1016/j.knosys.2023.111039>
- Hanelt, A., Hildebrandt, B., & Polier, J. (2015). Uncovering the role of is in business model innovation: A taxonomy-driven approach to structure the field. *ECIS 2015 Completed Research Papers*. <https://doi.org/10.18151/7217345>
- Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust: The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, 105, 105–120. <https://doi.org/10.1016/j.techfore.2015.12.014>
- Henry, A., & Liu, M. (2024). Jingle-jangle fallacies in l2 motivational self system research: A response to al-hoorie et al. (2024). *Applied Linguistics*, 45, 738–746. <https://doi.org/10.1093/applin/amae041>
- 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(106), 75–115.
- Janssen, A., Passlick, J., Rodríguez Cardona, D., & Breitner, M. H. (2020). Virtual assistance in any context: A taxonomy of design elements for domain-specific chatbots. *Business and Information Systems Engineering*, 62(3), 211–225. <https://doi.org/10.1007/s12599-020-00644-1>

- Jiang, S., & Arkin, R. C. (2015). Mixed-initiative human-robot interaction: Definition, taxonomy, and survey. *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 954–961. <https://doi.org/10.1109/SMC.2015.174>
- Khaokaew, Y., Holcombe-James, I., Rahaman, M. S., Liono, J., Trippas, J. R., Spina, D., Bailey, P., Belkin, N. J., Bennett, P. N., Ren, Y., Sanderson, M., Scholer, F., White, R. W., & Salim, F. D. (2022). Imagining future digital assistants at work: A study of task management needs. *International Journal of Human Computer Studies*, 168, 1–17. <https://doi.org/10.1016/j.ijhcs.2022.102905>
- Khennouche, F., Elmir, Y., Himeur, Y., Djebbari, N., & Amira, A. (2024). Revolutionizing generative pre-trained: Insights and challenges in deploying ChatGPT and generative chatbots for FAQs. *Expert Systems with Applications*, 246, 1–25. <https://doi.org/10.1016/j.eswa.2024.123224>
- Ki, C.-W. C., Cho, E., & Lee, J.-E. (2020). Can an intelligent personal assistant (IPA) be your friend? Para-friendship development mechanism between IPAs and their users. *Computers in Human Behavior*, 111, 1–10. <https://doi.org/10.1016/j.chb.2020.106412>
- Kim, J., & Im, I. (2023). Anthropomorphic response: Understanding interactions between humans and artificial intelligence agents. *Computers in Human Behavior*, 139. <https://doi.org/10.1016/j.chb.2022.107512>
- Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G., & Welch, G. F. (2018). Does a digital assistant need a body? The influence of visual embodiment and social behavior on the perception of intelligent virtual agents in ar. *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 105–114. <https://doi.org/10.1109/ISMAR.2018.00039>
- Kiseleva, J., Williams, K., Hassan Awadallah, A., Crook, A. C., Zitouni, I., & Anastasakos, T. (2016). Predicting user satisfaction with intelligent assistants. *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 45–54. <https://doi.org/10.1145/2911451.2911521>
- Klopfenstein, L., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. *DIS '17*, 555–565. <https://doi.org/10.1145/3064663.3064672>
- Knote, R., Janson, A., Eigenbrod, L., & Söllner, M. (2018). The what and how of smart personal assistants: Principles and application domains for is research. *Multikonferenz Wirtschaftsinformatik (MKWI)*, 1–13.
- Knote, R., Janson, A., Söllner, M., & Leimeister, J. M. (2019). Classifying smart personal assistants: An empirical cluster analysis. *52th HICSS*, 2024–2033. <https://doi.org/10.24251/HICSS.2019.245>
- Knote, R., Janson, A., Söllner, M., & Leimeister, J. M. (2021). Value co-creation in smart services: A functional affordances perspective on smart personal assistants. *Journal of the Association for Information Systems*, 22(2), 418–458. <https://doi.org/10.17705/1jais.00667>
- Koivunen, S., Ala-Luopa, S., Olsson, T., & Haapakorpi, A. (2022). The march of chatbots into recruitment: Recruiters' experiences, expectations, and design opportunities. *Computer Supported Cooperative Work: CSCW: An International Journal*, 31(3), 487–516. <https://doi.org/10.1007/s10606-022-09429-4>
- Kugele, S., Petrovska, A., & Gerostathopoulos, I. (2021). Towards a taxonomy of autonomous systems. In S. Biffl, E. Navarro, W. Löwe, M. Sirjani, R. Mirandola, & D. Weyns (Eds.), *Software Architecture* (pp. 37–45). Springer. https://doi.org/10.1007/978-3-030-86044-8_3
- Kundisch, D., Muntermann, J., Oberländer, A. M., Rau, D., Röglinger, M., Schoormann, T., & Szopinski, D. (2022). An update for technology designers. *Business & Information Systems Engineering*, 64(4), 421–439. <https://doi.org/10.1007/s12599-021-00723-x>
- Lee, S., Park, Y., & Park, G. (2024). Using AI chatbots in climate change mitigation: A moderated serial mediation model. *Behaviour and Information Technology*, 43(16), 4016–4032. <https://doi.org/10.1080/0144929X.2023.2298305>

- Li, C., Chrysostomou, D., & Yang, H. (2023). A speech-enabled virtual assistant for efficient human–robot interaction in industrial environments. *Journal of Systems and Software*, 205, 1–11. <https://doi.org/10.1016/j.jss.2023.111818>
- Li, Y., Gan, Z., & Zheng, B. (2023). How do artificial intelligence chatbots affect customer purchase? Uncovering the dual pathways of anthropomorphism on service evaluation. *Information Systems Frontiers*, 283–300. <https://doi.org/10.1007/s10796-023-10438-x>
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology*, 111(2), 331–353. <https://doi.org/10.1037/edu0000281>
- McKillop, M., South, B. R., Preininger, A., Mason, M., & Jackson, G. P. (2021). Leveraging conversational technology to answer common COVID-19 questions. *Journal of the American Medical Informatics Association*, 28(4), 850–855. <https://doi.org/10.1093/jamia/ocaa316>
- 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, 701–726. <https://doi.org/10.1007/s12525-021-00507-x>
- Nenni, M. E., De Felice, F., De Luca, C., & Forcina, A. (2024). How artificial intelligence will transform project management in the age of digitization: A systematic literature review. *Management Review Quarterly*. <https://doi.org/10.1007/s11301-024-00418-z>
- Ngai, E. W. T., Lee, M. C. M., Luo, M., Chan, P. S. L., & Liang, T. (2021). An intelligent knowledge-based chatbot for customer service. *Electronic Commerce Research and Applications*, 50. <https://doi.org/10.1016/j.elerap.2021.101098>
- 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. <https://doi.org/10.1057/ejis.2012.26>
- Nilsson, N. J. (2009). *The quest for artificial intelligence*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511819346>
- Nißen, M., Selimi, D., Janssen, A., Cardona, D. R., Breitner, M. H., Kowatsch, T., & von Wangenheim, F. (2022). See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons. *Computers in Human Behavior*, 127, 1–15. <https://doi.org/10.1016/j.chb.2021.107043>
- Noonia, A., Beg, R., Patidar, A., Bawaskar, B., Sharma, S., & Rawat, H. (2024). Chatbot vs intelligent virtual assistance (IVA). In *Conversational Artificial Intelligence* (pp. 655–673). John Wiley & Sons. <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781394200801.ch36>
- Okey, O. D., Udo, E. U., Rosa, R. L., Rodríguez, D. Z., & Kleinschmidt, J. H. (2023). Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis. *Computers and Security*, 135, 1–10. <https://doi.org/10.1016/j.cose.2023.103476>
- Oldemeyer, L., Jede, A., & Teuteberg, F. (2024). Investigation of artificial intelligence in SMEs: A systematic review of the state of the art and the main implementation challenges. *Management Review Quarterly*. <https://doi.org/10.1007/s11301-024-00405-4>
- Pais, S., Casal, J., Ponciano, R., & Lourenço, S. (2015, January 23). Unsupervised assistive and adaptive intelligent agent in smart environment. *ICIES 2015. International Conference on Intelligent Environments and Systems*. <https://doi.org/10.13140/RG.2.1.1242.0008>
- Porra, J., Lacity, M., & Parks, M. S. (2020). Can computer based human-likeness endanger humanness?: A philosophical and ethical perspective on digital assistants expressing feelings they can't have. *Information Systems Frontiers*, 22(3), 533–547. <https://doi.org/10.1007/s10796-019-09969-z>
- Preiß, N., & Westner, M. (2025). Unraveling jingle-jangle fallacies in digital assistant technologies: A comprehensive systematic review and research agenda. *ICICT 2025: 10th International Congress on Information and Communication Technology*, Preprint.
- Priya, P., Firdaus, M., & Ekbal, A. (2023). A multi-task learning framework for politeness and emotion detection in dialogues for mental health counselling and legal aid. *Expert Systems with Applications*, 224, 1–17. <https://doi.org/10.1016/j.eswa.2023.120025>

- Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human Computer Studies*, 151, 24. <https://doi.org/10.1016/j.ijhcs.2021.102630>
- Rashid, A. B., & Kausik, M. A. K. (2024). AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications. *Hybrid Advances*, 7, 1–34. <https://doi.org/10.1016/j.hybadv.2024.100277>
- Richards, D., & Bransky, K. (2014). ForgetMeNot: What and how users expect intelligent virtual agents to recall and forget personal conversational content. *International Journal of Human Computer Studies*, 72(5), 460–476. <https://doi.org/10.1016/j.ijhcs.2014.01.005>
- Rillig, M. C., & Kasirzadeh, A. (2024). AI personal assistants and sustainability: Risks and opportunities. *Environmental Science & Technology*, 58(17), 7237–7239. <https://doi.org/10.1021/acs.est.4c03300>
- Ripa, G., Torre, M., Urbietta, M., Rossi, G., Fernandez, A., Tacuri, A., & Firmenich, S. (2023). Generating voice user interfaces from web sites. *Behaviour and Information Technology*, 43(13), 3129–3152. <https://doi.org/10.1080/0144929X.2023.2272192>
- Ritchey, T. (2011). *General morphological analysis*. 17, 7–18. https://doi.org/10.1007/978-3-642-19653-9_2
- Rittelmeyer, J. D., & Sandkuhl, K. (2023). Amorphological box for AI solutions: Evaluation, refinement and application potentials. In M. Ruiz & P. Soffer (Eds.), *Advanced Information Systems Engineering Workshops* (pp. 5–16). Springer. https://doi.org/10.1007/978-3-031-34985-0_1
- Rosian, M., Hagenhoff, P., & Otto, B. (2021). Towards a holistic cloud computing taxonomy: Theoretical & practical findings. *AMCIS 2021 Proceedings*, 1–10.
- Rubio-Drosdov, E., Díaz-Sánchez, D., Almenárez-Mendoza, F., Cabarcos, P., & Marín-López, A. (2017). Seamless human-device interaction in the internet of things. *IEEE Transactions on Consumer Electronics*, 63, 490–498. <https://doi.org/10.1109/TCE.2017.015076>
- Russell, S. J. (with Norvig, P., & Davis, E.). (2016). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.
- Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Saldana, J. (2021). *The coding manual for qualitative researchers* (1st ed.). Sage.
- Schmidt-Kraepelin, M., Thiebes, S., Tran, M. C., & Sunyaev, A. (2018). What's in the game? Developing a taxonomy of gamification concepts for health apps. *Proceedings of the 51st HICSS*, 51, 1217–1226. <https://doi.org/10.24251/HICSS.2018.150>
- Schuetzler, R. M., Grimes, G. M., & Scott Giboney, J. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal of Management Information Systems*, 37(3), 875–900. <https://doi.org/10.1080/07421222.2020.1790204>
- Sharma, M., Joshi, S., Luthra, S., & Kumar, A. (2024). Impact of digital assistant attributes on millennials' purchasing intentions: A multi-group analysis using pls-sem, artificial neural network and fsQCA. *Information Systems Frontiers*, 26(3), 943–966. <https://doi.org/10.1007/s10796-022-10339-5>
- Shneiderman, B. (2020). *Human-centered artificial intelligence: Reliable, safe & trustworthy* (No. arXiv:2002.04087). arXiv. <http://arxiv.org/abs/2002.04087>
- Šmite, D., Wohlin, C., Galviņa, Z., & Prikladnicki, R. (2014). An empirically based terminology and taxonomy for global software engineering. *Empirical Software Engineering*, 19(1), 105–153. <https://doi.org/10.1007/s10664-012-9217-9>
- Somers, S., Mitsopoulos, K., Lebiere, C., & Thomson, R. (2019). Cognitive-level salience for explainable artificial intelligence. *Proceedings of the 17th Annual Meeting of the International Conference on Cognitive Modeling*, 19–22.
- Stieglitz, S., Mirbabaie, M., Möllmann, N. R. J., & Rzycki, J. (2022). Collaborating with virtual assistants in organizations: Analyzing social loafing tendencies and responsibility attribution. *Information Systems Frontiers*, 24(3), 745–770. <https://doi.org/10.1007/s10796-021-10201-0>

- Szopinski, D., Schoormann, T., & Kundisch, D. (2019). Because your taxonomy is worth it: Towards a framework for taxonomy evaluation. *27th ECIS*.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Valtolina, S., Barricelli, B. R., & Di Gaetano, S. (2020). Communicability of traditional interfaces VS chatbots in healthcare and smart home domains. *Behaviour and Information Technology*, 39(1), 108–132. <https://doi.org/10.1080/0144929X.2019.1637025>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000–6010. <https://doi.org/10.48550/arXiv.1706.03762>
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Wellsandt, S., Hribernik, K., & Thoben, K.-D. (2021). Anatomy of a digital assistant. In A. Dolgui, A. Bernard, D. Lemoine, G. von Cieminski, & D. Romero (Eds.), *Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems* (pp. 321–330). Springer International Publishing. https://doi.org/10.1007/978-3-030-85910-7_34
- Zhu, Z., Branzoi, V., Wolverton, M., Murray, G., Vitovitch, N., Yarnall, L., Acharya, G., Samarasekera, S., & Kumar, R. (2014). AR-mentor: Augmented reality based mentoring system. *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 17–22. <https://doi.org/10.1109/ISMAR.2014.6948404>
- Zumstein, D., & Hundertmark, S. (2017). Chatbots: An interactive technology for personalized communication, transactions and services. *LADIS International Journal on WWW/Internet*, 15, 96–109.

BRIDGING CYBER RESILIENCE AND IT BUSINESS VALUE: A LONGITUDINAL BIBLIOMETRIC ANALYSIS

ZIGGY VAN GIEL,¹ TIM HUYGH,² ANANT JOSHI,³
STEVEN DE HAES¹

¹ University of Antwerp, Faculty of Business and Economics, Antwerp, Belgium
ziggy.vangiel@uantwerpen.be, steven.dehaes@uantwerpen.be

² Open University, Faculty of Management, Science & Technology, Limburg,
the Netherlands
tim.huygh@ou.nl

³ University of Maastricht, School of Business and Economics, Limburg, the Netherlands
a.joshi@maastrichtuniversity.nl

Cyber resilience is often defined as the ability of an organization to continuously deliver the intended outcome despite adverse cyber events. While this definition takes the business and IT business value as a starting point, literature mainly focuses on technical or operational aspects of critical infrastructure or supply chains. This paper uses a longitudinal bibliometric analysis to analyze trends in both the current body of knowledge and the past body of knowledge regarding cyber resilience. The study shows that the domain shows a clear interest in emerging technologies (e.g., AI). However, it lacks fundamental conceptual research that tries to integrate cyber resilience with IT business value and the broader IT governance literature. To conclude, the results suggest that the domain would benefit from research that focuses on the business side of cyber resilience and the IT value it hopes to protect, instead of focusing on technical measures or sector-specific research.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.2](https://doi.org/10.18690/um.fov.4.2025.2)

ISBN

978-961-286-998-4

Keywords:

IT business value,
cyber resilience,
digital transformation,
bibliometric analysis,
IT governance



University of Maribor Press

1 Introduction

Organizations increasingly depend upon information technology (IT) for the achievement of their business goals, which increases the importance of cyber resilience to ensure continuous delivery of IT business value. While there appears to be an increasing interest in cyber resilience, the literature on resilience in organizational sciences remains fragmented. One reason for this is that resilience is often applied on a specific problem rather than focusing on creating conceptual clarity first (Linnenluecke, 2017). Indeed, cyber resilience research is still in its infancy (Bellini & Marrone, 2020; Eling et al., 2021; Linnenluecke, 2017). Given its fragmented body of knowledge, the cyber resilience domain would benefit from an analysis of the main research streams. This would enable future research to focus on integrating the fragmented research streams to provide conceptual clarity and improve overall maturity of the domain. Therefore the first research question aims to provide an overview of these research streams within organizational cyber resilience to enable the integration of the fragmented body of knowledge. *RQ1: What are dominant themes in the current literature on organizational cyber resilience? And what publications were of most influence?*

IT investments are driven by the potential realization of IT business value through benefits (e.g., increasing customer satisfaction, improved lead generation ...), cost reduction, or risk minimization (Gartner, 2023). Cyber resilience, defined as “*the ability to continuously deliver the intended outcome despite adverse cyber events*” (Bjorck et al., 2015), has traditionally been framed in relation to cybersecurity and IT risk management. First, while cybersecurity focuses on known complex threats, cyber resilience is said to focus on unpredictable and unknown threats to guarantee business continuity (Baikloy et al., 2020; Galinec & Steingartner, 2017). Next, Eling et al. (2021) mentioned that cyber resilience could be considered the next maturity level of IT risk management. Compared to other concepts, cyber resilience has been proposed to take the business as a starting point instead of focusing on technical approaches (Bjorck et al., 2015; Garcia-Perez et al., 2023). While some cyber resilience literature focuses on business and organizational aspects (e.g., Bagheri et al. (2023)), it is unclear whether the current body of knowledge has framed cyber resilience in relation to IT business value literature. However, the domain would benefit from a clear direct conceptual relationship between IT business value and cyber resilience. As a result, this paper aims to increase conceptual clarity by explicitly

linking organizational cyber resilience to the continuous realization of IT-enabled business value. Hence the second research question focuses on exploring the extent to which this perspective is present or is being integrated in the current body of knowledge. *RQ2: To what extent is IT business value considered in the current literature on organizational cyber resilience?*

The remainder of this paper is structured as follows. Section 2 provides theoretical background on both IT business value and organizational cyber resilience with a specific focus on their conceptual definitions and operationalizations. Next, section 3 presents the methodology used for this paper. Section 4 then presents the findings of this research. Before concluding in section 6, the discussion in section 5 reflects on the implications and limitations of this paper while also presenting avenues for future research.

2 Theoretical Background

This section presents the results of a qualitative analysis of existing literature on both IT business value and organizational cyber resilience. While the former is based on an ad-hoc literature review, the cyber resilience section is based on a qualitative analysis of literature that was identified in 2024 for the bibliometric analysis.

2.1 IT Business Value

IT business value requires the continuous alignment of business and IT on a strategic, operational and structural level (Maes, 1999), which in turn is enabled through an appropriate IT governance (De Haes et al., 2020). As such, IT governance's final aim is the delivery of IT business value, while mitigating IT-related risks (Parent & Reich, 2009). Depending on how IT business value is defined, controlling IT risks might be part of IT business value. In this context, Gartner (2023) states that business value of IT can be realized in three different ways: increasing revenue, improving cost-efficiency, or mitigating risks. As such, IT business value can be seen as the impact IT has on those three different dimensions (i.e., benefits, costs, and risks). As IT risk mitigation can be considered a dimension of IT business value, the question becomes what a proper definition for IT business value would be.

There exist numerous definitions of business value in the context of IT. For example, Melville et al. (2004, p. 287) defined IT business value as “[...] *the organizational performance impacts of information technology at both the intermediate process level and the organization wide level, and comprising both efficiency impacts and competitive impacts.*” While this definition only stresses the impact on efficiency and competitiveness, mitigation of IT risks can be considered an investment in sustaining competitiveness. Next, Schryen (2013, p. 141) defined information system (IS) business value as “[...] *the impact of investments in particular IT assets on the multidimensional performance and capabilities of economic entities at various levels, complemented by the ultimate meaning of performance in the economic environment.*” This definition stresses the impact investments in IT might have on performance and business capabilities. Finally, Riera and Iijima (2019) defined digital business value as “[...] *the level of achievement of business objectives using information technologies.*” This definition focuses on the link between the realization of business objectives and IT. To conclude, these definitions establish a link between IT investments and either organizational performance or the realization of business objectives.

The current definitions of IT business value highlight a proportional impact of IT on performance or on the achievement of business objectives. However, measuring IT business value has proven to be difficult. It is namely highly contingent upon organizational characteristics and the type of IT investment. In this context, Weill and Broadbent (1998) discussed the business value hierarchy, which is visualized in Figure 1. They showed that it will be difficult to measure IT business value by using high-level measures such as revenue growth. This aligns with the findings of Davern and Wilkin (2010) that higher-level measures are inadequate to measure IT business value effectively. Using inappropriate measures of IT business value might risk false statements, which has been called the ‘IT productivity paradox’ (Brynjolfsson, 1993; Davern & Wilkin, 2010; Schryen, 2013). For example, using high-level measures such as financial metrics to evaluate IT business value introduces dilution (Weill & Broadbent, 1998) and neglects non-monetary value (e.g., employee/user satisfaction, improved collaboration). Just as with measures for IT business value, IT risks and their mitigation strategies need to be considered on the appropriate level (e.g., application-level versus organization-level).

There is however no one-size-fits-all for IT business value and managing IT risks. In this regard, Schryen (2013) mentioned three types of contextual factors: macro-economic, industry, and firm contextual factors. Similarly, Melville et al. (2004) distinguished between characteristics in the macro environment, competitive environment, and of the firm itself. In the context of industry contextual factors, McAfee and Brynjolfsson (2008) analyzed the competitive dynamics during a period of surging IT investments from the 1990s onwards. They found that turbulence is the highest in IT intensive industries. Because of this, Zmud and Sambamurthy (2012) concluded that organizations that invest heavily in IT, are operating in more turbulent business environments. Conversely, organizations that are operating in relatively stable environments do not need to invest heavily in IT. The environment in which an organization is operating clearly has an impact on IT business value by dictating the intensity and type of IT investments.

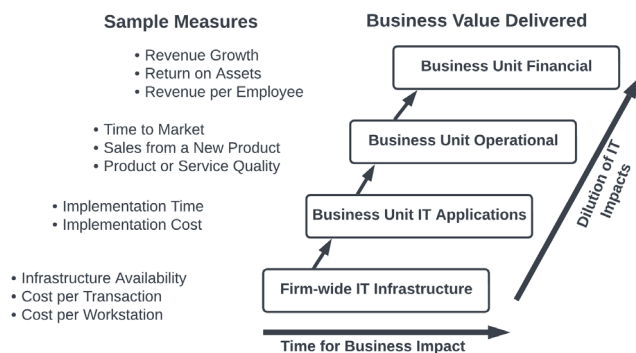


Figure 1: The Business Value Hierarchy

Source: Reproduced from Weill and Broadbent (1998)

While the external environment is an important factor influencing IT business value, also organization specific characteristics are important to consider. As mentioned before, the alignment between business and IT strategy has been considered an important precondition to realizing IT business value (e.g., (De Haes et al., 2020; Maes, 1999)). The business strategy, the accompanying business objectives, and the way in which IT supports the achievement of those objectives are important factors to consider when looking at IT business value and IT risks (Riera & Iijima, 2019).

Both perspectives (i.e., external environment factors, and organizational factors) can be combined in, what Nolan and McFarlan (2005) called, the IT strategic impact grid, as visualized in Figure 2. This model differentiates organizations based on the need for new IT, which is said to be dictated by market pressures (i.e., the external environment), and based on the dependency upon reliable IT, which is driven by the functioning and interdependence of business and IT. This model is also adopted by ISACA (2018) as a design factor in their COBIT 2019 on the IT governance operating model. Finally, from an IT risk perspective, the strategic role of IT will influence the importance of different types of IT risks (e.g., operational IT risks versus strategic IT risks, project risks versus business continuity risks etc.). As such, not every company might benefit from incorporate the same (cyber) resilience perspective (e.g., organizations that are in turnaround mode might not benefit from resilience as their dependence upon reliable IT is relatively low).

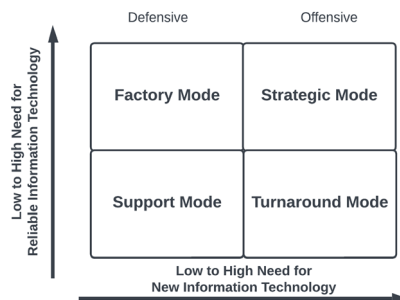


Figure 2: The IT Strategic Impact Grid

Source: Reproduced from (Nolan & McFarlan, 2005)

2.2 Organizational cyber resilience

Organizational cyber resilience has been defined in numerous ways. Firstly, one widely used definition of cyber resilience focuses on four events stages: (1) plan and prepare, (2) absorb, (3) recover, and (4) adapt and learn (Linkov et al., 2013; Linkov & Kott, 2019). It can be defined as “the ability of an organization to plan and prepare for, respond to, recover from, and adapt to a cyber attack” (Annarelli & Palombi, 2021; Hausken, 2020; Onwubiko, 2020). Similarly, Bagheri et al. (2023) state cyber resilience consists out of “anticipation, support, recovery and adaptation in a changing environment”. The goal is then to improve the organization’s capability of

facing adverse situations (Carias et al., 2020). While the event-based definitions of cyber resilience offer valuable insights, their approach is too narrow. An incident or crisis can namely be considered as process rather than an event (Williams et al., 2017). From this perspective, adverse events are solely events that manifested because of a process that preceded the event (e.g., mismanagement of risks, inadequate governance of an increasing reliance upon specific information systems).

Instead of an event-based approach, process-based approaches embed resilience thinking in the culture of an organization to build a sustainable business model (Sarkar et al., 2016). Cyber resilience then becomes “*a function of an organization’s situation awareness [...], management of [...] vulnerabilities, adaptive capacity, risk intelligence, flexibility and agility [...] in a complex, dynamic, and interconnected environment*” (Sarkar & Wingreen, 2015; Sarkar et al., 2016). It becomes clear that cyber resilience extends beyond the technical issues to include behavioral and organizational aspects (Bagheri et al., 2023). Instead of the technical domain, a holistic cyber resilience approach considers the physical, information, cognitive and social domains (Garcia-Perez et al., 2023; Linkov & Kott, 2019). It involves coordinated efforts on organizational, technological, and human factors (Safitra et al., 2023). Also, scholars argue that cyber resilience should be included in and aligned with the overall business strategy (Galinec & Steingartner, 2017; Sarkar et al., 2016). As an incident should be accepted as a likely event, cyber resilience should be considered a long-term endeavor that benefits from technology-neutral policy actions (Greiman, 2023). For that reason cyber resilience is not limited to specific processes (e.g., IT service management, IT risk management ...), nor is it limited to a specific organizational sub-unit. Finally, cyber resilience is defined as “*the ability to continuously deliver the intended outcome despite adverse cyber events*” (Bjorck et al., 2015). Indeed, cyber resilience aims at sustaining the delivery of IT business value while acknowledging the potential impact of disruptions, which necessitates an organization-wide approach.

When combining the above definitions, a nuanced view on cyber resilience is presented. First, cyber resilience takes the business as a starting point by ensuring the ability to deliver value. Second, to achieve that a holistic perspective view on the organizations should be taken which extends beyond the technical. Finally, by acknowledging the potential impact of IT-related risks, cyber resilience aims to make the organization able to withstand both known and unknown disruptions. While the existing literature on cyber resilience does not explicitly incorporate IT business

value, there is a clear link through their conceptualizations. Therefore, future conceptualizations and definitions of cyber resilience could benefit from the explicit integration of IT business value.

3 Methodology

This paper uses the process described by Zupic and Cater (2015) on bibliometric methods to analyze the current trends in cyber resilience research while considering the potential evolution towards IT business value. Using “cyber resilien*” as keyword relevant literature was identified from the Web of Science database. After an initial search that included synonyms (e.g., information systems resilience, organizational resilience of IT), the decision was made to not include synonyms. The number of papers that were not identified because of this was limited. However, the improved data quality and reduced need for manual screening outweighed the additional quantity of papers while increasing replicability. In Web of Science the results were refined to only include publications from the SCI-EXPANDED and SSCI indexes, and to only include articles, review articles and early access. This search was evaluated on February 6th, 2024, and re-evaluated on February 7th, 2025. The former yielded 195 results and the latter 311 results. The results of 2024 were also refined to only include references that focus on organizational aspects of cyber resilience as compared to other aspects of cyber resilience (e.g., resilience of individual applications, resilience of urban communities ...). By doing that the final number of publications included in the final analysis was 66 instead of the original 195. That refinement enables an adequate analysis of the most influential references and sources based on relevant literature, while the unrefined set from 2025 gives a clear view on the different research streams that relate to cyber resilience.

As we are interested in the potential integration of IT business value perspectives in the cyber resilience domain, a co-word analysis is used to analyze the co-occurrence of bi-grams in the abstracts of the publications through the application of the Louvain clustering algorithm. This is repeated for both the 2024 and 2025 dataset to uncover potential evolution over time. Next, based on the refined set of 2024 the most local cited references, and the most relevant sources (i.e., journals or conferences) are presented.

4 Results

4.1 Trends in cyber resilience literature

When looking at the co-occurrence network of the 2024 dataset (see Figure 3), four different clusters can be identified: (1) a cluster focusing on cyber resilience of critical infrastructure (purple), (2) a cluster focusing on cyber resilience of supply chains (blue), (3) a cluster linking cyber resilience to IT risk management and cybersecurity (green), and (4) a cluster that focuses on the external perspective of cyber resilience in terms of external threats, incidents, cyberattacks and preventive measures to deal with those.

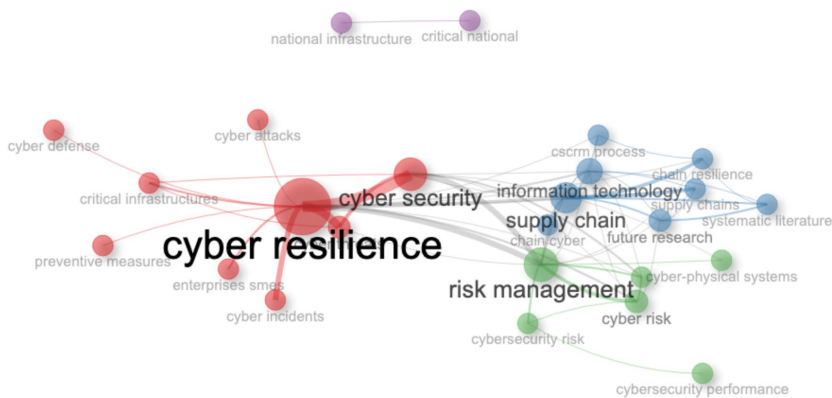
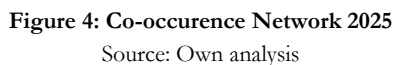


Figure 3: Co-occurrence Network 2024

Source: Own analysis

Comparing these to the co-occurrence network of 2025 we again can identify four clusters (see Figure 4). Firstly, a cluster of research focuses on the cyber resilience of the energy grid and power systems (green). Secondly, the main cluster of research focuses on cybersecurity and supply chain related aspects (red). Thirdly, a cluster focuses on operational aspects of cyber resilience monitoring such as anomaly detection and injection attacks (purple). Finally, one cluster focuses on the application of emerging technologies such as artificial intelligence, machine learning, and deep learning in the context of cyber resilience and industrial control systems.



By using the most local cited references, and the most relevant sources, we can identify the influence of individual publications and identify seminal articles that lie at the basis of the domain. In Table 1 we can see that the publications of Bjorck et al. (2015), Linkov et al. (2013), and Linkov and Kott (2019) were the most influential for the cyber resilience domain. While the former presented a clear definition of cyber resilience, which was also used in this paper, the latter outlined the cyber resilience matrix that combines different event-phases of resilience (i.e., plan/prepare, absorb, recover, and adapt/learn) with four different domains (i.e., physical, information, cognitive, and social).

Reference	Number of Citations
(Bjorck et al., 2015)	13
(Linkov et al., 2013)	12
(Linkov & Kott, 2019)	11
(Boyes, 2015)	9
(Bodeau & Graubart, 2011); (Davis, 2015); (von Solms & van Niekerk, 2013)	6

Source: Own analysis

Next, Table 2 provides an overview of the most important sources for cyber resilience literature. 7 out of the 66 publications originate from Computer & Security journal. Next, IEEE Access and Supply Chain Management complete the top three of most relevant sources.

Table 2: Most relevant sources

Source	Number of documents
Computer & Security (ISSN: 01674048)	7
IEEE Access (ISSN: 21693536)	5
Supply Chain Management: An international journal (ISSN: 13598546)	4
Applied Sciences – Basel (ISSN: 20763417)	3
Sustainability (ISSN: 20711050)	3

Source: Own analysis

Combining the above trends and the most influential references and sources, we can see a focus on technical cyber resilience publications (security, monitoring, applications of emerging technologies such as AI), on risk-related aspects (risk management, external threats), and on specific sectors (supply chain, critical infrastructure, energy and power). Next to these main trends there are a limited number of references that were of significantly higher influence. The publications of Bjorck et al. (2015), Linkov et al. (2013), and Linkov and Kott (2019) were identified as most influential. These papers were conceptual papers that were already discussed in section 2. As concluded then, although there is a conceptual link, IT business value is not explicitly incorporated in cyber resilience literature.

5 Discussion

This paper started from two main research questions that focused on uncovering trends within cyber resilience literature and exploring the integration of the IT business value perspective. Based on the bibliometric analysis and the theoretical background that was provided before we indeed observe that literature applies cyber resilience to specific problems (e.g., specific sectors, specific external risks ...), as mentioned by Linnenluecke (2017), instead of providing conceptual clarity first. While the most cited references were conceptual papers, they could be complemented by alternative perspectives. For example, instead of the event-based approach to cyber resilience as presented by (Linkov et al., 2013; Linkov & Kott, 2019), future research could explore a process-based perspective on cyber resilience.

Next, the definitions and conceptualizations of cyber resilience claim to take the business as a starting point (Bjorck et al., 2015; Garcia-Perez et al., 2023), and aim to enable continuous delivery of the intended outcome (i.e., continuous delivery of IT-enabled value) (Bjorck et al., 2015). However, the current literature does not reflect those perspective, evidenced by the lack of research trends in that direction and by the perspectives presented in the most cited conceptual papers. Indeed, there is a mismatch between the conceptual definition of cyber resilience and how it is used in academic literature. Therefore, future research could seek to integrate IT business value explicitly into cyber resilience conceptualizations in order to integrate the domains.

Finally, there is a stream of research focusing on IT business value that could be used to push the domain forward. Future research should take the business as a starting point when using cyber resilience and try to incorporate the final objective of cyber resilience on an organizational level: preserving IT business value. Also, past research has stated that cyber resilience is still in its infancy and lacks conceptual clarity (Bellini & Marrone, 2020; Eling et al., 2021; Linnenluecke, 2017). Nevertheless, this study has shown that there exist seminal articles that provides this conceptual clarity, but most papers neglect the broad perspective of the concept. Because of this, there is a clear need for research that applies cyber resilience holistically to an organizational context and takes the business as a starting point.

This research has different limitations which also offer opportunities for future research. Firstly, the scope is limited by a focus on IT specifically. Future research could incorporate OT as it significantly impacts the dependency upon reliable IT and cyber resilience. Next, the temporal difference shows some evolution in terms of trends. Nevertheless, future research could use a longer time horizon to uncover evolutions in the domain. Finally, the literature was selected only based on Web of Science and on only a single keyword. Future research could broaden this scope to incorporate more diverse perspectives.

6 Conclusions

This paper provided a theoretical background on IT business value and organizational cyber resilience. It claimed that there is a need to integrate IT business value because of their conceptual overlap. Next, it analyzed whether this business-

first perspective is being used in contemporary research on organizational cyber resilience. While we expected to see the emergence of a trend that focuses on IT business value or related concepts, the past and current literature mainly focuses on technical aspects and emerging technologies (e.g., AI) in that context. Although cyber resilience is claimed to take the business as a starting point, as compared to cybersecurity, the literature does not reflect this evolution.

Acknowledgements

Special thanks to the two anonymous reviewers for their valuable input. Additionally, Flanders Innovation & Entrepreneurship is acknowledged together with AE, as the lead author (Ziggy Van Giel) is being funded through a Baekeland mandate at the time of conducting this research.

References

- Annarelli, A., & Palombi, G. (2021). Digitalization Capabilities for Sustainable Cyber Resilience: A Conceptual Framework. *Sustainability*, 13(23). <https://doi.org/10.3390/su132313065>
- Bagheri, S., Ridley, G., & Williams, B. (2023). Organisational Cyber Resilience: Management Perspectives [Article]. *Australasian Journal of Information Systems*, 27, 28. <https://doi.org/10.3127/ajis.v27i0.4183>
- Baikloy, E., Praneetpolgrang, P., & Jirawichitchai, N. (2020). Development of Cyber Resilient Capability Maturity Model for Cloud Computing Services [Article]. *Tem Journal-Technology Education Management Informatics*, 9(3), 915-923. <https://doi.org/10.18421/tem93-11>
- Bellini, E., & Marrone, S. (2020, 2020). Towards a novel conceptualization of Cyber Resilience.
- Bjorck, F., Henkel, M., Stirna, J., & Zdravkovic, J. (2015). Cyber Resilience - Fundamentals for a Definition. *New Contributions in Information Systems and Technologies, Vol 1, Pt 1*, 353, 311-316. https://doi.org/10.1007/978-3-319-16486-1_31
- Bodeau, D. J., & Graubart, R. (2011). *Cyber Resiliency Engineering Framework*. https://www.mitre.org/sites/default/files/media/publication/11_4436_2.pdf
- Boyes, H. (2015). Cybersecurity and Cyber-Resilient Supply Chains. *Technology Innovation Management Review*, 5(4), 28-34. <http://timreview.ca/article/888>
- Brynjolfsson, E. (1993). The Productivity Paradox of Information Technology. *Communications of the ACM*, 36(12), 67-77. <https://doi.org/10.1145/163298.163309>
- Carias, J. F., Arrizabalaga, S., & Hernantes, J. (2020, Dec 03-04). Cyber Resilience Strategic Planning and Self-assessment Tool for Operationalization in SMEs. *IFIP Advances in Information and Communication Technology* [Information technology in disaster risk reduction, itdrr 2020]. 5th IFIP WG 5.15 International Conference on Information Technology in Disaster Risk Reduction (ITDRR), Univ Natl & World Econ, ELECTR NETWORK.
- Davern, M. J., & Wilkin, C. L. (2010). Towards an integrated view of IT value measurement. *International Journal of Accounting Information Systems*, 11(1), 42-60. <https://doi.org/10.1016/j.accinf.2009.12.005>
- Davis, A. (2015). Building Cyber-Resilience into Supply Chains. *Technology Innovation Management Review*, 5(4), 19-27. <http://timreview.ca/article/887>
- De Haes, S., Van Grembergen, W., Joshi, A., & Huygh, T. (2020). *Enterprise Governance of Information Technology: Achieving Alignment and Value in Digital Organizations* (Third Edition ed.). Springer.
- Eling, M., McShane, M., & Nguyen, T. (2021). Cyber risk management: History and future research directions. *Risk Management and Insurance Review*, 24(1), 93-125. <https://doi.org/10.1111/rmir.12169>

- Galinec, D., & Steingartner, W. (2017). Combining Cybersecurity and Cyber Defense to achieve Cyber Resilience. *2017 IEEE 14th International Scientific Conference on Informatics*, 87-93. <https://doi.org/10.1109/INFORMATICS.2017.8327227>
- Garcia-Perez, A., Sallos, M. P., & Tiwasing, P. (2023). Dimensions of cybersecurity performance and crisis response in critical infrastructure organisations: an intellectual capital perspective [Article; Early Access]. *Journal of Intellectual Capital*, 24(2), 465-486. <https://doi.org/10.1108/jic-06-2021-0166>
- Gartner. (2023). *Proving the Business Value of IT*. <https://www.gartner.com/en/information-technology/topics/business-value-of-it>
- Greiman, V. A. (2023, Mar 09-10). Nuclear Cyber Attacks: A Study of Sabotage and Regulation of Critical Infrastructure. [Proceedings of the 18th international conference on cyber warfare and security iccws]. 18th International Conference on Cyber Warfare and Security (ICCWS), Towson Univ, Baltimore, MD.
- Hausken, K. (2020). Cyber resilience in firms, organizations and societies [Review]. *Internet of Things*, 11, 9. <https://doi.org/10.1016/j.iot.2020.100204>
- ISACA. (2018). COBIT 2019 framework: Introduction & methodology.
- Linkov, I., Eisenberg, D. A., Plourde, K., Seager, T. P., Allen, J., & Kott, A. (2013). Resilience metrics for cyber systems. *Environment Systems and Decisions*, 33(4), 471-476. <https://doi.org/10.1007/s10669-013-9485-y>
- Linkov, I., & Kott, A. (2019). Fundamental Concepts of Cyber Resilience: Introduction and Overview. In (pp. 1-25). Springer International Publishing. https://doi.org/10.1007/978-3-319-77492-3_1
- Linnenluecke, M. K. (2017). Resilience in Business and Management Research: A Review of Influential Publications and a Research Agenda [Review]. *International Journal of Management Reviews*, 19(1), 4-30. <https://doi.org/10.1111/ijmr.12076>
- Maes, R. (1999). Reconsidering Information Management Through A Generic Framework.
- McAfee, A., & Brynjolfsson, E. (2008). Investing in the IT that makes a competitive difference. *Harvard Business Review*, 86(7-8), 98-+. <Go to ISI>://WOS:000257047500024
- Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organizational performance: An integrative model of IT business value. *Mis Quarterly*, 28(2), 283-322. <https://doi.org/10.2307/25148636>
- Nolan, R., & McFarlan, F. (2005). Information technology and the board of directors. *Harvard Business Review*, 83(10), 96-106, 157. <https://www.ncbi.nlm.nih.gov/pubmed/16250628>
- Onwubiko, C. (2020, Jun 15-19). Focusing on the Recovery Aspects of Cyber Resilience. 2020 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA), Dublin, Ireland.
- Parent, M., & Reich, B. H. (2009). Governing information technology risk. *California Management Review*, 51(3), 134-152. <https://doi.org/10.2307/41166497>
- Riera, C., & Iijima, J. (2019). The Role of IT and Organizational Capabilities on Digital Business Value. *Pacific Asia Journal of the Association for Information Systems*, 67-95. <https://doi.org/10.17705/1pais.11204>
- Safitra, M. F., Lubis, M., & Fakhurroja, H. (2023). Counterattacking Cyber Threats: A Framework for the Future of Cybersecurity [Article]. *Sustainability*, 15(18), 32, Article 13369. <https://doi.org/10.3390/su151813369>
- Sarkar, A., & Wingreen, S. (2015, Sep 21-22). How CEOs of the Small Firms Make Decisions to Ensure Information Systems Resilience? *Proceedings of the European Conference on Information Management and Evaluation* [Proceedings of 9th european conference on is management and evaluation (ecime 2015)]. 9th European Conference on Information Management and Evaluation (ECIME), Univ W England, Bristol, ENGLAND.
- Sarkar, A., Wingreen, S., Ascroft, J., & Assoc Informat, S. (2016). Top Management Team Decision Priorities to Drive IS Resilience: Lessons from Jade Software Corporation. [Amcis 2016 proceedings]. 22nd Americas Conference on Information Systems (AMCIS), San Diego, CA.

- Schryen, G. (2013). Revisiting IS business value research: What we already know, what we still need to know, and how we can get there. *European Journal of Information Systems*, 22(2), 139-169. <https://doi.org/10.1057/ejis.2012.45>
- von Solms, R., & van Niekerk, J. (2013). From information security to cyber security. *Computers & Security*, 38, 97-102. <https://doi.org/10.1016/j.cose.2013.04.004>
- Weill, P., & Broadbent, M. (1998). *Leveraging the new infrastructure—How market leaders capitalize on Information Technology*. Harvard Business School Press.
- Williams, T. A., Gruber, D. A., Sutcliffe, K. M., Shepherd, D. A., & Zhao, E. Y. F. (2017). Organizational Response to Adversity: Fusing Crisis Management and Resilience Research Streams. *Academy of Management Annals*, 11(2), 733-769. <https://doi.org/10.5465/annals.2015.0134>
- Zmud, R. W., & Sambamurthy, V. (2012). *Guiding the digital transformation of organizations*. Legerity Digital Press.
- Zupic, I., & Cater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods*, 18(3), 429-472. <https://doi.org/10.1177/1094428114562629>

DIGITALIZATION AS A CATALYST FOR RESILIENCE IN THE HOSPITALITY SECTOR

ALINDA KOKKINO, MICHEL ALTAN, ABDON DANTAS,
ONDREJ MITAS

Breda University of Applied Sciences, Breda, the Netherlands
kokkinou.a@buas.nl, altan.m@buas.nl, dantas.a@buas.nl, mitas.o@buas.nl

Hospitality organizations are particularly vulnerable to changes in demand caused by disruptive events such as natural catastrophes, geopolitical events, and pandemic diseases. Nevertheless, the development of organizational resilience by hospitality organizations has remained under-explored. The ongoing digitalization trend provides a unique opportunity for hospitality organizations to combine the adoption of digitalization tools with the development of data analytic capability as a way to anticipate disruptive events and mitigate their impact on operations and performance. Through a cross-sectional survey design and using Partial Least Square Structural Equation Modeling, the present study demonstrates that hospitality organizations can improve their organizational resilience by developing data analytic capability. This can be achieved by (1) investing in the digital tools and IT infrastructure that allows them to sense their environment and (2) adapting their organizational infrastructure to quickly be able to use this information in decision-making. A limitation of the study lies in the use of cross-sectional data which limits temporal causality inferences.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.3](https://doi.org/10.18690/um.fov.4.2025.3)

ISBN
978-961-286-998-4

Keywords:
hospitality,
services,
resilience,
digitalization,
data analytic capability



University of Maribor Press

1 Introduction

Service organizations, and hospitality organizations in particular (e.g. hotels, restaurants, campsites), are vulnerable to disruptions (He et al. 2023). Due to their fixed capacity, simultaneous production and consumption, perishable inventory, and high fixed cost/ low variable cost structure, hospitality organizations have difficulty in absorbing the large shifts in customer demand that can be caused by disruptions (Kokkinou et al. 2023c). COVID-19 was the most recent high-impact disruption affecting the hospitality industry (Gursoy and Chi 2020). However, while the COVID-19 pandemic was unprecedented, hospitality organizations have a long history of vulnerability to a variety of disruptions such as disease outbreaks, terrorism attacks, and natural disasters such as earthquakes or floods (Hall et al. 2023).

These disruptive events have accentuated the importance of resilience for hospitality organizations. Resilience is defined as an organization's ability to *either* absorb shocks and continue operating, *or* bounce back from these shocks in a short enough time that there is no lasting effect on its performance (Hall et al. 2023). Resilience has received extensive attention in the tourism literature (Lew 2014; Luthe and Wyss 2014; Jiang et al. 2019; Usher et al. 2019), where it is typically treated from a systems perspective (e.g. destination level and/or ecosystem). Hospitality organizations are typically considered as part of such a tourism ecosystem, as they act both as an enabler of tourism and a beneficiary thereof (Melían-Alzola et al. 2020). Nevertheless, in the context of hospitality research, resilience is typically examined from an organizational perspective, using the organization as the level of analysis.

On an organizational level, resilience is seen as a capability that organizations can develop through investments in assets and organization-specific processes (Jiang et al. 2019). The resilience capability is further enabled by other capabilities, such as organizational flexibility, organizational agility, and organizational learning. In other industries, a promising avenue for the development of resilience has been shown to be the development of Data Analytic Capability (DAC) (Kokkinou et al. 2023b). This ability to quickly deploy data, people, and technology allows organizations to generate the insights needed for complex decision-making. For hospitality organizations this might include a better understanding of market segments, how to allocate scarce capacity to them, and where to find additional sources of business

and revenue in crisis situations. Nevertheless, developing DAC requires significant investments in acquiring new knowledge, developing the necessary IT infrastructure and organizational infrastructures and moving the organizational culture towards a more data driven approach (Kokkinou et al. 2023a).

Ongoing digitalization is encouraging hospitality companies to embrace technological innovations as a way to remain competitive (Law et al. 2022), and thus also improve the resilience of their organizations. Nevertheless, and despite the rising importance of digitalization, the role that digital technologies and data analytics can play in developing resilience remains limited (He et al. 2023). The purpose of the present study is to examine how hospitality organizations can harness the opportunities provided by digitalization to develop their resilience. Our study thereby contributes to the growing body of literature examining how hospitality organizations can develop their resilience capabilities through a combination of assets and organization-specific processes.

The next section reviews existing literature on resilience capabilities, with special attention to resilience research in hospitality and tourism. We subsequently review literature on digitalization, examining the role that DAC plays in developing resilience, and which organizational assets and processes are needed for this, leading to the study hypotheses and conceptual model. The third and fourth sections describe the sampling, data collection, and data analysis procedures used to test the study hypotheses. The fourth section discusses the findings of the study and compares them to literature. We conclude with the implications for practice, a review of the limitations of the study and suggestions for future research.

2 Review of the Literature

2.1 Resilience as an Organizational Capability

Several conceptualizations of resilience exist. Hall et al. (2023) distinguishes between socio-ecological resilience, ecological resilience, and engineering resilience. Socio-ecological resilience is applicable when changes happen at the system level, while ecological resilience is primarily concerned with changes at the sectoral level. From this perspective, the ongoing digitalization trend can be seen as a change that affects the hospitality industry in its entirety. The present study adopts the engineering

perspective of resilience, viewing it as an organization's ability to return to its previous state after a disruption (Hall et al. 2023). This conceptualization of resilience is also consistent with how other industries view organizational and/or supply chain resilience (Dubey et al. 2021; Iftikhar et al. 2024).

Through the lens of Dynamic Capability Theory (DCT), resilience is perceived to be an organizational capability that (1) allows organizations to sense their environment and identify opportunities and/or threats, (2) seize opportunities or neutralize threats, and (3) transform themselves to remain aligned with their changing environment (Teece et al. 1997; Teece 2007; Helfat and Raubitschek 2018). This conceptualization has also been adopted in the context of tourism where Jiang et al. (2019) argued that tourism organizations build resilience by using dynamic capabilities and slack resources to transform their operational routines into new ones, more resilient to disruptions. Organizations can develop their resilience capability by investing in other capabilities, assets and company-specific routines.

While the scope of the present study is on organizational resilience, resilience can also be seen from the individual and group perspectives (Sutcliffe and Vogus 2003). At the individual level, employees contribute to the resilience of the organization by identifying potential threats early on, and/or providing creative and flexible solutions (He et al. 2023). At the group level, teams can perform better in adjusting to new conditions when they oriented towards acquiring new knowledge and skills (Sutcliffe and Vogus 2003). In the context of individuals at work, resilience is enhanced when they have access to resources and perceive to have the needed expertise and efficacy. At the organizational level, resilience can be achieved when the organization has sufficient information to manage vulnerabilities and/or adjust its direction in case of a disruption (He et al. 2023).

Across industries, antecedents of organizational resilience have included (1) other capabilities such as resourcefulness, dynamic competitiveness, organizational learning, organizational culture; (2) assets such as cash flow, skilled employees, and (3) organization-specific processes and routines such as networks, leadership, and operational flexibility (Pal et al. 2014; Cotta and Salvador 2020; Vakilzadeh and Haase 2021). In the context of hospitality organizations, Melián-Alzola et al. (2020) further identified strategy and culture as antecedents of resilience.

More recently, ongoing digitalization across industries has drawn attention to the potential of digital technologies to improve organizational and supply chain resilience by improving the information flows and reducing uncertainty across the organization (Jia et al. 2020).

2.3 Digitalization

Broadly speaking, digitalization refers to society's transition from analog to digital, and the corresponding changes in customer and organizational behavior (Vial 2019). Digitalization is contributing to a shift in customer behavior and expectations. Social media, Virtual Reality (VR) and Artificial Intelligence (AI)-enabled chatbots are diversifying the information sources that customers can use to form service expectations. Concurrently, novel digital technologies are providing hospitality companies with innovative ways to meet these expectations by customizing service experiences and empowering customers (Neuhofer et al. 2015). Digitalization is also a source of competitive advantage for organizations that can successfully transform their operating model. Organizations are harnessing the power of digital technologies to develop new IT capabilities, using them to design new digital strategies (Bharadwaj et al. 2013).

Digitalization has been linked to improved organizational and supply chain resilience (Dubey et al. 2021; He et al. 2023). In the context of hospitality organizations, He et al. (2023) demonstrated that a hospitality's digital maturity was positively associated with its organizational resilience. Nevertheless, given how broad the concept of digitalization is, further elaboration is needed to better understand how digitalization contributes to organizational resilience. In the section below we focus on an IT capability, DAC (Wamba et al. 2017), that has been linked to supply chain resilience (Yu et al. 2021; Kokkinou et al. 2023b), and the organizational assets and processes that are known to contribute to it.

2.3.1 Data Analytic Capability (DAC)

In the context of DCT, DAC can also be seen as an organizational capability. DAC enables organizations to identify new demand and consumer behavior patterns (e.g., changes in booking patterns, evolution in consumer preferences and requirements), thereby improving their ability to *sense* their environments (Mikalef et al. 2019). DAC

furthermore facilitates organizations in *seizing* these opportunities by fueling the use of data-driven decision-making when assessing business opportunities and potential investment opportunities. Finally, DAC supports organizations in *transforming and reconfiguring* their activities as it pushes the organization culture towards becoming more data-driven (Kokkinou 2024). DAC has been shown to be a source of competitive advantage, particularly in uncertain environments (Wamba et al. 2017). DAC is also commonly viewed as enabling other capabilities (Helfat and Peteraf 2009), including organizational and supply chain resilience.

DAC has been linked to organizational and supply chain resilience in a variety of industries including manufacturing and supply chains (Dubey et al. 2021). In this context, DAC helps organizations to increase their capacity to process information and generate the insights that help them mitigate risks associated with disruptions (Iftikhar et al. 2024). This finding can extend to service organizations who similarly can develop their DAC to more quickly identify (potential) disruptions and which measures they can take to mitigate their effect on their operations. We therefore propose: in the context of hospitality organizations, DAC is positively related to organizational resilience (H1).

2.3.2 Enablers of Data Analytic Capability

For organizations to develop their DAC, they need to orchestrate people, processes, and technology assets and capabilities (Gupta and George 2016). The development of DAC is qualified as a complex socio-technical process, requiring a multi-disciplinary perspective (Legner et al. 2017; Mikalef and Krogstie 2020). This requires concurrent investments in digital transformation technologies and the corresponding organizational infrastructure (Kokkinou et al. 2025). We therefore propose that in the context of hospitality organizations, IT infrastructure is positively related to DAC (H2) and organizational infrastructure is positively related to DAC (H3).

We posit that DAC is the capability by which hospitality organizations can translate investments in digital tools and processes into increased organizational resilience (He et al. 2023). We therefore propose the corresponding mediation hypotheses: DAC mediates the relationship between IT Infrastructure and organizational

resilience (H2') and DAC mediates the relationship between organizational infrastructure and organizational resilience (H3').

In the context of the hospitality and tourism industries, SMEs are perceived to be more vulnerable to disruptions due to their lack of organizational capabilities and slack resources (Jiang et al. 2019). Similarly, brand-affiliated hotels are perceived as having better access to the resources needed in case of disruptions as compared to independent hotels. Hotel size and affiliation are therefore used as control variables. The conceptual model is shown in Figure 1.

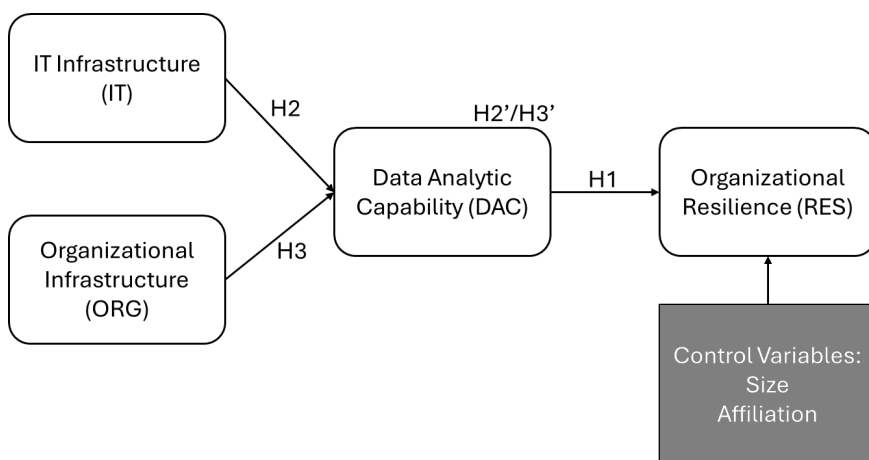


Figure 1: Conceptual Framework

Source: Own work

3 Methodology

To examine the study hypotheses, a cross-sectional survey design was employed as described below. The unit of analysis for the study was the organization, with each organization being represented by a single key informant (Yu et al. 2021; Kokkinou et al. 2023b). Cross-sectional data was used as it is more efficient and inexpensive to collect and reduces the risk of attrition typical of longitudinal studies (Maier et al. 2023). However, for the purpose of the present study, the use of cross-sectional data limits the study of temporal order and the differentiation between cause and effect.

3.1 Survey Development

For this study, a survey was designed and administered using the software Qualtrics. The survey consisted of five sections. First, participants were provided with information about the study and informed consent. In this first section, participants also had to confirm their eligibility for the study. The second section was used to collect demographic information about the hotel company. The third section of the questionnaire measured hospitality resilience, while the fourth section measured the digitalization and DAC of the hotel. The fifth section comprised demographic questions and attention checks.

Previously validated scales obtained from the literature were used to measure the variables of interest of this study. To measure hospitality organization's resilience, the scale developed by He et al. (2023) was adapted. He et al.'s (2023) scale is based on the combination of individual-level and organizational-level resilience. The 15 items were measured on a scale from 1 (completely disagree) to 5 (completely agree). The scale for DAC consisted of four items (Srinivasan & Swink, 2018). The hypothesized antecedent of DAC, namely IT Infrastructure and Organizational Infrastructure, were measured using scales adapted by Proksch et al. (2024) using a scale from 1 to 5. Hotel size and whether the hotel is independent or affiliated were included as control variables.

Given the risk for Common Method Bias (CMB), several survey design best practices were employed. These included keeping the language unambiguous and simple (Lindell and Whitney 2001; Podsakoff et al. 2003), ensuring participants anonymity and providing reassurance that there were no incorrect answers (Li et al. 2020). Furthermore, the scales for the dependent variables (resilience) were shown in a separate page. Halfway through the survey, an open question was used to collect information about the participants' role in the organization. This was also used as an attention check.

3.2 Data Collection Procedures and Sampling

Participants in the study were recruited through the platform Prolific. Participants were eligible for the study if they were currently employed by a hotel, or had been employed in the past six months, in a full-time or part-time position. In addition to

the initial screening question, participants were asked to describe their position in an open question. They were paid £2 to complete the study. The median survey duration was 10:10. Participants who took less than 50% of the median time were removed from the study, as did participants who failed the attention checks. The final sample for this study included 167 participants, of which 44.2% identified as male and 55.2% as female. Regarding position, 6.1% of participants indicated working on a management level, 30.9% on a middle-management level, 27.9% as team leader/ shift supervisor, and 32.1% at operational level. Participants' mean tenure at the company was 5.02 years ($sd = 3.895$). Most participants worked for an affiliated hotel (65.5%). Affiliated hotels included franchised hotels (19.2%), small chains (9.6%) and large chain hotels (35%). 32.1% of participants indicated working for an independent hotel, and 2.4% owned and operated their hotel or a small chain of hotels. Small and large hotels were relatively equally represented in the sample. 60.5% of hotels in the sample had fewer than 100 rooms. The largest two groups were hotels with 11-50 rooms (26.6%) and 51-100 rooms (26.6%). 7.3% of hotels in the sample had fewer than 10 rooms, while 4.0% had over 500 rooms.

3.3 Data Analysis Procedures

Prior to analysis, data were reviewed to remove responses from participants that completed the study too fast and/or failed the attention checks. Subsequently, demographics were computed using the software SPSS. To analyze the data and test the study hypotheses, Partial Least Squares Structural Equation Modeling (PLS-SEM) was applied using the software SMART-PLS (Ringle et al. 2022). This was done in two steps. First, the reliability, convergent validity, and discriminant validity of the focal constructs were examined using the PLS-SEM algorithm. Second, the full model shown in figure 1 was calculated using the bootstrapping function to estimate the significance of the coefficients.

4 Findings

First, and prior to testing the study hypotheses, reliability, convergent validity, and discriminant validity were assessed (Melián-Alzola et al. 2020). Reliability for each construct was determined by examining Cronbach Alpha (CA) and Scale Composite Reliability (SCR). As table 1 shows, for all constructs, CA and SCR exceeded the

recommended threshold of 0.7 (Fornell and Larcker 1981). Convergent validity was assessed by examining the Average Variance Extracted (AVE).

As shown in table 1, AVE exceeded the recommended value of 0.5 for all study constructs. Discriminant validity was assessed by applying two well-known criteria (1) Fornell-Larcker criterion (Fornell and Larcker 1981) and (2) Heterotrait-Monotrait (HTMT) ratio criterion (Henseler et al. 2015). For (1) for each construct, the square root of AVE (see diagonal in table 2) needed to exceed the construct's correlation with each of the other constructs (see bottom left quadrant of table 2), which was the case. For (2), the value of HTMT needed to remain below 0.850 for each combination of constructs (see top right quadrant for table 2), which was also the case. Therefore, discriminant validity was established.

Table 1: Scale Statistics

Construct	Mean	SD	CA	SCR	AVE
IT Infrastructure	3.751	0.869	0.878	0.908	0.623
Organizational Infrastructure	3.662	0.888	0.911	0.929	0.621
Data analytic capability	3.619	0.928	0.879	0.917	0.734
Resilience	4.022	0.730	0.938	0.946	0.558

Note: SD: Standard Deviation / CA: Cronbach Alpha/ SCR: Score Composite Reliability / AVE: Average Variance Extracted

Source: Own work

Table 2: Discriminant Validity (Fornell-Larcker criterion and HTMT)

Construct	AVE	IT	ORG	DAC	RES
IT Infrastructure (IT)	0.623	0.790	0.824	0.783	0.733
Organizational Infrastructure (ORG)	0.621	0.737	0.788	0.816	0.721
Data analytic capability (DAC)	0.734	0.688	0.730	0.857	0.723
Resilience (RES)	0.558	0.666	0.666	0.657	0.744

Note: Bottom left quadrant: correlations / Diagonal: square root of AVE / Top right quadrant: HTMT

Source: Own work

Since the requirements for reliability, convergent validity and discriminant validity were met, the structural model (shown in figure 1) could be examined. This model explained 44.0% of the variance in organizational resilience, and 58.3% of the variance in DAC. The results show that the control variables, size and affiliation,

were not statistically significant (respectively $\beta=-0.037$, $p=.796$ and $\beta=0.294$, $p=.095$). Using bootstrapping (5000 samples), the significance of the coefficients corresponding to each hypothesis were estimated (shown in tables 3 and 4).

Table 3: Direct Effects

Path	Coeff	t-value	p-value	Hyp.	Conclusion
DAC $\rightarrow$ RES	0.685	12.917	.000	H1	Supp.
IT $\rightarrow$ DAC	0.328	4.066	.000	H2	Supp.
ORG $\rightarrow$ DAC	0.488	5.698	.000	H3	Supp.
Size $\rightarrow$ RES	-0.037	0.259	.796	Control	Not
Affiliation $\rightarrow$ RES	0.291	1.668	.095	Control	Not

Source: own work

Table 4: Mediation Hypotheses

Path	Coeff	t-value	p-value	Hyp.	Conclusion
IT $\rightarrow$ DAC $\rightarrow$ RES	0.225	3.665	.000	H2'	Supported
ORG $\rightarrow$ DAC $\rightarrow$ RES	0.334	5.254	.000	H3'	Supported

Source: own work

The results show that DAC has a statistically significant positive effect on organizational resilience ($\beta=0.685$, $p<0.000$), supporting H1. Furthermore, IT infrastructure and organizational infrastructure both have a statistically significant positive effect on DAC (respectively $\beta=0.328$, $p<0.000$ and $\beta=0.448$, $p<0.001$), supporting H2 and H3. Furthermore, an examination of the indirect effects shows that IT infrastructure and organizational infrastructure both have a positive effect on organizational resilience through DAC (respectively $\beta=0.225$, $p<0.000$ and $\beta=0.334$, $p<0.001$), lending support to H2' and H3'.

5 Discussion

Due to their characteristics, hospitality organizations are vulnerable to disruptions that affect demand and supply of their services and therefore need to develop organizational resilience (He et al. 2023; Kokkinou et al. 2023c). The purpose of this study was to examine how hospitality organizations can develop their DAC in such a way as to improve their organizational resilience.

First, DAC was associated with organizational resilience of hospitality organizations. This mirrors research conducted in other contexts such as manufacturing and supply chains (Dubey et al. 2021). However, whereas manufacturing organizations can somewhat mitigate the impact of disruptions on their supply chains through proactive measures such as strategically positioning inventory in their supply chains (Ivanov and Dolgui 2020), hospitality organizations are characterized by their fixed capacity and the simultaneous production and consumption of their services which renders them even more vulnerable to unexpected variations in demand (Melián-Alzola et al. 2020). DAC can therefore help hospitality organizations better anticipate on their demand, through for example improved forecasting (Kokkinou 2013).

Second, and consistent with previous research, we found that to develop organizational resilience, hospitality organizations needed to develop their DAC through the concurrent orchestration of digital technology, processes and people assets and capabilities (Kokkinou et al. 2023a). Hospitality organizations need to invest in the digital technologies that will help them develop an IT infrastructure capable of sensing their environment and translate data into insights (He et al. 2023). They also need to adapt the processes and structure of their organization to incorporate these insights into decision-making to mitigate the effect of disruptions on their performance. Furthermore, organizations also need to invest in the development of employees. Merely investing in digital technologies without providing appropriate support for employees is not sufficient (Motamarri et al. 2017; Kokkinou et al. 2024). This supports the view that the development of DAC is a complex socio-technical process which requires a more comprehensive and multi-disciplinary approach (Legner et al. 2017; Mikalef and Krogstie 2020).

5.1 Implications for Practitioners

The development of resilience is a strategic and long-term endeavor that is frequently postponed in favor of dealing with more immediate problems, especially by SMEs (Jiang et al. 2019). For unaffiliated hotels and other organizations that are not part of a network, the development of a resilience capability might be particularly daunting (Usher et al. 2019). Nevertheless, given the vulnerability of hospitality organizations to changes in demand, hospitality organizations are strongly advised to invest in developing their resilience capability. The ongoing digitalization trend

provides a unique opportunity for hospitality organizations to combine the adoption of digitalization tools with the development of their DAC as way to anticipate on disruptive events and mitigate their impact on operations and performance (Melían-Alzola et al. 2020). By investing in digital tools and their IT infrastructure, hospitality organizations can improve their ability to sense disruptions and take swift action. However, it remains necessary to couple the implementation of new digital tools to the adaptation of the organization's structure and processes. Hospitality organizations can adopt a step-by-step to the development of digital maturity, focusing on developing and adopting a digital strategy, and implementing it through investments in leadership, employee, operational, cultural and governance capabilities (Rossmann 2018).

5.2 Limitations and Future Recommendations

Our study has several limitations that need to be addressed. First, a cross-sectional design, even with ample support from academic literature, limits the ability to draw conclusions regarding causality. Temporal causality between investments in resources, the development of DAC, and the development of resilience, and how variance in investments in resources relate to resilience can be better proven using multi-wave data (Maier et al. 2023). Longitudinal research is thus needed to better understand the causal effects associated with investments in digitalization. Second, the unit of analysis for this study is the organization, however an organization's resilience can also be affected by the ecosystem in which it operates (Hall et al. 2023). For example, previous research has shown that during severe disruptions such as COVID-19, hospitality operators are able to draw on the goodwill of their suppliers to improve their digitalization capabilities (Kokkinou et al. 2023c) and local government for resources (Usher et al. 2019). Future research should therefore incorporate contextual variables such as environmental complexity, competitive intensity, government support, and legal environment as potential moderators.

Acknowledgements

We would like to acknowledge funding received from the Dutch Taskforce for Applied Research (SIA). This publication is part of the project A Collaborative Data Driven Approach to Developing Resilience for Hospitality SMEs (project number MV.KIEM.01.115) funded by the Taskforce for Applied Research (SIA) and the Dutch Research Council (NWO).

References

- Bharadwaj A, El Sawy OA, Pavlou PA, Venkatraman N v (2013) Digital business strategy: toward a next generation of insights. *MIS quarterly* 471–482
- Cotta D, Salvador F (2020) Exploring the antecedents of organizational resilience practices – A transactive memory systems approach. *IJOPM* 40:1531–1559. <https://doi.org/10.1108/IJOPM-12-2019-0827>
- Dubey R, Gunasekaran A, Childe SJ, Fosso Wamba S, Roubaud D, Foropon C (2021) Empirical investigation of data analytics capability and organizational flexibility as complements to supply chain resilience. *International Journal of Production Research* 59:110–128. <https://doi.org/10.1080/00207543.2019.1582820>
- Fornell C, Larcker DF (1981) Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research* 18:382–388. <https://doi.org/10.1177/002224378101800313>
- Gupta M, George JF (2016) Toward the development of a big data analytics capability. *Information & Management* 53:1049–1064. <https://doi.org/10.1016/j.im.2016.07.004>
- Gursoy D, Chi CG (2020) Effects of COVID-19 pandemic on hospitality industry: review of the current situations and a research agenda. *Journal of Hospitality Marketing & Management* 29:527–529. <https://doi.org/10.1080/19368623.2020.1788231>
- Hall CM, Safonov A, Naderi Koupaei S (2023) Resilience in hospitality and tourism: issues, synthesis and agenda. *IJCHM* 35:347–368. <https://doi.org/10.1108/IJCHM-11-2021-1428>
- He Z, Huang H, Choi H, Bilgihan A (2023) Building organizational resilience with digital transformation. *JOSM* 34:147–171. <https://doi.org/10.1108/JOSM-06-2021-0216>
- Helfat CE, Peteraf MA (2009) Understanding dynamic capabilities: progress along a developmental path. *Strategic Organization* 7:91–102. <https://doi.org/10.1177/1476127008100133>
- Helfat CE, Raubitschek RS (2018) Dynamic and integrative capabilities for profiting from innovation in digital platform-based ecosystems. *Research Policy* 47:1391–1399. <https://doi.org/10.1016/j.respol.2018.01.019>
- Henseler J, Ringle CM, Sarstedt M (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J of the Acad Mark Sci* 43:115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Iftikhar A, Ali I, Arslan A, Tarba S (2024) Digital Innovation, Data Analytics, and Supply Chain Resiliency: A Bibliometric-based Systematic Literature Review. *Ann Oper Res* 333:825–848. <https://doi.org/10.1007/s10479-022-04765-6>
- Ivanov D, Dolgui A (2020) Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International Journal of Production Research* 58:2904–2915. <https://doi.org/10.1080/00207543.2020.1750727>
- Jia F, Blome C, Sun H, Yang Y, Zhi B (2020) Towards an integrated conceptual framework of supply chain finance: An information processing perspective. *International Journal of Production Economics* 219:18–30. <https://doi.org/10.1016/j.ijpe.2019.05.013>
- Jiang Y, Ritchie BW, Verreynne M (2019) Building tourism organizational resilience to crises and disasters: A dynamic capabilities view. *International Journal of Tourism Research* 21:882–900
- Kokkinou A (2024) How organizations can harness continuous improvement practices to develop their data analytic capability: a conceptual paper. In: van Kollenburg T, Kokkinou A, McDermott O (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 (2013) Forecasting for the Start-Up Restaurant Owner. *Journal of Foodservice Business Research* 16:101–112. <https://doi.org/10.1080/15378020.2013.761026>
- Kokkinou A, Kollenburg T van, Mandemakers A, Hopstaken E, Elderen J van (2023a) The Data Analytic Capability Wheel: an implementation framework for digitalization. In: Pucihar A, Kljajić Borštnar M, Bons R, Ongena G, Heikkilä M, Vidmar D (eds) *Proceedings of the*

- 36th Bled eConference Digital Economy and Society: balancing act for digital innovation in times of instability: June 25 - 28,. Univerza v Mariboru, Univerzitetna založba, Bled, Slovenia
- Kokkinou A, Mandemakers A, Mitas O (2023b) 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>
- Kokkinou A, Mandemakers A, Mitas O (2024) Cultivating a Data-Driven Culture: Exploring the Role of Data Analytic Capability, Employee Empowerment, and Transformational Leadership. In: Pucihar A (ed) 37th Bled eConference – Resilience Through Digital Innovation: Enabling the Twin Transition: June 9 – 12, 2024, Bled, Slovenia, Conference Proceedings. Bled, Slovenia
- Kokkinou A, Mitas O, Mandemakers A (2025) Leveraging digital transformation to develop a data-driven culture for supply chain resilience: an empirical investigation. *CRR*. <https://doi.org/10.1108/CRR-08-2024-0023>
- Kokkinou A, Mitas O, Zeroual S (2023c) Hospitality Status: It's Complicated - the Impact of Stakeholder Relationships on Restaurant Resilience During COVID-19. *Journal of Quality Assurance in Hospitality & Tourism* 1–23. <https://doi.org/10.1080/1528008X.2023.2206173>
- Law R, Ye H, Chan ICC (2022) A critical review of smart hospitality and tourism research. *IJCHM* 34:623–641. <https://doi.org/10.1108/IJCHM-08-2021-0986>
- Legner C, Eymann T, Hess T, Matt C, Böhm T, Drews P, Mädche A, Urbach N, Ahlemann F (2017) Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. *Bus Inf Syst Eng* 59:301–308. <https://doi.org/10.1007/s12599-017-0484-2>
- Lew AA (2014) Scale, change and resilience in community tourism planning. *Tourism Geographies* 16:14–22. <https://doi.org/10.1080/14616688.2013.864325>
- Li G, Li L, Choi T, Sethi SP (2020) Green supply chain management in Chinese firms: Innovative measures and the moderating role of quick response technology. *J of Ops Management* 66:958–988. <https://doi.org/10.1002/joom.1061>
- Lindell MK, Whitney DJ (2001) Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology* 86:114–121. <https://doi.org/10.1037/0021-9010.86.1.114>
- Luthe T, Wyss R (2014) Assessing and planning resilience in tourism. *Tourism Management* 44:161–163. <https://doi.org/10.1016/j.tourman.2014.03.011>
- Maier C, Thatcher JB, Grover V, Dwivedi YK (2023) Cross-sectional research: A critical perspective, use cases, and recommendations for IS research. *International Journal of Information Management* 70:102625. <https://doi.org/10.1016/j.jinfomgt.2023.102625>
- Melián-Alzola L, Fernández-Monroy M, Hidalgo-Peñate M (2020) Hotels in contexts of uncertainty: Measuring organisational resilience. *Tourism Management Perspectives* 36:100747. <https://doi.org/10.1016/j.tmp.2020.100747>
- 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* 29:260–287. <https://doi.org/10.1080/0960085X.2020.1740618>
- Mikalef P, Van de Wetering R, Krogstie J (2019) From Big Data Analytics to Dynamic Capabilities: The Effect of Organizational Inertia. *AIS Electronic Library*, pp 1–14
- Motamarri S, Akter S, Yanamandram V, Wamba SF (2017) Why is Empowerment Important in Big Data Analytics? *Procedia Computer Science* 121:1062–1071. <https://doi.org/10.1016/j.procs.2017.11.136>
- Neuhofer B, Buhalis D, Ladkin A (2015) Smart technologies for personalized experiences: a case study in the hospitality domain. *Electron Markets* 25:243–254. <https://doi.org/10.1007/s12525-015-0182-1>
- Pal R, Torstensson H, Mattila H (2014) Antecedents of organizational resilience in economic crises—an empirical study of Swedish textile and clothing SMEs. *International Journal of Production Economics* 147:410–428. <https://doi.org/10.1016/j.ijpe.2013.02.031>

- Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP (2003) Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88:879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Proksch D, Rosin AF, Stubner S, Pinkwart A (2024) The influence of a digital strategy on the digitalization of new ventures: The mediating effect of digital capabilities and a digital culture. *Journal of Small Business Management* 62:1–29. <https://doi.org/10.1080/00472778.2021.1883036>
- Ringle C, Wende S, Becker J-M (2022) SmartPLS 4
- Rossmann A (2018) Digital Maturity : Conceptualization and Measurement Model. In: *Proceedings of the International Conference on Information Systems : Bridging the Internet of People, Data, and Things* (39th ICIS 2018). Association for Information Systems, Atlanta, GA
- Sutcliffe KM, Vogus T (2003) Organizing for resilience. In: Cameron KS, Dutton JE, Quinn RE (eds) *Positive organizational scholarship: Foundations of a new discipline*. Berrett-Koehler Publishers
- Teece DJ (2007) Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strat Mgmt J* 28:1319–1350. <https://doi.org/10.1002/smj.640>
- Teece DJ, Pisano G, Shuen A (1997) Dynamic capabilities and strategic management. *Strat Mgmt J* 18:509–533. [https://doi.org/10.1002/\(SICI\)1097-0266\(199708\)18:7<509::AID-SMJ882>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z)
- Usher LE, Yusuf J-E (Wie), Covi M (2019) Assessing tourism business resilience in Virginia Beach. *IJTC* 6:397–414. <https://doi.org/10.1108/IJTC-02-2019-0019>
- Vakilzadeh K, Haase A (2021) The building blocks of organizational resilience: a review of the empirical literature. *CRR* 3:1–21. <https://doi.org/10.1108/CRR-04-2020-0002>
- Vial G (2019) Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems* 28:118–144. <https://doi.org/10.1016/j.jsis.2019.01.003>
- Wamba SF, Gunasekaran A, Akter S, Ren SJ, Dubey R, Childe SJ (2017) Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research* 70:356–365. <https://doi.org/10.1016/j.jbusres.2016.08.009>
- Yu W, Wong CY, Chavez R, Jacobs MA (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>

FRAMING KNOWLEDGE: ENHANCING COMMUNICATION STRATEGIES IN THE DUTCH SECOND-HAND CLOTHING INDUSTRY

MATS BELGERS,¹ MARJA EXALTO-SIJBRANDS,¹

MICHÈLLE BAL,² PASCAL RAVESTEIJN¹

¹ Utrecht University of Applied Sciences, Process Innovation and Information Systems,
Utrecht, Netherland

mats.belgers@hu.nl, marja.exalto-sijbrands@hu.nl, pascal.ravesteijn@hu.nl

² University of Utrecht, Faculty of Social Sciences, Utrecht, Netherland
m.bal@uu.nl

The increasing growth of clothing consumption substantially impacts the environment. Although consumers may be aware of this impact, the 'green gap' between their environmental intentions and their actual purchase behaviour is notable. This study explores the effects of different types of environmental knowledge (land- and water-use and CO₂ emissions) and different types of framing (positive and negative) on Dutch second-hand clothing consumption behaviour. Through an experiment with six groups, anticipated shame and pride were measured in hypothetical situations. Results indicate that positive framing has more effect on second-hand clothing consumption behaviour than negative framing, while effects between different types of knowledge were minimal. Open questions showed that, positive, memorable, communication strategies, with sufficient background information, might help in bridging the green gap in clothing consumption. These insights may help policymakers and stakeholders to enhance the effectiveness of their communication strategies, thereby increasing second-hand clothing consumption.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.4](https://doi.org/10.18690/um.fov.4.2025.4)

ISBN

978-961-286-998-4

Note

This proceeding is based on master thesis Belgers, M. (2024). *Moving towards More Second-hand Clothing Consumption: Measuring Anticipated Shame and Pride on Different Types of Concrete Environmental Knowledge*.
Utrecht University.

Available at:

[https://studenttheses.](https://studenttheses.uu.nl/handle/20.500.12932/47121)

[uu.nl/handle/20.500.
12932/47121](https://studenttheses.uu.nl/handle/20.500.12932/47121)

Keywords:

second-hand clothing
consumption,
behaviour change,
Value-Belief-Norm theory,
environmental knowledge,
framing



University of Maribor Press

1 Introduction

The amount of clothing consumption has nearly doubled over the past two decades, mainly due to the uprising of fast-fashion companies such as H&M and Zara (Ellen Mac Arthur Foundation, 2017; Koszewska, 2018). The consumption culture of buying much at low prices (that fast-fashion supports) impacts the environment with energy consumption, water and chemical consumption, the amount of solid waste, and the amount of CO₂ emissions (Koszewska, 2018). Most clothing items worldwide are being produced using non-renewable resources, being worn for a short time, and end up in landfills or being incinerated (Ellen Mac Arthur Foundation, 2017). Within the Netherlands, 5% of the total environmental damage by individuals is caused by the consumption, use, and disposal of clothing (Milieu Centraal, 2020). The average Dutch consumer buys around 46 clothing items per year and keeps around 173 items in their wardrobe. Of these 173 items, 50 have not been worn the past year, and only 16 of the 40 disposed clothing items are being reused or recycled (Maldini et al., 2017). For the industry to be more sustainable, it should adopt the idea of a circular economy, including the aspects of reducing and reusing (Koszewska, 2018; Rebel, 2023). Within the clothing industry, this implies more consumption of second-hand clothing and less new produced clothing.

Current literature shows that some consumer groups have biospheric values, which means that they care about the climate and, to some extent, are aware that buying less clothing is better (Khandual and Pradhan, 2019; Motivaction, 2023). In the Netherlands, 40% of consumers see the importance of buying fewer clothes, and 33% feel responsible for preventing pollution of the environment (Motivaction, 2023). However, only 7% state that they deliberately bought fewer clothes last year because of the impact on humans and the environment (Motivaction, 2023) and only 4% of the items in consumers' wardrobes are second-hand (Maldini et al., 2017). These statistics show a discrepancy between the consumer's intention and behaviour (Gleim and Lawson, 2014; Juvan and Dolnicar, 2014; National Geographic & GlobeScan, 2012; Young et al., 2010), which is commonly known as the 'green gap' within the sustainability context (ElHaffar et al., 2020). The green gap exists e.g. due to the lack of knowledge of consumers on the negative environmental consequences of clothing consumption (Bray et al., 2010; Connel, 2010; ElHaffar et al., 2020; Ronda, 2024). Although consumers have a certain level of awareness, the exact negative impacts of new clothing consumption misses.

Since the influence of concrete consumer knowledge on second-hand clothing consumption has not been studied yet, this study will explore the relationship between different types of knowledge, both negatively framed and positively framed, and sustainable clothing consumption behaviour. The idea is that concrete knowledge will raise more awareness and therefore lead to more second-hand clothing consumption. This will be measured by assessing anticipated shame and anticipated pride in hypothetical situations, as these anticipated emotions mediate the relationship between personal norms and behaviour.

2 Theoretical framework

2.1 The NAM and VBN theory

The norm activation model (NAM) posits that pro-environmental behaviour depends on the personal norms (PN) of an individual (Schwartz 1977; Ünal et al. 2018). To activate PN, an individual will need awareness of the adverse consequences (AC), followed by the idea that changing its behaviour has an actual positive impact on the environment (ascription of responsibility to self, or AR). Stern's (1999) extension of the NAM, the Value-Belief-Norm (VBN) theory, states personal values affect someone's ecological worldview, which in turn affects someone's awareness. Values are described as "desirable goals, varying in importance, that serve as guiding principles in people's lives" (Schwartz 1992, 21). For understanding sustainable consumption, four types of values are relevant (Stern 1999): Hedonic values (pursuit of pleasure and avoidance of pain), Egoistic values (personal interests), Altruistic values (welfare and well-being of others), and Biospheric values (the well-being of the environment). Unsurprisingly, individuals who prioritize altruistic and biospheric values are more likely to engage in sustainable consumption. However, as values are relatively stable over one's lifetime (Schwartz, 1992), it is hard to change someone's values.

2.2 Environmental knowledge and Anticipated emotions

Research by Bolderdijk et al. (2013) reveals that, providing environmental knowledge to people of whom it is proven that they strongly endorse biospheric values, will lead to environmental action. When these people obtain information about the negative consequence of a certain product, they are confronted with their behaviour

not being in line with their prioritized values, which will cause them to change their behaviour (Steg, 2023). Steg therefore states that it is important to ensure people's awareness of their environmental impact of their actions, resulting in acting upon their biospheric values.

This confrontation of not aligning your behaviour with your values might lead to feelings of guilt and shame (Amatulli et al., 2019; Onwezen et al., 2013; Tian and Liu, 2022). Subsequently, when an individual's behaviour does align with their values, it might lead to feelings of pride. Onwezen et al. (2013) state that anticipated emotions of guilt and pride influence future decision-making, as people strive to accomplish positive emotions and avoid negative emotions. Using the NAM, Onwezen et al. conclude that "anticipated pride and guilt mediate the effects of personal norms on behaviour." Applied on consumers with strong biospheric values, the provision of both the negative consequences of newly produced clothing consumption and the positive consequences of second-hand clothing consumption could move them towards more second-hand clothing consumption.

2.3 Framing

The way messages are framed shapes consumer behaviour (Amatulli et al., 2019). Amatulli et al. describe message framing as "highlighting specific aspects and making them more salient in communication" (p. 1113). Various research has demonstrated that negative framing is more effective at promoting sustainable purchases than positive framing (Amatulli et al., 2019; Olsen et al. 2014). According to White et al. (2011), negative framing is processed in more detail by consumers and it makes them feel more empowered. Contradicting with Onwezen (et al. 2013), Amatulli et al. (2019) highlight the emotion of shame instead of guilt. They argue that shame is more effective in changing behaviour than guilt, as guilt is a negative judgment about an action, and shame is a negative judgment about the self.

On the other hand, arguments can be made for the application of positive framing. Some studies (Bouman & Steg, 2022; Schneider et al., 2021) see the potential for a 'positive spiral' through the emergence of positive emotions and productive engagement. Emotions such as optimism, hope, gratitude, anticipated pride, and a 'warm-glow', can contribute to this positive spiral (Schneider et al., 2021). Transmitting messages that evoke positive emotions about people's abilities to

contribute to an environmentally healthy planet could therefore lead to more sustainable purchase practices.

2.4 This study

The abovementioned concepts are assembled into the conceptual model in Figure 1. Which type of environmental knowledge has the most effect on consumer's behaviour has not been studied yet. It is also unknown which type of framing (positive or negative) works best in the context of second-hand clothing consumption. Therefore this study explores the effects of different types of knowledge (land use, water use, and CO2 emissions) and different types of framing (positive and negative) on consumer behaviour with the research question: *Which types of knowledge about the environmental impact of clothing consumption, negatively framed or positively framed, encourages second-hand clothing consumption the most?*

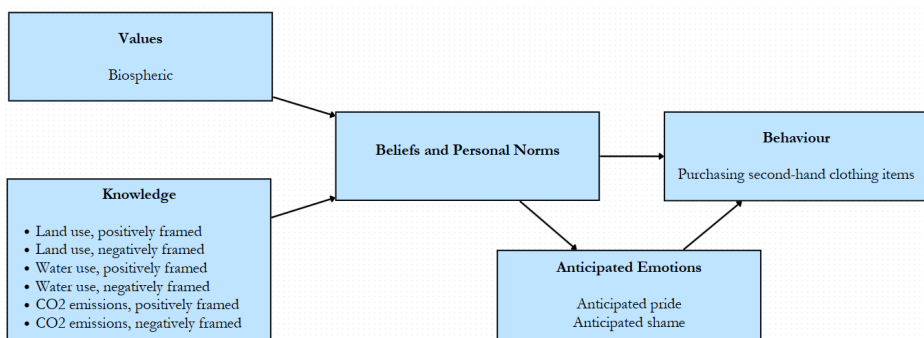


Figure 1: Conceptual model of integration of knowledge, framing and anticipated emotions in the Value-Belief-Norm theory

Source: Own

3 Methods

Through a quantitative approach including open-ended questions, both high generalizability and deeper insights were ensured. The research was conducted among students and employees of HU University of Applied Sciences and Utrecht University (UU) via an online survey reaching 234 respondents. The sample taken into analyses was 169 participants, (40 male, 123 female, 5 non-binary/third gender and 1 respondent preferred not to say). Their ages ranged between 17 and 69 ($M =$

23.28, SD = 6.01). The survey consisted partly of an experiment with six conditions to compare the differences between the effects of providing different types of knowledge (about land, water and CO₂), and different types of framing (negative and positive) on anticipated emotions and purchase intentions. For example, knowledge about land could be negatively framed: “For each EU citizen, it currently takes 400 square meters of land yearly to provide clothes and shoes” or positively framed: “For each EU citizen who only consumes second-hand clothing, it saves 200 square meters of land yearly.” Between 25 and 31 participants took part in each of these six conditions. Additionally, respondents were asked questions which were unrelated to the experiment.

3.1 Survey procedure and Materials

Participants were recruited using online and offline methods. They were randomly put in one of the six conditions. The survey started with demographic questions, followed by an adjusted version of Schultz’ (2001) Environmental Concerns scale (7-point Likert), which evaluates the degree of concern individuals have about environmental issues, divided into three primary values: Egoistic values, Biospheric values, and Altruistic values. Next, the participants took part in the experiment, in which they were brought in a hypothetical situation, to imagine that they wanted to purchase a clothing item: in one situation a newly produced clothing item and in the other situation a second-hand clothing item. Answering with only the knowledge from one of the six conditions, they needed to anticipate in how far they would experience a feeling of shame and pride when purchasing the newly produced clothing item and the second-hand clothing item (7-point Likert). Consequently, respondents were asked whether they would purchase the newly produced clothing item and whether they would purchase the second-hand clothing item (5-point Likert), and were asked to explain their choice. Finally, to explore their opinions, respondents were given all three types of knowledge, both negatively framed and positively framed, followed by the question of which knowledge would persuade them the most to purchase second-hand clothing items over new clothing items and why this knowledge would persuade them the most.

3.3 Data analysis

Data were first examined for accuracy, missing values, and outliers. Descriptive statistics were computed for all variables, including means and standard deviations for continuous variables (e.g. age) and frequencies for categorical variables (e.g. gender). Due to group comparison, one-way ANOVA's (analysis of variance) were performed to examine causal effects between the six conditions and the anticipated emotions and purchase intentions (Tabachnick and Fidell, 2007). The independent variables were the types of knowledge, either negatively framed or positively framed. The dependent variables were the anticipated emotion and the purchase intentions. Post-hoc tests with Bonferroni correction identified specific group differences.

Concerning the open-ended questions, all responses were gathered into a single document for analysis and reviewed for completeness. Codes were created across the dataset to capture key concepts and ideas. Behind all codes numbers were put to assess the frequency of the codes.

4 Results

This section involves the results of this study's population's values on Schultz' Environmental Concerns scale, this study's population's anticipated emotions and purchase intentions given their different conditions, and this study's population's most persuasive types of knowledge and the reasons behind this.

4.1 Values on Schultz' Environmental Concerns scale

Results in Table 1 show that this study's population scores the highest (7-point Likert) on biospheric values, followed by altruistic values. This study's population scores the lowest on egoistic values.

Table 1: values on Schultz' Environmental Concerns scale

	Biospheric values	Altruistic values	Egoistic values
Study's population	6.06	5.98	5.43

4.2 Conditions and anticipated emotions/purchase intentions

Table 2 shows which comparisons between the conditions show

significant results on anticipated emotions and purchase intentions. Results show that positive framing (PF) makes participants in our study less proud of their newly produced clothing item than negative framing (NF). Additionally, some combinations show that positive framing, CO2 knowledge, and water knowledge, are more persuasive towards purchasing second-hand clothing items for participants in our study. Lastly, a negative framing of land makes that participants in our study rather purchase a second-hand clothing item than a positive framing of land.

Table 2: significant results for Anticipated emotions / Purchase intentions with conditions

Dependent variable		Significant result for	F-value	Df	P-value	Effect size	Mean (SD)
Newly produced clothing item	Anticipated pride	PF vs. NF	6.153	2, 163	0.014	0.036	Positive: 3.3 (0.14), Negative: 3.8 (0.15)
Second-hand clothing item	Anticipated pride	Water PF vs. Water NF Interaction	6.151	1, 162	0.014	0.037	Water Positive: 6.00 (0.24), Water Negative: 5.15 (0.24)
	Purchase intentions	CO2 PF vs. Land PF Interaction	3.595	2, 163	0.008	0.042	CO2 Positive: 4.16 (0.16), Land Positive: 3.55 (0.16)
		Water PF vs. Land PF Interaction	3.595	2, 163	0.027	0.042	Water Positive: 4.07 (0.17), Land Positive: 3.55 (0.16)
		Land NF vs. Land PF Interaction	3.595	2, 163	0.029	0.042	Land Negative: 4.08 (0.18), Land Positive: 3.55 (0.16)

4.4 The most persuasive type of knowledge

Table 3 shows which type of knowledge, either positively framed or negatively framed, respondents thought would be most persuasive to them. The open-ended questions revealed that the most prevalent reason to choose CO₂ is that respondents find this the most urgent climate issue, partly because they are most familiar with the CO₂ issue, as one respondent answered: “Carbon emissions is used more frequently when talking about climate change so I feel like I can imagine that effect better.” They often hear it on the news or read about it. Some respondents are not familiar at all with the necessity to save water and land and believe CO₂ emissions are the only real climate issue. Additionally, because the knowledge on CO₂ was put in percentages instead of in numbers of square metres or litres of water, this knowledge was more convincing to them. They were able to put this knowledge in perspective and many respondents stated that it was hard to imagine how much land and water it was. One respondent answered for example: “I think percentage of total carbon emissions is something that people can grasp easier than a number that they cannot compare to anything. 400 sqm great. But how much space is that really? I don’t know.”

Table 3: Choices for type of knowledge

	Land	Water	CO ₂	Cumulative
Negative	8	15	43	66
Positive	6	29	54	89
Cumulative	14	44	97	

Reasons for respondents to opt for water are that they are familiar with water issues, such as clean water scarcity, and water scarcity for agriculture, and they found water could be used for better things than clothing production. One respondent argued: “Carbon footprint is so commonly used I don’t think it triggers people anymore. Water is now being reported as becoming a scarcity.” Another prevalent reason to choose water was that it had the biggest number (9000 litres) of all three types of knowledge and respondents were therefore most impressed by this number.

Land was not chosen much because respondents found it hard to imagine how much land a number in square metres actually represents and some questioned the importance of saving land for the environment, as one respondent expressed:

“Square meters of land does not seem so relevant, as in the countries of production (e.g. Central Asia) 400 square meters is comparatively little.” Additional information was necessary according to a respondent.

Concerning the framing, most respondents chose positive framing because it shows the positive effects of our actions and behaviour and it provides solutions and alternatives. It consequently gives a good feeling and some respondents stated that this made them to want to come in action, as one respondent expressed: “It’s an encouraging message which drives me to purchase second hand clothes.” Arguments to opt for negative framing were that this made a bigger impact and therefore gives more urgency to come in action: “It’s about doing something to prevent the bad consequences of the clothing industry. And 10% sounds more than decreasing it to 5%”, one respondent argued. Furthermore, some argued that the positive framing falls short, because it only says how much you are saving, while you have no idea how much it costs the environment.

5 Discussion and conclusion

5.1 Types of values and knowledge

To assess whether the target group scores high on biospheric values, the mean scores of the three types of values were compared with the mean scores of the study of Dornhoff et al. (2019), who also used Schultz’ Environmental Concerns scale. The current target group has mean scores of 6.06 (biospheric), 5.98 (altruistic), and 5.43 (egoistic), while students of Dornhoff et al.’s study have mean scores of 4.5 (biospheric), 4.29 (altruistic), and 4.44 (egoistic) in Ecuador; 4.01 (biospheric), 4.12 (altruistic), and 3.87 (egoistic) in Germany. Comparing these results demonstrates that the current target group has relatively high biospheric values as well as altruistic values, and relatively low egoistic values. This could indicate that, for a lot of the respondents, the provision of environmental knowledge about the negative impact did have an effect on their anticipated emotions and purchase intentions.

Although the experiment does not provide evident findings throughout all the questions, some significant results indicate that knowledge on the impact of water use and CO₂ emissions is more effective than knowledge on land use, which might indicate that consumers are less concerned about saving land than about saving

water or reducing CO₂ emissions. It may also indicate that understanding the impact of land use is beyond the current imagination of world's climate impact on land use.

Moreover, a clear preference for knowledge on CO₂ emissions was found, with 63% of the respondents opting for CO₂, as many respondents are most familiar with the issue of CO₂ emissions, and less familiar with the importance of saving land and water for the environment. This means that providing numbers of issues consumers know nothing about, such as land loss, will miss its impact on consumers and therefore cannot lead to emotions as shame and pride. Consumers should first be better informed about what those number mean. Therefore, future research should focus on whether providing background knowledge makes a difference in the decision-making process of consumers and if providing such knowledge is feasible to convey.

Another interesting finding is that respondents state that 'percentages' as unit make a bigger impact than absolute numbers, as percentages can be better placed in perspective. Comparing these three different knowledge types of impacts was thus never a fair comparison. However, the followed method now provides this insight in the chosen knowledge types. Translating knowledge about land and water into percentages would therefore be better equipped for research, which corresponds with the study of Brase (2002) that shows small-scale (simple frequency) and percentage (relative frequency) formats should be used to create fast and easy understanding of numbers at respondents. Brase (2000) argues that, in situations where percentages are very small and thus not able to make an impact, absolute numbers can be more effective. Alternatively, respondents in the current study argue that knowledge is also easier to imagine when it is translated into comprehensible images, such as amount of shower sessions for water or soccer fields for land use. Future research should therefore focus on making the knowledge better conceivable, either through percentages, simple frequencies, or visual representations.

5.2 Framing

We can conclude that, for respondents in the current study, positive framing works slightly better to promote second-hand clothing consumption. However, these findings contradict to most literature, which state that negative framing is more effective, as this is processed in more detail by consumers and it makes them feel

more empowered (Amatulli et al. 2017; White et al. 2011). The respondents who preferred positive framing reaffirm the arguments of Schneider et al. (2021) that positive framing can make consumers experience positive emotions such as hope, optimism, and pride. However, one named disadvantage of the positive framing in this study is that respondents were not able to put this in perspective, as they did not know the actual costs. It is therefore recommended to add some sort of comparison or percentage to the positive framing.

5.3 Conclusion

The outcomes of this study contain implications for the Dutch society on two different levels: for (governmental) policymakers and practitioners within the clothing industry. First, implications for (governmental) policymakers correspond with earlier recommendations of Motivaction (2023) which emphasize increasing moral consciousness and the sense of responsibility, and offer perspective for action. This could be in the form of awareness campaigns, education classes, and collaborations with influencers. Support seems to be present in society, as Motivaction (2023) states that 54% of Dutch citizens are positive towards awareness campaigns about the environmental consequences of clothing consumption. Secondly, practitioners working for second-hand clothing stores could improve second-hand clothing consumption by communicating the environmental impact of clothing consumption via messages in the store, advertisements, and social media.

In the communication to consumers, consumers should be informed about the negative consequences of clothing consumption, but emphasis should be placed on creating a positive spiral within society, in which consumers can be proud of their second-hand clothing purchases and in which they have reason to be hopeful for the future. According to Bouman and Steg (2022), the actions of governments and businesses are crucial in creating a positive spiral of climate action. Furthermore, the knowledge should be conveyed in a conceivable and memorable way in order to have a lasting impact. This could be in the form of percentages or visualizations. Finally, the knowledge should come with enough background information about the topic, so that consumers can comprehend the whole picture and are able to make a well-considered purchase decision.

5.4 Limitations

In conducting research, it is crucial to reflect on how the researcher positionality may have influenced the study. In recruiting participants, most participants were either the social circle or were asked in person, running the chance that the respondents have similar personalities, values, and interests in sustainability as the researcher. Furthermore, interest in sustainability might have shaped the interpretations, the data and drawn conclusions. All these aspects might have caused a non-representative sample of students and employees of the HU and UU, which in itself may be argued as lack of a representative population. Additionally, females were most present in the target group. Therefore future research should have respondents from all layers of society and a more equal distribution of genders.

References

- Amatulli, C., De Angelis, M., Peluso, A. M., Soscia, I., & Guido, G. (2019). The effect of negative message framing on green consumption: An investigation of the role of shame. *Journal of Business Ethics*, 157, 1111-1132.
- Bolderdijk, J. W., Gorsira, M., Keizer, K., & Steg, L. (2013). Values determine the (in) effectiveness of informational interventions in promoting pro-environmental behavior. *PloS one*, 8(12), e83911.
- Bouman, T., & Steg, L. (2022). A spiral of (in) action: Empowering people to translate their values in climate action. *One Earth*, 5(9), 975-978.
- Brase, G. L. (2002). Which statistical formats facilitate what decisions? The perception and influence of different statistical information formats. *Journal of Behavioral Decision Making*, 15(5), 381-401.
- Bray, J., Johns, N., & Kilburn, D. (2011). An exploratory study into the factors impeding ethical consumption. *Journal of business ethics*, 98, 597-608.
- Connell, K. Y. H. (2010). Internal and external barriers to eco-conscious apparel acquisition. *International Journal of Consumer Studies*, 34(3), 279-286.
- Dornhoff, M., Sothmann, J. N., Fiebelkorn, F., & Menzel, S. (2019). Nature relatedness and environmental concern of young people in Ecuador and Germany. *Frontiers in Psychology*, 10, 422312.
- ElHaffar, G., Durif, F., & Dubé, L. (2020). Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions. *Journal of cleaner production*, 275, 122556.
- Ellen MacArthur Foundation. (2017). *A new textiles economy: Redesigning fashion's future*. Ellen MacArthur Foundation. https://www.ellenmacarthurfoundation.org/assets/downloads/publications/A-New-Textiles-Economy_Full-Report.pdf
- Gleim, M., & Lawson, S. J. (2014). Spanning the gap: An examination of the factors leading to the green gap. *Journal of Consumer Marketing*, 31(6/7), 503-514.
- Juvan, E., & Dolnicar, S. (2014). The attitude-behaviour gap in sustainable tourism. *Annals of tourism research*, 48, 76-95.
- Khandual, A., & Pradhan, S. (2019). Fashion brands and consumers approach towards sustainable fashion. In S. Muthu (Ed.), *Fast fashion, fashion brands and sustainable consumption* (pp. 37-54). Springer.

- Koszewska, M. (2018). Circular economy—Challenges for the textile and clothing industry. *Autex Research Journal*, 18(4), 337-347.
- Maldini, I., Duncker, L., Bregman, L., Piltz, G., Duscha, L., Cunningham, G., & Balgooi, F. (2017). Measuring the Dutch clothing mountain.
- Milieu Centraal. (2020). *Consumer insights: Tweedehands kleding kopen – Mei 2020*.
- Motivaction. (2023, October 9). *Input circulair textielbeleid. Wat stimuleert tot minder kleding kopen?* Ministerie van Infrastructuur en Waterstaat. https://www.motivaction.nl/content/uploads/2012/07/Greendex_2012_Full_Report_NationalGeographic_GlobeScan.pdf
- National Geographic, & GlobeScan. (2012). *Greendex 2012: Consumer choice and the environment – A worldwide tracking survey*. National Geographic Society. https://globescan.com/wp-content/uploads/2012/07/Greendex_2012_Full_Report_NationalGeographic_GlobeScan.pdf
- Olsen, M. C., Slotegraaf, R. J., & Chandukala, S. R. (2014). Green claims and message frames: How green new products change brand attitude. *Journal of Marketing*, 78(5), 119-137.
- Onwezen, M. C., Antonides, G., & Bartels, J. (2013). The norm activation model: An exploration of the functions of anticipated pride and guilt in pro-environmental behaviour. *Journal of Economic Psychology*, 39, 141-153.
- Rebel. (2023). *Minder textiel: Literatuuronderzoek naar de preventie van consumptie en productie van textiel in Nederland*.
- Ronda, L. (2024). Overcoming barriers for sustainable fashion: bridging attitude-behaviour gap in retail. *International Journal of Retail & Distribution Management*, 52(1), 44-61.
- Schneider, C. R., Zaval, L., & Markowitz, E. M. (2021). Positive emotions and climate change. *Current Opinion in Behavioral Sciences*, 42, 114-120.
- Schultz, P. W. (2001). The structure of environmental concern: Concern for self, other people, and the biosphere. *Journal of environmental psychology*, 21(4), 327-339.
- Schwartz, S. H. (1977). Normative influences on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221-279). Academic Press.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). Academic Press.
- Steg, L. (2023). Psychology of climate change. *Annual Review of Psychology*, 74, 391-421.
- Stern, P. C. (1999). Information, incentives, and proenvironmental consumer behavior. *Journal of Consumer Policy*, 22(4), 461-478.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA* (Vol. 724). Belmont, CA: Thomson/Brooks/Cole.
- Tian, H., & Liu, X. (2022). Pro-environmental behavior research: Theoretical progress and future directions. *International Journal of Environmental Research and Public Health*, 19(11), 6721.
- Ünal, A. B., Steg, L., & Gorsira, M. (2018). Values versus environmental knowledge as triggers of a process of activation of personal norms for eco-driving. *Environment and behavior*, 50(10), 1092-1118.
- White, K., MacDonnell, R., & Dahl, D. W. (2011). It's the mind-set that matters: The role of construal level and message framing in influencing consumer efficacy and conservation behaviors. *Journal of Marketing Research*, 48(3), 472-485.
- Young, W., Hwang, K., McDonald, S., & Oates, C. J. (2010). Sustainable consumption: green consumer behaviour when purchasing products. *Sustainable development*, 18(1), 20-31.

WHERE IS EMOTION? COMPARING EMOTION MAPPING TECHNOLOGIES

ONDREJ MITAS,¹ TAMARA SURLA,² BEN WIELENGA,³
JELENA FARKIĆ,¹ PETER WARD,⁴ ALINDA KOKKINO,⁵
SANDY THOMAS^{1,6}

¹ Breda University of Applied Sciences, Academy for Tourism, Breda, the Netherlands
mitas.o@buas.nl, farkic.j@buas.nl, 236538@buas.nl

² University of Novi Sad, Department of Geography, Tourism and Hotel Management,
Faculty of Sciences, Novi Sad, Serbia
tamara.bozovic@dgt.uns.ac.rs

³ NHL Stenden University of Applied Sciences, Academy of Leisure, Tourism,
Hospitality, European Tourism Futures Institute (ETF), Leeuwarden, the Netherlands
ben.wielenga@nhlstenden.com

⁴ Brigham Young University, Department of Experience Design and Management, Provo,
Utah, United States of America
peter_ward@byu.edu

⁵ Breda University of Applied Sciences, Academy for Built Environment and Logistics,
Breda, the Netherlands
kokkinou.a@buas.nl

⁶ Wageningen University, Wageningen, the Netherlands
236538@buas.nl

Understanding how experiences unfold requires measuring participants' emotions, especially as they move from location to location. Measuring and mapping emotions over space is technically challenging, however. While a number of technologies to record and spatially resolve emotion data exist, they have not been systematically compared. We present emotion data collected at a natural and military heritage site in the Netherlands using three different methods, namely retrospective self report, experience reconstruction, and physiology. These data are applied to three corresponding mapping methods. The resulting maps lead to divergent findings, demonstrating that spatial mapping of emotion data accentuates differences between distinct dimensions of emotions.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.5](https://doi.org/10.18690/um.fov.4.2025.5)

ISBN

978-961-286-998-4

Keywords:

forts,
nature,
heritage,
emotions,
location tracking,
spatial analysis



University of Maribor Press

1 Introduction

Experiences are a key construct in understanding how people use and value space. *Spaces* become *places* when imbued with social meanings (Tuan, 1977), and such meanings arise from individuals' experiences. Experiences are defined differently by fields such as philosophy and neuroscience, but a common thread in existing definitions is that an experience comprises the conscious content of a person's mind. In this paper we follow the Bastiaansen et al. (2019) theory of experience as a process, in which sensory inputs are filtered through mental models, creating a stream of consciousness in a person's mind. According to Bastiaansen et al. (2019), emotions are crucial in the process of experiencing. Emotions translate the inputs of experience into both memory and behavior. An experiential episode only changes a person's behavior, and is only remembered, if it makes a person emotional. Thus, measuring emotions is crucial to designing and managing experiences.

In recent years, scholars have asserted that experiences should be managed based on measuring participants' emotions, and that emotion data should be spatially visualized in 'emotion maps' to inform management decisions (Mitas, Cuenen, et al., 2020; Mitas, Mitasova, et al., 2020). This paper explores three emotion mapping methods: 1) experience reconstruction maps; 2) emotion physiology maps, and; 3) emotion effectiveness maps. Importantly, no single study has applied all three of these methods to the same context. Thus, little information currently exists to help researchers choose an optimal emotion mapping method to address spatial management demands. The objective of this paper is to **compare and contrast the affordances and caveats of each of these methods of emotion mapping**. Extending previous research, we apply all three emotion mapping methods to the same location for the first time, a Napoleonic fort in the Netherlands which offers visitors both natural and built heritage settings. We acknowledge that this is a primarily methodological objective, with potential theoretical and practical implications playing a secondary role.

2 Background

Measuring experiences with useful spatial resolution is technically challenging. Emotions arise and fade in the mind quickly and unpredictably. Memory and recollection of emotions is biased (Bastiaansen et al., 2019). Therefore, technologies

to measure emotion all have limitations. These limitations must also be kept in mind when applying spatial resolution to emotion data.

Bastiaansen et al. (2019) propose the use of emotion measurement technologies from psychology and neuroscience. Their approach enables real-time tracking of emotional reactions through EEG, heart rate, facial expressions, and skin conductance. These may be complemented using more traditional methods of retrospective self report.

2.1 Emotion measuring technologies

2.1.1 Retrospective self-report

Many studies rely on self-reports to measure experiences, whether through post-experience questionnaires and interviews or experience sampling. It is difficult to fully capture emotional ebb and flow with self-report methods. First, they measure memory of the experience rather than the experience itself, which can lead to distortions due to aggregation processes and temporal bias (Tonetto & Desmet, 2016). Second, time resolution by experience sampling, although seen by some as 'real time' measurement, in fact disrupts the natural flow of an experience, particularly in short or immersive settings. Third, self-reports are susceptible to social desirability bias and coping strategies (Larsen & Fredrickson, 1999), while fourth, they fail to capture unconscious emotional processes that play a crucial role in forming memorable experiences (Winkielman & Berridge, 2004). Therefore, existing methods should be complemented with modern research tools from psychology and neuroscience, enabling more precise and comprehensive measurement of emotions throughout experiential episodes (Bastiaansen et al., 2019).

2.1.2 Experience reconstruction self-report

Kahneman et al. (2004) developed the 'day reconstruction method,' a narrative interview at the end of the day that segmented the narrative into several experience episodes, and elicited emotion ratings for each episode. Because the process of narrative interviewing and segmenting is laborious, Strijbosch et al. (2021) developed the experience reconstruction method one step further, identifying obvious episodes

(scenes) in a structured experience (a musical) *a priori* and using photos and descriptions to represent these in a questionnaire, eliciting self-report emotion ratings of each.

Experience reconstruction has several important advantages. If limited in scope, it is convenient and fairly low-burden for participants, fitting easily into a post-experience exit questionnaire. It can accommodate a variety of self-report emotion items and formats. If the episodes of the experience represent locations rather than moments in time, it also can produce data with moderate spatial resolution. The method carries limitations as well, however. First, it is still susceptible to social desirability and recall biases, just like retrospective self-report (Bastiaansen et al., 2019) though perhaps less so (Kahneman et al., 2004). Second, it depends greatly on participants understanding and recalling episode or location the researcher is referring to. Third, there is a tradeoff between resolution and participant burden. If more than (approximately) 10 episodes or locations are asked about, there is a real risk of participants getting irritated or bored with the questionnaire, to the extent that the validity of their responses declines.

2.1.3 Physiology

Emotions affect the processes of the human body. Thus, measurement technologies have been under development since the 1960's to capture emotions' physiological manifestations. Physiological measurement technologies can be categorized across two dimensions: central (brain) and peripheral physiology, and measures for lab settings with immobile participants, versus measures that capture emotion even when participants are moving.

While it is technically possible to measure central (brain) physiology while participants are moving, using mobile electroencephalography, muscle activity and the richness of visual stimuli completely overwhelms any signal due to emotion. Thus, central physiological measures of emotion are exclusively the domain of research in a laboratory. Peripheral physiological signals include heart rate, breathing rate, facial electromyography, and skin conductance. Heart rate and breathing rate are dramatically affected by the need of muscles for oxygenated blood, signals due to movement and posture overwhelms any sign of emotion, much as for brain activity. Thus, despite publications reporting heart rate measurements in mobile

settings (i Agustí et al., 2019; Kohn et al., 2019; Mitas, Cuenen, et al., 2020), heart rate and breathing rate measurement are also best reserved for lab settings.

This leaves skin conductance as a sole promising pathway to record emotion signals from participants who are moving around. Skin conductance is usually recorded from hands or fingers, where sweat secretion is uniquely responsive to emotional arousal (Braithwaite et al., 2015). By passing a small electrical current between two electrodes attached to palms or fingers, wrist-worn devices such as Shimmer GSR are able to accurately record fluctuations in skin conductance. The resulting signal is precisely time-resolved, and can thus also be spatially resolved, and is completely unbiased by conscious processes.

Skin conductance recording in the field has a robust track record. In the field of leisure, tourism, and hospitality experiences, we refer readers to the systematic review by Li et al. (2022). While previous skin conductance research has uncovered technical affordances as well as theoretical contributions, two important limitations have also come to light. First, while skin conductance on the palms is relatively robust to the physiological effects of physical activity, movement can mechanically affect the skin conductance signal (e.g., by loosening or pressing the electrodes ‘into’ the sweat). The resulting artifacts must be removed by trained researchers supported by detection algorithms like ArtifactZ (Bastiaansen et al., 2022) or EDAExplorer (Taylor et al., 2015). A second limitation is that skin conductance signal is unidimensional, representing emotional arousal exclusively. Skin conductance contains no information whatsoever about the valence or type of emotion being experienced.

2.2 Spatial resolution of emotion data

2.2.1 Overview

Combining emotion data with spatial data from any of the above sources accentuates, rather than compensates, the limitations of emotion measuring technologies. Emotion maps with any measure of validity are only possible with theoretically sensible combinations of emotion and location measuring technologies. In this paper, we review and illustrate three such combinations: experience reconstruction maps, emotion physiology maps, and emotion effectiveness maps.

2.2.2 Experience reconstruction maps

An experience reconstruction is a set of self-report emotion ratings given by participants after an experience, for each separate episode of that experience. In a spatial context, this means that participants report on their emotions each specific location they visited. A map can use polygons covering areas of interest (AOI) and shade them in different colors based on recalled ratings averaged over participants.

To our knowledge, few if any experience reconstruction maps exist. There are, however, instances of maps based on self-report using data from experience sampling. These are usually organized by municipal governments, such as the Happiness Map of London. A particularly rigorous example is the transportation happiness map of the Minneapolis St. Paul region in the United States. Unfortunately, only a non-reviewed methodological document about this effort exists (Fan et al., 2020). Other examples by Panek and colleagues are noteworthy for addressing transport routes (Pánek & Benediktsson, 2017) as well as neighborhood areas (Panek, 2019). As Pánek and Benediktsson (2017) point out, their maps address correlates of emotional experience, such as aesthetics (e.g., ‘ugly’) or overall evaluation (‘good’ and ‘bad’) rather than emotions themselves. Yet another approach to mapping emotions uses sentiment analysis linked to geolocated social media posts (Gupta, 2018). As with much social media research, however, this approach must include a within-user averaging step, as volume of social media postings is heavily skewed and thus risks creating maps shaped by within-individual autocorrelation.

2.2.3 Emotion physiology maps

Emotion physiology maps comprise a simple visualization of location-resolved emotion physiology measurement, usually skin conductance. Skin conductance measurements from single or (usually) multiple participants are either averaged over a regular (Shoval et al., 2018) or irregular spatial bin, or form the input to a kernel density function (Mitas, Mitasova, et al., 2020), portraying smooth gradients in skin conductance in a given location.

Emotion physiology maps have existed for over a decade. The first examples of which we are aware concerned museum exhibits where Bluetooth signals were used to determine participant location, while skin conductance was measured with a

special electrode-filled glove (Kirchberg & Tröndle, 2015). Soon thereafter, several urban studies groups began to record location using GPS alongside skin conductance or heart rate, producing the first emotion physiology maps in outdoor settings (i Agustí et al., 2019; Shoval et al., 2018). Soon thereafter, maps of indoor and outdoor museum tours (Mitas, Mitasova, et al., 2020) as well as roller coasters (Bastiaansen et al., 2022) appeared.

While these studies have been useful for identifying which locations during an experience are associated with the strongest emotional responses, limitations of GPS or Bluetooth location tracking and skin conductance reinforce one another. Skin conductance data is highly skewed and unidimensional, while location tracking data tends to be noisy. Thus, many participants (> 50) are needed to cover a study site with sufficient data so that the map represents true averages of emotional arousal across space. Furthermore, the resulting map affords no information about the emotional valence, content, or effect of the experience. These must be inferred from other data, usually obtained by self-response questionnaires or interviews.

2.2.4 Emotion effectiveness maps

In an attempt to visually integrate skin conductance, location, and self-response data, Mitas, Mitasova, et al. (2020) developed the *emotion effectiveness map*. This map essentially visualizes between-participant correlations between *local* within-participant skin conductance and *retrospective* self-reports. In this map, location-resolved skin conductance data is binned, either in evenly-spaced hexbins or AOI polygons. Within each bin, each participant's average skin conductance is computed. This results in a single number representing each participant's emotional arousal within each bin. A Pearson product-moment correlation of emotional arousal within each bin to each participant's self-response on a questionnaire item, for example emotional valence, is then computed. The outcome of these calculations is a correlation coefficient for each spatial bin, representing the between-participant relationship of retrospective self-reported emotion valence to emotional arousal within that bin. On the map, color indicates the direction and strength of the correlation coefficients within each bin.

The resulting map visualizes spatial differences in between-participant associations between momentary emotional arousal and self-reported valence. In other words, it shows *where* individuals with a more positive overall experience felt stronger emotions, and where they felt milder emotions. Besides conveying richer information, a second strength of emotion effectiveness maps is that in principle, they can be implemented with various self-reported experience evaluations besides valence. Mitas, Mitasova, et al. (2020) used intent to recommend for example, while Mitas et al. (2024) used self-reported connection with nature. Given that multiple methods exist of spatially resolving emotion data, each with inherent strengths and limitations, the objective of this paper is to compare and contrast the affordances and caveats of each of these methods of emotion mapping. In this paper we take the three methods presented here—experience reconstruction maps, emotion physiology maps, and emotion effectiveness maps—and for the first time, apply all three methods to the same location. In doing so, we answer the following questions:

- What are the relative affordances and limitations of experience reconstruction maps, emotion physiology maps, and emotion effectiveness maps?
- How do portrayals of an experience offered by reconstruction maps, emotion physiology maps, and emotion effectiveness maps differ?

It is worth noting that these data may also be compared *statistically*, as already demonstrated in comparisons of skin conductance to experience reconstruction (Strijbosch et al., 2021) and spatial change in skin conductance compared to spatial change in emotion effectiveness (Mitas, Mitasova, et al., 2020). The present paper focuses instead of *graphical representation* of data on maps. The implications of differences between statistical and graphical representations will also be touched on, though graphical representations remain the focus of the current inquiry.

3 Methods

The present study is based an intercept sample at a park-like fort site in the Netherlands, where we combined intake and exit questionnaires with mobile GPS location and skin conductance recording to collect data.

3.1 Data collection and sampling

We collected data at Fort Sabina, in the southwest of the Netherlands, on 14 and 15 September, 2024. We asked each visitor entering the site on foot from the main parking lot during opening hours if they would be willing to participate. We obtained informed consent according to procedures under approval TSB-RP698 of the University of Tilburg ethics committee. We asked participant who gave their consent to fill out an intake questionnaire on a tablet, and fitted them with sensors to measure their location and emotion arousal. Subsequently they were welcome to visit the fort as they had normally intended to. Finally, when they completed their visit and prepared to leave the site, we collected the sensors from them and asked them to fill out an exit questionnaire. Of 75 participants who provided data, either questionnaire, wearable, or GPS data were missing for 20 participants, making for a final sample of 55 participants. The primary technical error leading to missing data involved researchers accidentally turning off or on recording on a device at an unintended moment (e.g., turning recording off instead of on at the beginning of a visit). As such, participants with missing data are unlikely to have had a different experience from those who provided complete data.

3.2 Passive mobile sensor measures

To record location, we lent participants a smartphone with the popular workout application Strava that recorded GPS location once per 2 seconds. We separately up-sampled latitude and longitude data from 0.5 Hz to 16 Hz, the frequency of processed skin conductance data, using linear interpolation.

To measure skin conductance as a proxy for emotional arousal, we used Shimmer GSR+ wristbands to measure skin conductance, a proxy for emotional arousal, at 64 Hz. The wristband records skin conductance from 2 wires which attach to pre-gelled electrodes worn on the fingers. The skin conductance signal was cleaned from motion artifacts using the ArtifactZ function of the Breda Experience Lab Toolbox (Bastiaansen et al., 2022), then downsampled to 16 Hz. Tonic changes in the signal due to temperature and wearing of the device were filtered out using deconvolution (Benedek & Kaernbach, 2010). Skin conductance signals were also Z-standardized to cancel out differences in skin responsiveness between participants.

3.3 Self-report measures

3.3.1 Emotion before and after visit

We used self-report questionnaires to measure participants' emotional state at the beginning and end of their visit. In the pre-visit questionnaire, we asked participants to rate to what extent they felt each of a list of emotions “right now.” In the post-visit questionnaire, we asked them to recall to what extent they had experienced each emotion “during their visit.” Both questionnaires included an identical 5-point response scale ranging from “not at all” to “extremely,” and an identical list of 12 emotions based on the SPANE (Diener et al., 2010) including positivity in general, negativity in general, and several specific positive and negative emotions such as joy, contentment, anger, and sadness. As intended by the design of the SPANE, we computed an average of positive items to create a positive emotion index. We omitted negative emotions from the current analysis.

3.3.2 Experience reconstruction

We selected 6 locations as being clearly recognizable and likely to be visited by many, though not all, visitors. We photographed each location from an angle a visitor would be likely to see and recognize. Photos and names of each location were included in the exit questionnaire. For each location, we asked participants to recall and rate how they felt while there on two 5-point semantic differential items reflecting valence (»very negative« to »very positive«) and arousal (»calm« to »excited«). This implementation of experience reconstruction has been validated in previous research (Strijbosch et al., 2021).

4 Findings

4.1 Descriptive findings

The 55 visitors to Fort Sabina spent an average of 70 minutes at the site. Their self-reported positive emotions were mild before (mean = 2.77, sd = 0.71) and elevated after (mean = 3.27, sd = 0.65) their visit.

4.2 Mapping the data

4.2.1 Experience reconstruction map

To create the experience reconstruction maps (Figure 1), we first created polygons to correspond with the approximate extent of each of the locations addressed by photo-questions in the questionnaire. This polygon layer was visualized over a base map layer. For each location, responses to the valence item (»negative« to »positive«) were averaged together. Polygons were colored according to the percentile rank of their average valence, so the most positive-average locations were colored yellow, and the least positive-average locations colored blue. The resulting map showed that the area in front of the cafe is the most positively-rated of the fort, while the high wall facing the estuary on the west side of the fort, featuring five large gun installations, was the least positively rated. Incidentally, this western wall is also rated as the least arousing. The exhibitions at the eastern side of the barracks were rated as the most arousing.



Figure 1: Experience reconstruction maps of Fort Sabina (valence left, arousal right)

Note: Yellow indicates highest ratings of arousal and valence; blue indicates lowest ratings of arousal and valence.

Source: Authors

4.2.2 Emotion physiology map

Based on the high resolution of GPS and skin conductance data, a hexbin visualization with relatively small hexbins of 6 meters—following a heuristic of two times the typical margin of error for GPS—was chosen. Thus, the visualization aimed to balance the noise-to-signal ratio of spatially resolved skin conductance data, keeping noise in both skin conductance and GPS in mind. Our emotion physiology map (Figure 2) shows several distinct areas of relatively high arousal. The largest covers the entire cafe area, extending to the back of the barracks and exhibition space, and outside over the northern wall of the fort, where visitors can climb for the views of the landscape and the delapidated fort warehouse across the road. On the right side of Figure 2, we also averaged skin conductance over the same AOI as the emotion effectiveness map.

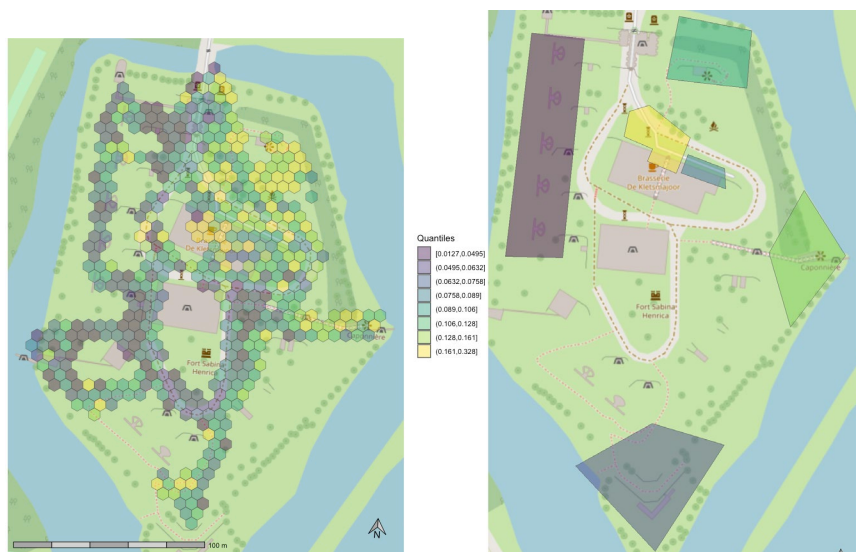


Figure 2: Experience physiology maps of Fort Sabina (6 meter hexbin left, AOI right)

Note: Yellow indicates highest quantiles of arousal; blue indicates lowest quantiles of arousal.

Source: Authors

4.2.3 Emotion effectiveness map

The emotion effectiveness map is based on the same data as the emotion physiology map, but subsequent steps to the data analysis are added. Instead of all skin conductance data being averaged within each bin, regardless of participant, skin conductance data are instead averaged within each bin, within each participant. A filter is used to exclude bins with fewer than 5 participants. Then, a correlation coefficient is computed between the average skin conductance within each hexbin and self-reported positive emotion change from before to after the visit. These correlation coefficients therefore represent the between-participant relationships between skin conductance within each hexbin and positive emotions recalled from the experience as a whole. This procedure leads to a map with implausible differences, from very positive to very negative correlations, in adjacent hexbins. Emotion does not plausibly swing from very positive to very negative in an average participant from meter to meter. Thus, the size of the hexbin is incrementally increased (in this case to 12 meters) until most abrupt adjacent differences in correlations disappear. As a final step to compare the emotion effectiveness map with the emotion reconstruction map, the procedure was applied to the six AOI polygons used to visualize the experience reconstruction map (Figure 3).



Figure 3: Emotion Effectiveness Maps of Fort Sabina (10 meter hexbin left, AOI right)

Note: Blue and dark green correlations are negative; Yellow and light green are positive.

Source: Authors

The resulting emotion effectiveness maps show somewhat differing areas of emotion effectiveness, depending on if correlations are computed over hexbins or AOI polygons. The hexbin map suggests the entrance bridge, the passage to the back of the barracks, and part of the grassy north-eastern wall are locations where the most positive emotion was associated with the greatest arousal. In other words, in these places, participants with higher arousal reported feeling more positive afterward. The eastern caponniere, the northwest corner of the fort wall, and the western section of the barracks show the most negative correlations between arousal and later self-reported positive emotion. In terms of AOI's, all show a positive correlation between arousal later self-reported positive emotion, except the southern double caponniere. The highest positive correlation, as the highest self-reported valence and the highest physiological arousal, occurred at the cafe.

5 Discussion

We visualized emotion data using recalled and reconstructed self-report, as well as wearable physiological measurement, using three different approaches to spatial resolution of emotion data. The goal was to compare and evaluate existing technologies for mapping emotion in terms of their affordances and limitations. While developments such as wearables to measure physiology afford new possibilities, they also introduce new caveats which must be considered in their application. Our maps show that each of the mapping techniques studied is useful, albeit in different situations. This is an important contribution to academic efforts at emotion mapping, as most previous studies visualized one (Shoval et al., 2018) or at most two (Mitas, Mitsova, et al., 2020) of these maps as the »best« visualization with existing technology. Furthermore, we were able to explain *how* these different map portrayals of emotion over space differ.

5.1 The value of physiology maps

Emotions are multidimensional by definition (Rosenberg, 1998) and physiology is just one of these many dimensions, which related to but not identical to other dimensions (Mauss & Robinson, 2009). A skin conductance recording is therefore no more or less »true« than any other emotion recording. It indexes physiological arousal from emotion, no more and no less. The emotion physiology maps in the present study demonstrate that this metric can be combined with GPS data for a

precise, but unidimensional map of individuals' emotional experience across locations. It is clearly visible where participants became emotional on average, and to what extent, but not why, with which emotion, and with which consequence. Therefore, our maps show that an emotion physiology map might be best when researchers would like to know exactly where and when individuals become emotional as they visit a site. In practical contexts, a skin conductance map is therefore useful when a manager has a fairly clear idea of *which* emotions their participants feel during a visit, but not *where* these emotions become the most intense.

Thus, important emotion dimensions such as valence and action tendency (Fredrickson, 1998) are not present in this map. It follows that emotion physiology maps are not useful when researchers would like to know *in what way* participants became emotional. This is especially important as in the current dataset, the physiology maps had more in common with the self-report-based experience reconstruction maps of *valence* rather than arousal. This apparent discrepancy could be due to recall bias or to misinterpretation of self-report item wording by participants, but this is merely conjecture. Such conjecture highlights that a physiology map must be viewed as a rather unidimensional portrayal of emotion, precise in resolution but minimal in content. It cannot be interpreted without companion self-report data or companion knowledge from previous research.

5.2 The *added* value of emotion effectiveness maps

To directly address the limitation that physiology maps show no emotional content besides arousal, emotion effectiveness maps visualize correlations between momentary physiology and retrospective self-report. As such, they add some emotional content to the visualization, by showing where higher arousal was associated with, in our example, a more emotionally positive experience overall. Previous emotion effectiveness maps instead looked at correlations with experience evaluations (Mitas, Mitasova, et al., 2020). Thus, we extend previous research by linking two different dimensions of emotion—momentary arousal and recalled valence—in a single map, a more coherent outcome than momentary arousal and evaluation, which is an outcome rather than a component of emotion.

Adding a second dimension of emotional experience to physiology maps has clear added value, because stakeholders can now see where something positive and emotional was occurring, and where something was not working well. For this added value, emotion effectiveness maps present tradeoffs. First, the spatial resolution and coverage both decrease, necessarily, as the data represent correlations in between-participant variation. Also, the visualization does not control for within-participant autocorrelation. Thus, the map may show different effects than multilevel statistical models, with the latter being a truer representation of within participant development. Statistical analyses of the present data are beyond the scope of this article, but such analyses in conjunction with emotion effectiveness maps have been conducted in previous studies (Mitas, Mitsova, et al., 2020).

Managers can use emotion effectiveness maps when they have at least anecdotal knowledge of the site and which emotions visitors generally experience, but would like to have a heuristic of locations which work »well« or »less well.« However, here also they need to be aware of whether each location is generally negative or positive for the average visitor, and which stimuli at that site might be dominant in triggering emotions.

5.3 Caveats

When synchronizing emotion and location data, there is often an unstated assumption that something about the location triggers emotions in participants. That is, the location contains stimuli which make people emotional, such as a beautiful building or a busy crowd (Bastiaansen et al., 2019). Since maps show participants averaged together, it is reasonable to assume that individual variations unrelated to location such as a participant having an argument with their partner, while another participant becomes hungry, and so forth, are averaged out. However, it cannot be assumed that peaks in emotion in the average participant shown on the map are triggered by any specific physical feature of that location. For example, we saw relatively high emotion at entrance/exit gates of the fort. These may have been triggered by the massive gatehouse, but also the concealing and revealing of the beautiful grassy lawn around the barracks. It is obvious but worth reiterating: emotion maps link emotion to location, not to a specific stimulus.

Second, as noted in the caveats applying to emotion effectiveness maps, mapping data is not the same as analyzing it statistically. This means that differences between locations or participant groups, or even correlations between emotion signals, remain untested against their standard errors to determine statistical significance. While this information could be useful to both researchers and practitioners, there is another more serious limitation to merely mapping without statistical modeling, namely that within-participant autocorrelation is not controlled for. This is especially a problem for emotion physiology and experience reconstruction maps, where there is no between-participants variable brought to bear on within-participant effects. As Mitas, Mitasova, et al. (2020) demonstrate, the conclusions implied by an emotion effectiveness map and an appropriate multilevel model can be similar, but are not identical.

The third major caveat is that emotion is multidimensional, and to say that one place is more »happy« than another ignores the need to specify which *dimension* of emotion the »happy« is referring to. Is it »happy« because individuals express many positive emotional words in their verbal expressions there, or because they recall having been happy there when asked at the end of the day? We know from Zajchowski et al. (2016) and Mauss and Robinson (2009) that these comprise separate dimensions of the complex phenomenon of emotion. With the current findings, showing that maps based on self report and physiology lead to different visualizations and different conclusions, we extend these theoretical distinctions to emotion mapping. Our contribution is to assert that there is no one »emotion map« or »happy map.« Rather, multiple valid »emotion maps« are possible, each with a different ideal use case.

Acknowledgements

We appreciate the input of Keri Schwab and Marcel Bastiaansen to the conceptualization of this research, Jona Broothaerts, Geert Coninx, Wim Debaene, Eric Goosen, Anne-Wil Maris, and Nico Verwimp to the data collection, and Sait Durgun and Hans Revers to the data processing and analysis.

References

- 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.

- Benedek, M., & Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47(4), 647-658.
- 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*.
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D.-w., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, 97(2), 143-156.
- Fan, Y., Ormsby, T., Wiringa, P., Liao, C.-F., & Wolfson, J. (2020). Visualizing Transportation Happiness in the Minneapolis-St. Paul Region.
- Fredrickson, B. L. (1998). What good are positive emotions? *Review of general psychology*, 2(3), 300-319.
- Gupta, N. (2018). *Exploring happiness indicators in cities and industrial sectors using Twitter and Urban GIS data* [University of Warwick].
- i Agustí, D. P., Rutllant, J., & Fortea, J. L. (2019). Differences in the perception of urban space via mental maps and Heart Rate Variation (HRV). *Applied Geography*, 112, 102084.
- Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306(5702), 1776-1780.
- Kirchberg, V., & Tröndle, M. (2015). The museum experience: Mapping the experience of fine art. *Curator: The Museum Journal*, 58(2), 169-193.
- Kohn, L., Dastageeri, H., Bäumer, T., Moulin, S., Müller, P., & Coors, V. (2019). *Hot or not – identifying emotional “hot spots” in the city* 3rd International Conference on Smart Data and Smart Cities, Delft, Netherlands.
- Li, S., Sung, B., Lin, Y., & Mitas, O. (2022). Electrodermal activity measure: A methodological review. *Annals of tourism research*, 96, 103460.
- Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & emotion*, 23(2), 209-237.
- Mitas, O., Cuenen, R., Bastiaansen, M., Chick, G., & van den Dungen, E. (2020). The War from both Sides: how Dutch and German Visitors Experience an Exhibit of Second World War Stories. *International Journal of the Sociology of Leisure*, 3(3), 277-303. <https://doi.org/https://doi.org/10.1007/s41978-020-00062-3>
- Mitas, O., Mitasova, H., Millar, G., Boode, W., Neveu, V., Hover, M., van den Eijnden, F., & Bastiaansen, M. (2020). More is not better: The emotional dynamics of an excellent experience. *Journal of hospitality & tourism research*, 1096348020957075.
- Mitas, O., Wielenga, B., Evers, K., Farkić, J., Ward, D. P., Klijs, J., & Bastiaansen, M. (2024). *Relaxing or exciting? Mobile emotion recording in natural settings*. 37th Bled eConference., Bled, Slovenia.
- Panek, J. (2019). Emotional and subjective volunteered geographical information. In *Crowdsourcing: Concepts, Methodologies, Tools, and Applications* (pp. 1347-1362). IGI Global.
- Pánek, J., & Benediktsson, K. (2017). Emotional mapping and its participatory potential: Opinions about cycling conditions in Reykjavik, Iceland. *Cities*, 61, 65-73.
- Rosenberg, E. L. (1998). Levels of analysis and the organization of affect. *Review of general psychology*, 2(3), 247-270. <https://doi.org/https://doi.org/10.1037/1089-2680.2.3.247>
- Shoval, N., Schvimer, Y., & Tamir, M. (2018). Real-time measurement of tourists' objective and subjective emotions in time and space. *Journal of Travel Research*, 57(1), 3-16.
- Strijbosch, W., Mitas, O., van Blaricum, T., Vugts, O., Govers, C., Hover, M., Gelissen, J., & Bastiaansen, M. (2021). When the parts of the sum are greater than the whole: Assessing the peak-and-end-theory for a heterogeneous, multi-episodic tourism experience. *Journal of Destination Marketing & Management*, 20, 100607.
- Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015). Automatic identification of artifacts in electrodermal activity data. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),
- Tuan, Y.-F. (1977). Space and place: The perspective of experience. *U of Minnesota P*.
- Zajchowski, C. A., Schwab, K. A., & Dustin, D. L. (2016). The Experiencing Self and the Remembering Self: Implications for Leisure Science. *Leisure Sciences*, 1-8.

BUSINESS BENEFITS AND CUSTOMER PERCEPTIONS IN CUSTOMER SERVICE CHATBOTS

JASON VAN HAMOND,¹ DENYS BESPALKO,¹

MARGOT NEGGERS,¹ ALINDA KOKKINO²

¹ Breda University of Applied Sciences, Data Science & AI, North-Brabant,
the Netherlands

232567@buas.nl, 231452@buas.nl, neggers.m@buas.nl

² Breda University of Applied Sciences, ABEL Research team Zhan, North-Brabant, the
Netherlands

kokkinou.a@buas.nl

The implementation of chatbots offers benefits but also causes risks, making it important for companies to understand their customers' perceptions. Benefits and risks were previously researched, but were often limited to specific features. The focus of this research paper is to determine whether businesses have a comprehensive understanding of potential dangers, benefits, and consumer perceptions regarding customer service chatbots. The objectives are to find and understand the relationship between company expectations and customer perceptions, and to describe the risks and benefits associated with customer service chatbots. The conducted thematic analysis on interviews with companies and customers showed that the expectations and customer perceptions align. This analysis describes how overall chatbot implementation and quality of interactions can positively and negatively influence customer loyalty, interest, and experience. In conclusion, companies seem to comprehend the perspectives of their customers and are prepared for the benefits and risks regarding chatbots in customer service.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.6](https://doi.org/10.18690/um.fov.4.2025.6)

ISBN
978-961-286-998-4

Keywords:
customer service chatbots,
business understanding,
customer perceptions,
benefits,
risks



University of Maribor Press

1 Introduction

In recent years, Artificial Intelligence (AI) has influenced various factors in our daily lives. AI is implemented in various systems and programs, including chatbots. Chatbots are programs that often use Natural Language Processing for communication in comprehensible human language, but they can also work through keywords. Their use has become more widespread, especially in business settings like customer service, entertainment, and e-commerce (Hasal et al., 2021).

Businesses can benefit from chatbots by providing continuous customer service, personalizing interactions, reducing costs, automating routine tasks, improving customer interaction, personalization, and enhancing data utilization (Hasal et al., 2021; Kedi et al., 2024). This may even provide automation of repeated tasks, thus ensuring less need for hiring additional staff. These benefits might increase customer engagement, service, satisfaction, and loyalty. While chatbots provide benefits for companies, they also introduce new risks. Chatbots often learn from user data, which can create security risks. The study of Hasal et al. (2021) on security, privacy, data protection, and social aspects reveals that great responsibility is involved with chatbot implementation.

The use of chatbots may also bring positive and negative outcomes in interactions. Chatbots may reduce stress, improve tasks, and increase trust, but can also reduce trust, frustrate users, or provide overall unpleasant interaction (Ranieri et al., 2024). Negative outcomes might occur when the chatbot doesn't fit in its own context, which might cause unsatisfied expectations (Gnewuch et al., 2017). To ensure positive outcomes, evaluating company comprehension of the benefits, risks and overall company perceptions is important, which is the objective of the current research. This involves comparing the expected advantages, risks, and perceptions of chatbot implementation from the perspectives of companies with those of the customers. Customer loyalty can also be impacted by chatbot benefits (Kedi et al., 2024).

This study's research question is the following: 'Is there a relationship between the company expectations regarding the benefits of customer service chatbots and the customer perceptions of their experiences?' By comparing perspectives, this study aims to provide insights into how well companies comprehend chatbot risks and

benefits. A hypothesis indicating a relationship between interaction and loyalty will also be conducted to provide further insights into the research question.

2 Literature Review

Selamat and Windasari (2021) found that customers value chatbot features like responsiveness, personalization, and human-like interaction. They narrowly focused on conversations where human conversation styles are mimicked, leaving other important features unexplored. In contrast, Johari and Nohuddin (2021) identified seven features preferred by users for good quality chatbots, including functionality, efficiency, humanity, effectiveness, technical satisfaction, and ethics. The study was only limited to vital features and left the eventual benefits unexplored. Similar findings were concluded in the study by Jenneboer et al. (2022), which indicated that customer satisfaction and loyalty could be raised through optimization of the customer experience. This optimization could best be done through offering quick and relevant information and ensuring system, information, and service quality. These results might indicate that there is a possible connection between chatbot interaction and customer loyalty. In the study, there are some limitations regarding a lack of focus on age demographics and functions and contexts of a chatbot. The current study aims to address these limitations by ensuring deeper investigation into customer preferences and perceptions, company expectations, and business benefits regarding customer service chatbots.

Gümüş and Çark (2021) conducted a study on the effects of chatbots on customer satisfaction, where results indicated that perceived ease of use, usefulness, and enjoyment positively influenced customer experience and behavior. A study on chatbot security by Hasal et al. (2021) revealed other impacts on customer attitudes. That study reveals that security practices should be implemented with chatbot usage within a company. Communication with chatbots can raise security issues, which might cause concern among customers and negatively influence a company. The study by Tran et al. (2021) found topics unexplored by the previous two studies, revealing that a user's attitude towards chatbots is more positive than towards human assistants after chatbot implementation. Chatbot implementation might possibly lead to dissatisfaction when customers must wait to receive support. However, it is also indicated that the difference of products and services also impacts this attitude.

All three studies had some limitations, where the study by Gümüş and Çark (2021) had cross-sectional data, a lack of focus on specific sections, and used moderate sample sizes. Meanwhile the study by Hasal et al. (2021) made use of data manipulation, as companies are unable to share security secrets. The study by Tran et al. (2021) contained sampling limitations and a lack of focus on different types of chatbots. This research addresses each of these limitations by ensuring a focus on generative and keyword-based chatbots along with applying multiple research methods.

A study on user experience of ChatGPT by Skjuve et al. (2023) revealed that a chatbot's effectiveness, usefulness, and interactions improve user experience. Meanwhile, user experience is negatively impacted when the chatbot is irrelevant, unresponsive or has poor interactions.

3 Methodology

3.1 Data Collection

3.1.1 Qualitative Approach

Qualitative approaches are used as the main source of data for answering the question. Interviews were conducted with two male and three female customers with previous chatbot experience and four company representatives interested in chatbot implementation. These companies include an instance of homecare, technical, and chatbot focused companies. Two of the companies were selected through personal connections, and the other two were found through networking events. Through the interviews, the perceptions of customers and company expectations regarding customer service chatbots and their benefits were collected through questions focused on a customer's experiences and a company's expectations of benefits and risks of chatbots. Qualitative approaches were selected with the intent to receive a deeper understanding of the reason why companies or customers feel a certain way. Further analysis was provided for customer perceptions by a company that could provide information on how their customers perceived chatbot usage in the past, which was first evaluated for its reliability and then combined with customer perceptions. To ensure reliability and validity, interviews were conducted in a flexible

way so that participants could answer questions based on their knowledge or experiences.

3.1.2 Quantitative Approach

In the current study, a hypothesis suggesting interactions to impact customer loyalty is tested through quantitative research. Interaction quality is measured through likert-scale questions describing a customer's opinion on the statements related to efficiency, conveniency, and quality of interactions. With the survey focused on customers, responses from 200 participants found through Prolific of various ages and with previous chatbot experience were collected. Loyalty was described using the participant's agreements to the following statements: "Chatbots that I previously used have encouraged me to return to the business," "Chatbots have enhanced my overall loyalty and trust in the company's ability to provide good services," and "A chatbot's quality can impact my view on a company". Meanwhile interaction was described using answers on the following statements: "Chatbots have made my interaction with businesses more efficient and convenient," "Chatbots often improve the quality of customer service," "Chatbots provide faster service compared to human customer support," and "Chatbots efficiently resolved my issues without the need of further assistance." These questions help measuring customers' opinions on interaction quality and a chatbot's impact on loyalty. The questions and measures used are self-developed; they are designed to best capture and describe the key components of interaction quality and customer loyalty in chatbot usage. While not drawn from existing literature, the items were aligned with the topics while being easy to understand for participants.

3.2 Data Analysis

First, the interviews were transcribed, and words related to the research objective were highlighted inside the file. Python was used to automate extraction of these highlighted phrases and saving into a comma-separated values (CSV) file, ensuring consistency and efficiency in data processing. Thematic analysis was then performed by assigning themes and keywords and summarizing the main takeaways of the phrase. These themes and keywords were further analyzed, interpreted, and compared. Customer perceptions were collected through interviews with chatbot users and a company's research on customer interactions with chatbots. The

company's findings were shared to support the interviews and must be compared to evaluate reliability and validity.

4 Results

In this section, the results of the qualitative and quantitative analyses will be explored and evaluated. Qualitative results are divided into three groups: customer perception based on previous analysis of a company, actual customer perception, and the expected benefits from companies. Important patterns of this analysis were tested further using a conducted survey on chatbot interactions.

4.1 Topic Distribution

Thematic analysis revealed differences in the distribution of key topics between customer perception and company expectations. Customers frequently mentioned chatbot implementation and service quality as important aspects. Chatbots were often compared to human services, which might indicate that this comparison could play a role in customer experience. Many customers had similar ideas, shortly described by a male interviewee: "The answers are great, but I always have this gnarly feeling when I talk to a chatbot or any AI as it still isn't human so the answers could still be wrong". Overall customer perceptions and loyalty were not mentioned in interviews, but this does not outline its importance.

Company expectations on benefits and customer perceptions regarding customer service chatbots are vital for this study. Thematic analysis revealed common company expectations in terms of benefits, risks, or overall customer perceptions. The analysis shows that companies mentioned service and implementation many times in interviews. That could mean that companies expect a chatbot's service and implementation quality to greatly influence risks, benefits, and customer perceptions. Customers are often expected to compare chatbots with human agents. The benefits a chatbot provides to businesses are very important to companies, as it was mentioned frequently in the interviews. Loyalty, security, and overall customer perceptions were not mentioned many times in interviews but might still provide valuable insights.

4.2 Main Takeaways

4.2.1 Customer Perception through Company Analysis

The interview about company-analyzed customer perception has revealed the quality of information and quality of overall implementation to be the most important themes. “People often drop out if the quality of the asked information is not good,” said a male participant. “The quality of a conversation is crucial.” The participant mentioned many vital subjects related to service, such as correctly answering Frequently Asked Questions (FAQ) having a positive effect on a customer's perceptions. They also mentioned how correct information must be translated correctly.

Participants regularly stated that overall implementation may have a strong impact on customer perceptions. Keyword-based chatbots typically provide incorrect answers and cause negative associations. Common User Interface (UI) decisions are typically perceived as annoying, and people from older generations are more frequently negative about chatbot implementation. Participants describe that bad experiences, interactions, or inaccurate information could harm a company's reputation and reduce interest in its products or services. Fast and accurate information was indicated to provide positive impacts on the view of your company.

4.2.2 Customer Preferences

In addition to interviews related to previous company research, interviews on customer preferences have been personally conducted for comparisons and have indicated many vital details for this research. Their statements indicate that customers prefer to receive short answers to simple questions and long answers to complex questions. Customers commonly believe that chatbots can handle urgent or complex issues with accuracy, but still occasionally become dissatisfied when a chatbot does not provide a full answer, which, along with not understanding questions, causes annoyance.

Many customers have expressed a dislike for keyword-based chatbots due to the lack of a broad selection of responses and choices. Having to go through a chatbot before the ability to talk to a human also causes many negative perceptions. A customer is

satisfied once the answers are accurate, brief, and straight to the point. Customers feel like chatbots are the best fit for simple queries, whereas human services are preferred for more complex questions. The answers of human services are perceived to be more correct than chatbot answers.

4.2.3 Company Expectations

Companies expect that the increase in benefits is caused by a chatbot's speed, ease of use, clarity in response, focus on service, quick solutions, presentation, and ability to answer FAQs. They are expected to solve the problem and find products more easily. Companies expect chatbots to fail more likely on broad and complex questions, and a loss of interest is expected with insufficient implementation. They also stated that poor UI quality might cause aggression or frustration for the customer. Chatbots are anticipated to influence loyalty by ensuring easy contact or increasing interest in products or services. Wrong implementation of chatbots is expected to influence negative characteristics, yet right UI decisions, information selection, or a combination of a chatbot with human agents or AI is expected to bring new benefits. Not all companies might receive advantages for using a chatbot, since company demographics can often differ. Security issues are expected to impact customer perceptions due to possible leakage of data.

4.3 Quantitative Findings

Important patterns were found during qualitative data analysis and were evaluated with quantitative methods to judge reliability. A frequent pattern in qualitative data was good interaction with chatbots causing more loyalty. Survey questions were assigned to a specific theme and further analyzed. The distribution analysis in Figures 1 and 2 show the level of impact and improvement on loyalty and interaction based on the answers of male and female participants, aged between 18 and 71 years. Figure 1 shows a chatbots having a medium to high impact on loyalty, and Figure 2 shows that chatbots cause an improvement on interaction quality.

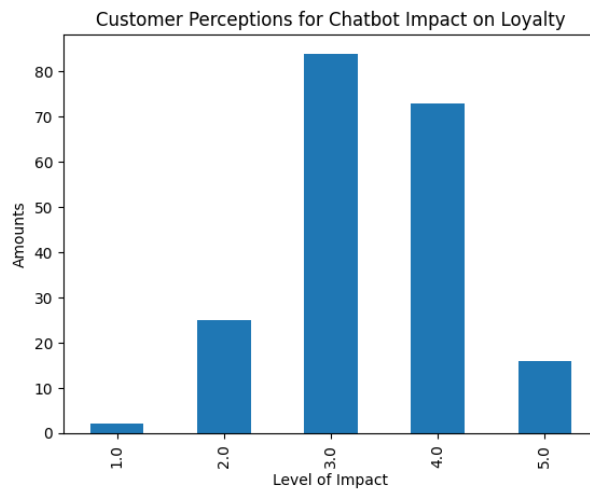


Figure 1: Histogram of customer opinions: chatbot impact on loyalty

Source: Own

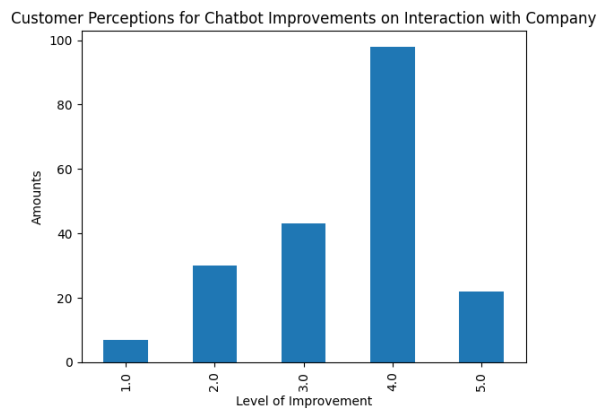


Figure 2: Histogram of customer opinions: chatbots improve interaction

Source: Own

Figure 3 tests a hypothesis indicating a relationship between interaction and loyalty, in which Pearson and Spearman both indicated that a statistically significant positive relationship is found. With a correlation of 0.71, a p-value below 0.001, and a chi-square statistic of 172, a strong correlation and association are indicated.

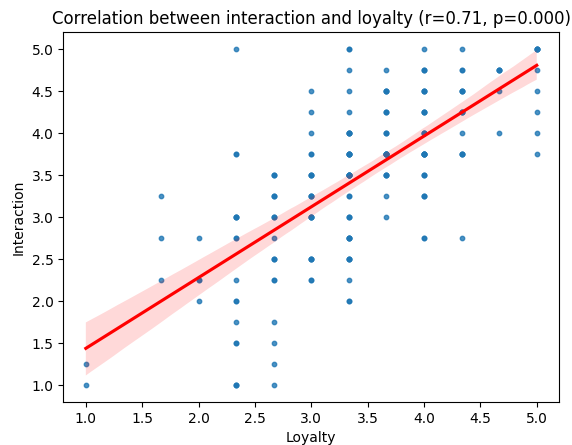


Figure 3: Pearson correlation

Source: Own

5 Discussion

To answer the research question, thematic analysis was performed on the qualitative data. The results of this analysis were then compared with other groups.

5.1 Customer Comparisons

Before comparing company expectations with customer perceptions, it is important to compare results of the customer perceptions according to a company with the preferences according to customers. The results largely aligned, showing customers to be more positive towards chatbots providing high-quality, quick, accurate, and direct answers, increasing loyalty. However, incorrect answers, bad interaction, or poor-quality conversations may lead to a decrease in loyalty and loss of interest for customers. These findings seem to align with the study by Jenneboer et al. (2022), that study mentioned that speed and relevancy of information and service quality can impact the loyalty of the customer. The comparisons also showed that while customers prefer human interaction, demand decreases when chatbots provide high quality service. As stated previously in the results section, keyword-based chatbots cause negative associations among customers, and UI implementation greatly impacts the customer perceptions too. The age demographic of a company should

also be evaluated before implementing a chatbot to ensure that the company benefits from it.

5.2 Final Comparisons

Using the findings of customer perceptions, the comparison between company expectations and the actual customer perception can be evaluated. In this section, comparisons are made between the various results and existing literature. The important comparisons will be made first and afterwards interpreted into the research question.

Regarding chatbot service, both company expectations and customer preferences align well with each other. Both groups expect chatbots to be most effective when providing quick, direct, accurate, and easy-to-access information. This aligns well with the studies from Gümüş & Çark (2021), indicating that the ease of use, usefulness, and enjoyment influence customer experience, and Jenneboer et al. (2022), describing an impact on customer satisfaction through offering quick and relevant information along good quality service. The study by Skjuve et al. (2023) also aligns with the findings, mentioning how a chatbot's efficiency, usefulness, relevancy, accuracy, interaction quality, and intelligence improves user experience. Both groups expect early chatbots to perform poorly due to insufficient data and poor-quality conversations, potentially harming customer interest. Misalignment was also present, where companies expect chatbots to solve customer problems, while customers often experience inaccurate or incomplete answers to their questions. This misalignment might be due to bad experiences with different implementations of chatbots or companies having high expectations for chatbots.

Companies and customers seem to have aligned ideas related to loyalty. Both parties expect loyalty to decrease with poor or inaccurate information and to increase with quick and accurate answers and easier contact. The increase in loyalty due to quick and accurate answers is a finding also mentioned in the study of Jenneboer et al. (2022). Companies expect chatbots to have the ability to sell new products to customers, although customers did not mention this. This misalignment could be either due to customers lacking that experience or customers not mentioning this idea.

Ideas about human services compared to chatbots seemed to mostly align. Both parties argue that customers prefer human services over chatbots due to accuracy and better help. Chatbots are preferred for simple questions, and demand for human services will decrease while chatbot quality increases. Some aspects still do not align, like companies expecting chatbot preference for language barriers, the feeling of personal contact reduction, and customers mentioning chatbot preferences when human assistance takes too long. These are not mentioned by opposite groups, likely due to a lack of data or experience. The findings seem to contradict the findings of Tran et al. (2021), describing the attitudes towards chatbots to generally be more positive, while the findings of this analysis indicate a likely more positive attitude towards human assistants.

The findings indicate that most ideas of customers and companies regarding implementation seem to align as well. Both mention the negativity toward keyword-based chatbots, the impact of UI changes, relevancy of information, varied implementation per company, and direct, brief, and accurate answers causing satisfaction. These findings seem to align with the study by Jenneboer et al. (2022), describing how quick and relevant information impacts customer experience. Some statements did not align, such as forcing customers to use a chatbot before offering human services, causing dissatisfaction, an idea also mentioned in the study by Tran et al. (2021). The customers' preferences toward human-like speech, which Selamat and Windasari (2021) also found in their study, was not stated to be expected by the company either, so accurate comparisons can't be made related to that topic.

Opinions on security and business benefits did not align, where companies expect chatbots to cause security risks, causing customer mistrust. However, customers have not mentioned these topics in the interviews, likely showing either a sampling bias or customers overall not thinking about it. The study of Hasal et al. (2021) does mention the importance of security and responsibility with chatbot implementations, indicating that interviewed customers have no previous experience on this topic. Security not being mentioned by customers during interviews does not show the importance of that topic but rather indicates that they don't often think about the topic.

The findings of the quantitative analysis indicate a pattern where chatbot interaction impacts customer loyalty, which was then further evaluated through hypothesis testing using quantitative data. The results of Pearson and Spearman tests indicate a strong positive relationship, and the chi-square test also points towards an association between the values, indicating the relationship between chatbot interaction and loyalty to be statistically significant. This indicates that increasing the chatbot interaction quality should increase the loyalty of the customer, helping a further understanding of how to achieve extra benefit through chatbots.

5.5 Limitations

Some limitations were present in this research, such as sampling limitations within the qualitative methods. These limitations happened due to the difficulty of finding companies interested in participating and strict limitations on time. This has likely influenced some interpretations due to some statements not being addressed by the other party. Second, the research was conducted within the Netherlands, meaning that there could be potential bias in opinions. Future research could expand the sample size and include a more diverse scope.

6 Conclusion

This research paper aims to answer the following research question: "Is there a relationship between company expectations regarding the benefits of customer service chatbots and customer perceptions of their experiences?" Through thematic analysis, a strong alignment between company expectations and customer perceptions is found, indicating a good company comprehension of risks, benefits, and customer's perceptions concerning chatbot implementation. Inconsistencies were still found, where security and business benefits seem to be a bigger concern to companies. Results aligned well with previously conducted studies and ways of improving benefits were also concluded, such as providing quick, accurate, and easy-to-access information increasing loyalty. Quantitative analysis allowed further investigation to strengthen the current research findings, establishing a statistically significant relationship between chatbot interactions and increased loyalty.

The risks of implementing a chatbot were also found in the results, such as customers preferring human agents over chatbots and loyalty decreasing with poor or inaccurate answers. The overall implementation of a chatbot, like UI decisions, information quality, and interaction quality, also impacts the experience of a customer.

This research paper has mentioned important features that can be applied to increase customer service quality. The recommended ways are by optimising the design for increasing loyalty, evaluating the preference of your target audience between humans and chatbots, and ensuring that the chatbot is implemented in the right way. Each of these should be evaluated based on your group demographics, as the design and implementation of a chatbot should reflect the preferences of your own customers.

Future research might analyse the benefits, risks, customer perceptions, and company expectations in different types of environments, such as e-commerce or transactional chatbots. Further insights regarding the best implementation might also be very useful to research in the future. This study contributed to a better understanding of the risks, benefits, customers' perceptions, and the company's knowledge surrounding the implementation of customer service chatbots.

References

- Gnewuch, Ulrich, Morana, Stefan; and Mädche, Alexander, Towards Designing Cooperative and Social Conversational Agents for Customer Service (2017). ICIS 2017 Proceedings. 1. <https://aisel.aisnet.org/icis2017/HCI/Presentations/1>
- Gümüş, N., & Çark, Ö. (2021). The effect of customers' attitudes towards chatbots on their experience and behavioural intention in Turkey. *Interdisciplinary Description of Complex Systems*, 19(3), 420–436. <https://doi.org/10.7906/indecs.19.3.6>
- Hasal, M., Nowaková, J., Saghair, K. A., Abdulla, H., Snášel, V., & Ogiela, L. (2021). Chatbots: Security, privacy, data protection, and social aspects. *Concurrency and Computation: Practice and Experience*, 33, e6426. <https://doi.org/10.1002/cpe.6426>
- Jenneboer, L., Herrando, C., & Constantinides, E. (2022). The impact of chatbots on customer loyalty: A systematic literature review. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(1), 11. <https://doi.org/10.3390/jtaer17010011>
- Johari, N. M., & Nohuddin, P. (2021). Quality attributes for a good chatbot: A literature review. *International Journal of Electrical Engineering and Technology (IJEET)*, 12(7), 109–119.
- Kedi, W. E., Ejimuda, C., Idemudia, C., & Ijomah, T. I. (2024). AI chatbot integration in SME marketing platforms: Improving customer interaction and service efficiency. *International Journal of Management & Entrepreneurship Research*, 6, 2332–2341. <https://doi.org/10.51594/ijmer.v6i7.1327>

-
- Ranieri, A., Di Bernardo, I., & Mele, C. (2024). Serving customers through chatbots: positive and negative effects on customer experience. *Journal of Service Theory and Practice*, 34(2), 191–215. <https://doi.org/10.1108/jstp-01-2023-0015>.
- Selamat, M. A., & Windasari, N. A. (2021). Chatbot for SMEs: Integrating customer and business owner perspectives. *Technology in Society*, 66, 101685. <https://doi.org/10.1016/j.techsoc.2021.101685>.
- Skjuve, M., Følstad, A., & Brandtzaeg, P. B. (2023). The user experience of ChatGPT: Findings from a questionnaire study of early users. Proceedings of the 5th International Conference on Conversational User Interfaces (CUI '23), Eindhoven, Netherlands. Association for Computing Machinery. <https://doi.org/10.1145/3571884.3597144>.
- Tran, A. D., Pallant, J. I., & Johnson, L. W. (2021). Exploring the impact of chatbots on consumer sentiment and expectations in retail. *Journal of Retailing and Consumer Services*, 63, 102718. <https://doi.org/10.1016/j.jretconser.2021.102718>.

ALIGNING CHATBOT FEATURES WITH INDUSTRY NEEDS: A COMPARATIVE STUDY OF USER EXPECTATIONS IN UTILITARIAN AND HEDONIC SECTORS FOR SMEs

DENYS BESPALKO,¹ JASON VAN HAMOND,¹

MARGOT NEGGERS,¹ ALINDA KOKKINOU²

¹ Breda University of Applied Sciences, Data Science and AI, North-Brabant,
the Netherlands
231452@buas.nl, 232567@buas.nl, neggers.m@buas.nl

² Breda University of Applied Sciences, ABEL Research Team Zhan, North-Brabant,
the Netherlands
kokkinou.a@buas.nl

This study explores the optimal chatbot orientation—socially oriented (SO) or task-oriented (TO)—for small and medium-sized enterprises (SMEs) in hedonic versus utilitarian industries. Through a mixed-methods approach combining quantitative data from 166 participants and qualitative interviews, it examines user preferences for chatbot interaction styles, focusing on perceived ease of use (PEOU) and usage intention (UI). Findings highlight a strong preference for SO chatbots in hedonic contexts, emphasizing engaging, human-like interactions, while utilitarian industries showed no significant preference, with efficiency taking precedence. Correlation analysis revealed a robust link between perceived ease of use and usage intention for SO chatbots, underscoring their potential to enhance user engagement. The research provides actionable insights for SMEs to tailor chatbot designs to industry-specific customer expectations, aligning digital tools with business goals and customer satisfaction.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.7](https://doi.org/10.18690/um.fov.4.2025.7)

ISBN

978-961-286-998-4

Keywords:

chatbot interaction design,
SME digital strategies,
hedonic vs utilitarian
industries,
user-centric AI tools,
social vs task orientation



University of Maribor Press

1 Introduction

Businesses face a critical imperative: adapt to digital integration or risk obsolescence. With global retail e-commerce sales projected to rise from 6.3 billion in 2024 to 8 billion by 2027 (Statista, 2024), small and medium-sized enterprises (SMEs) must adopt AI tools like chatbots to meet customer demands for seamless, personalized, engaging and 24/7 accessible solutions (Adamopoulou & Moussiades, 2020). However, SMEs face barriers such as high costs and technical complexity, compounded by industry-specific expectations (Hansen & Bøgh, 2020; Alford & Page, 2015). As described in Table 1, in utilitarian sectors (e.g., banking), chatbot users prioritize efficiency and task-oriented functionality, whereas, in hedonic industries (e.g., hospitality), they demand personalized, engaging interactions (Haugeland et al., 2022). This dichotomy underscores that chatbots must be tailored to sector-specific needs—a critical consideration for resource-limited SMEs (Bedué, 2020).

While existing research has explored chatbot attributes, such as hedonic (entertainment-focused) and utilitarian (task-driven) qualities, it has largely overlooked their alignment with industry-specific contexts, particularly for resource-constrained SMEs (Haugeland et al., 2022). This gap is significant, as lack of the context-aware chatbot risks misaligned investments, leading to suboptimal user experiences and diminishes returns of AI adoption. Without systematic evidence linking chatbot orientation to industry-specific user expectations, such companies risk deploying generic solutions that fail to address distinct customer needs in hedonic (e.g., hospitality) versus utilitarian (e.g., banking) sectors.

This study addresses this critical oversight by investigating: *How do social-oriented (SO) and task-oriented (TO) chatbot designs align with user preferences in hedonic versus utilitarian industries for SMEs?* At the same time contributing to the emerging field of context-aware conversational agents (Følstad & Brandtzæg, 2017). By providing empirical insights into industry-tailored chatbot design, this research advances theoretical understanding of user-centric AI tools and offers practical guidance for SMEs or digital solutions companies to optimize customer engagement and competitiveness.

Table 1: Definition and interpretation of chatbot/scenario features

	Features identifies	Explanations	Reference
TO/SO	Orientation: Task- or Social-oriented	Task-Oriented (TO): Focuses on efficiency and goal completion; concise, direct responses to complete tasks quickly, minimizing unnecessary interactions; value speed, functionality, and efficiency in the interaction. Social-Oriented (SO): Prioritizes personalization and engagement; friendly, conversational responses, including small talk or emotional cues; more human-like, personalized experience, even if it sacrifices task efficiency.	(Sheth, 1976)
PEOU	Perceived Ease of Use	Degree to which an individual believes that using a system will be free of effort.	(Davis et al., 1986)
UI	Usage Intention	Extent to which an individual intends to engage with a system, in the future. Closely associated with system adoption metrics.	(Davis et al., 1986)
IND	Industry: Hedonic or Utilitarian	Hedonic Industry: These industries, such as hospitality, travel, and entertainment, focus on creating enjoyable, immersive and personalized user experiences. Utilitarian Industry: Industries like banking, logistics, and technical support prioritize efficiency, accuracy, and the completion of tasks.	(Babin et al., 1994)

Source: Own

2 Literature research

2.1 Effectiveness of Conversational Agents

The increasing integration of AI-driven chatbots into business operations is reshaping the way companies interact with their customers. Recent studies have shown that these agents are evolving to match human performance in certain areas of customer engagement, particularly in structured, task-oriented contexts. For instance, as Luo et al. (2019) report “*undisclosed chatbots (whose nature is not disclosed during the chat, so users assume they are human) are as effective as proficient workers and four times more effective than inexperienced workers in engendering customer purchases*” (p. 938). In a field experiment involving over 6,200 customers, these chatbots performed on par with experienced salespeople and significantly outperformed less experienced workers. This finding is extremely practical for such businesses, as it presents them with a

great opportunity to streamline their budget into a reliable, cost-efficient and powerful system with measurable and proven performance.

However, the research also revealed a key caveat: when customers were made aware that they were interacting with a chatbot, purchase rates plummeted by 79.7%. This stark contrast between perceived and actual competence underscores the scepticism many consumers have towards AI-driven systems, despite their objective performance. The study offers insights into mitigating this bias, such as delaying the disclosure of chatbot identity and targeting customers with prior positive experiences with AI. These findings suggest that while chatbots have demonstrated their capabilities, businesses must strategically manage customer perceptions to fully leverage this technology.

2.2 User Motivation and Chatbot Effectiveness

Following on the previous section, it is extremely important to understand the origin of customer motivations that significantly influence how users interact with AI conversational agents. It was established that participants' interaction rates with AI models were positively correlated with purchasing functional (utilitarian) products (Ruan & Mezei, 2022). Conversely, when the product was perceived as experiential (hedonic), participants' satisfaction was lower if they received a consultation from a chatbot rather than from a human consultant. This finding indicates that there are differences in expectations toward chatbots based on users' motivations and the nature of the product or service.

While this study explored the relationship between the kind of service agent (chatbot or human) and user motivation, it did not delve into how specific chatbot features might impact user perceptions across different industry contexts. To address this gap, our research focuses on the type of industry in which service agents are used.

2.3 Aligning Features of Conversational Agents with User Expectations

This research builds upon established frameworks for understanding user attitudes and behavioral intentions, focusing specifically on the role of chatbot orientation in influencing PEOU and UI. Two significant studies provide the theoretical foundation for this approach. Firstly, Chuen et al. (2023) investigated how perceived

usefulness and perceived ease of use (described in Table 1) impact usage intention in the context of a digital assessment system. The research ended up concluding that while PEOU may not significantly affect usage attitude directly, it remains a critical factor in shaping user behavior and intentions. Secondly, the widely adopted Technology Acceptance Model (TAM) developed by Davis et al. (1986) offers a robust theoretical framework linking system characteristics to user acceptance, emphasizing the role of perceived ease of use as a determinant of usage intention.

2.4 Implications for SMEs and Chatbot Design

Focusing specifically on small and medium-sized enterprises (SMEs), factors that lead to higher usage intentions of conversational agents were investigated within these organizations (Selamat & Windasari, 2021). Given the limited research on the adoption of such technologies by smaller businesses, this study fills a crucial gap by identifying elements that align with the unique needs of these companies. Notably, the study introduces the concept of "relationship marketing," which emphasizes the importance of social presence and feeling of personalization in fostering effective customer interactions (Yi, 2018).

This approach is particularly relevant for small enterprises, as they often rely on building strong customer relationships to compete with larger firms. The emphasis on relationship marketing suggests that these businesses can significantly benefit from implementing socially oriented conversational models that enhance personalization and human-like interactions.

3 Methodology

To investigate the research questions, this study examines three primary variables: Chatbot Orientation—whether the chatbot is Socially Oriented (SO) or Task Oriented (TO); Perceived Ease of Use (PEOU); and Usage Intention (UI). The Chatbot Orientation serves as the independent variable, while PEOU and UI are the dependent variables, with constructs defined in Table 1. The research aims to examine the relationship shown in Figure 1.

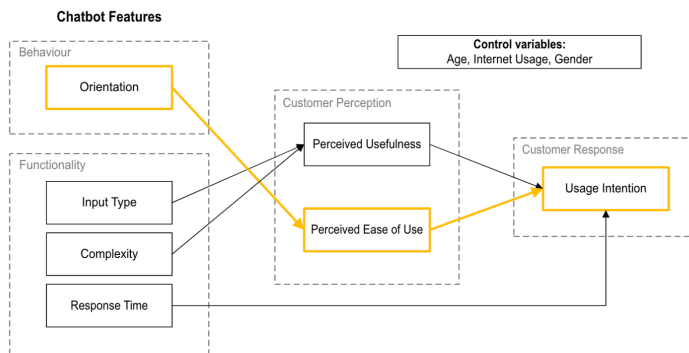


Figure 1: General research model (highlighted area is the scope covered by this research paper)

Source: Own

H₁: SO chatbots are perceived as significantly easier to use than TO ones across both industries, suggested by a difference in means of PEOU.

Socially oriented (SO) chatbots, designed to mimic human-like interactions (Selamat & Windasari, 2021), likely align naturally with users' intuitive expectations of communication.

H₂: Participants are significantly more likely to choose SO chatbots for hedonic industrial contexts, while TO agents are preferred more in utilitarian scenarios, reflecting an industry-dependent preference.

Ruan & Mezei (2022) found users prefer human consultants for hedonic products due to social interaction, suggesting SO chatbots better meet these needs. TO chatbots' efficiency suits utilitarian contexts, where task completion is prioritized (Haugeland et al., 2022).

H₃: A positive correlation can be established between PEOU and UI for SO chatbots among all industry types (Yi, 2018; Selamat & Windasari, 2021).

Selamat & Windasari (2021) and Yi (2018) emphasize relationship marketing in SMEs, where SO chatbots’ ease of use in social interactions likely strengthens UI, as social temperament is critical for customer retention.

3.1 Participants and Sampling

Table 2: Sample demographics information

Characteristic	Distribution
Sample size	168 (through Prolific: 151, convenience sample: 17) 156 (after cleaning and eliminating inattentive responders, explained in section 3.2)
Age	Range: 19-67 years Mean: 29.61 Median: 26 Std: 9.88
Sex	male: 43% female: 57%
Chatbot usage	Every day 22.2% At least 10 times a month 26.7% At least once a month 35% Less than once a month 16.1%

Source: Own

3.2 Research design

The research utilized a mixed-methods approach, combining quantitative questionnaires with qualitative interviews to enrich the data collection. The within-subjects design, involving scenario-based experiments, was chosen to allow participants to experience both chatbot orientations, enhancing the comparison reliability, minimizing individual variability and increasing data robustness (Charness et al., 2012).

Four scenarios were crafted, representing two hedonic industries (karting, board game store) and two utilitarian industries (IT solutions, car garage). To ensure ecological validity, scenarios were based on real-world SME customer interactions and direct input from sector-specific business owners (e.g., hospitality managers, IT support staff). Dialogues reflected authentic user queries—such as reserving a karting slot or resolving software issues—with hedonic scenarios using engaging, conversational tones and utilitarian scenarios prioritizing clarity and efficiency.

Participants accessed the survey via Qualtrics through Prolific. After informed consent, each was randomly assigned two scenarios (one hedonic, one utilitarian), each presenting SO and TO chatbot interactions side-by-side. Participants selected their preferred interaction style and rated Perceived Ease of Use (PEOU) and Usage Intention (UI) on a 5-point Likert scale (1 = “Strongly Disagree,” 5 = “Strongly Agree”):

Perceived Ease of Use (PEOU):

- **PEOU1:** "I found the chatbot easy to use during this interaction."
- **PEOU2:** "The chatbot was clear in guiding me through the process."
- **PEOU3:** "It was simple to complete the required task with the chatbot."
- **PEOU4:** "It was rather difficult to understand the solution from the chatbot."

PEOU4 was intentionally phrased negatively to assess participant attention and response consistency. This reverse-coded item helps identify inattentive respondents and improves the reliability of the data.

Usage Intention (UI):

- **UI1:** "I would use this chatbot again if I needed similar assistance."
- **UI2:** "I would prefer using this chatbot over other available options."
- **UI3:** "I am likely to recommend this chatbot for similar tasks."

This process was repeated for both assigned scenarios, resulting in each participant providing data on two industries and their corresponding chatbot preferences. Randomization of scenarios and chatbot presentation order was implemented to minimize order effects and biases.

3.3 Interview component

To complement the survey data, qualitative interviews were conducted with 5 participants with ages ranging from 19 to 71, providing a deeper understanding of user expectations and preferences across different industry contexts in addition to the survey. The interviews presented participants with chatbot interaction examples

for both hedonic and utilitarian industries. Participants were asked to reflect on these examples, discussing how tone, style, and the use of emojis or casual language influenced their perception of the chatbot's appropriateness and effectiveness within each industry type.

4 Results

4.1 Measurement Model Assessment

Table 3: Assessing the measurement model for reliability and convergent validity of latent variables to ensure robustness of the findings

Variable	Cronbach's alpha	Composite Reliability	Difference	AVE
PEOU	0.804	0.815	0.011	0.635
UI	0.872	0.889	0.027	0.775

Source: Own

The reliability analysis showed strong internal consistency for both Perceived Ease of Use (PEOU) and Usage Intention (UI), with Cronbach's alpha (PEOU: 0.804; UI: 0.872) and Composite Reliability (PEOU: 0.815; UI: 0.889) exceeding the 0.70 threshold. Convergent validity was confirmed, as all Average Variance Extracted (AVE) values surpassed 0.5 (PEOU: 0.635; UI: 0.775). Discriminant validity was established using the Fornell-Larcker criterion, with the square root of AVE for both constructs (PEOU: 0.797; UI: 0.880) exceeding their correlation (0.450), confirming their distinctiveness. These results mean that **these constructs, while related, are conceptually distinct**, which validates the reliability and validity of the measurement model (Fornell & Larcker., 1981).

4.2 Hypotheses testing

Paired t-tests assessed H_1 , which proposes that social-oriented (SO) chatbots are perceived as easier to use than task-oriented (TO) chatbots across industries. In the general context, SO chatbots ($M = 4.437$, $SD = 0.62$) outperformed TO chatbots ($M = 4.307$, $SD = 0.65$), with a mean difference of 0.130 ($t(309) = 1.8529$, $p = 0.065$, $d = 0.21$). This small effect size suggests a slight trend favouring SO chatbots' ease of use, though not significant at $p < 0.05$. Similarly, small effects were observed in hedonic (mean difference = 0.131, $t(154) = 1.3894$, $p = 0.167$, $d = 0.22$) and

utilitarian contexts (mean difference = 0.103, $t(153) = 0.9803$, $p = 0.328$, $d = 0.16$), with no significant differences. These modest effects, potentially influenced by a ceiling effect in PEOU ratings (see Figure 3), indicate limited differentiation in perceived ease of use between chatbot types.

Table 4: Studying H_1 by assessing the significance of differences in ease of use among industries and chatbot types

Construct + IND	Chatbot Type						Result
	SO	TO	Mean diff	T-score	P-value	df (Cohen's d)	
PEOU (General)	4.437	4.307	0.130	1.853	0.065	309 (0.22)	Not supported
PEOU (Hedonic)	4.492	4.361	0.131	1.389	0.167	154 (0.23)	Not supported
PEOU (Utilitarian)	4.368	4.265	0.103	0.980	0.328	153 (0.16)	Not supported

source: own

Table 5: Testing H_2 by comparing the choices of the participants for the preferred interaction style, given the industry scenario

	Orientation	SO	TO	Total
Industry				
Hedonic		93	63	156
Utilitarian		74	81	155
Total		167	144	311

Source: Own

The chi-square analysis ($\chi^2 = 3.94$, $p = 0.047$) reveals a statistically significant association between chatbot orientation (SO vs. TO) and industry context (hedonic vs. utilitarian). This result, significant at the 0.05 level, indicates that the distribution of preferences for SO and TO chatbots varies depending on the context, suggesting that user orientation choices are context sensitive.

To further investigate these preferences, z-tests were conducted to compare the proportions of SO choices within each context. For hedonic industries, the z-test yielded a statistic of 2.45 ($p < 0.014$), indicating a significant preference for SO

chatbots. This suggests that users in hedonic contexts are more inclined to choose chatbots emphasizing social interaction.

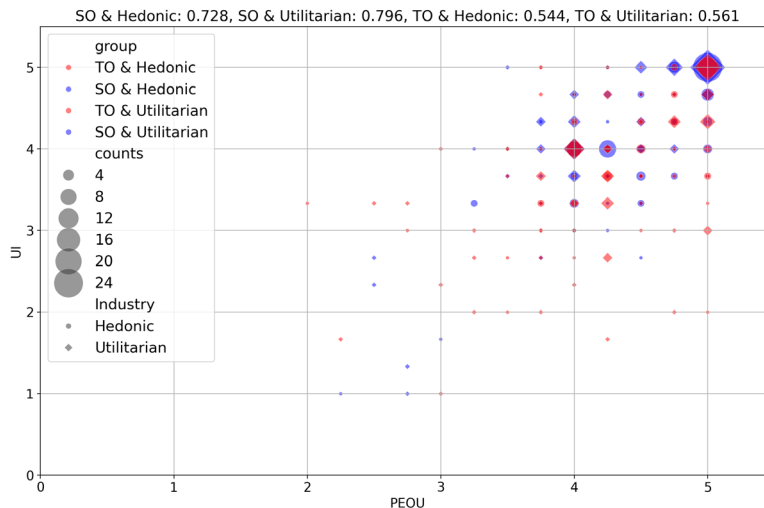


Figure 2: H₃ Correlation combinations of industry and agent types

Source: Own

Conversely, the z-test for utilitarian contexts yielded a non-significant statistic ($z = -0.57$, $p = 0.57$), suggesting no strong preference for either SO or TO chatbots within utilitarian settings, with no apparent inclination toward one interaction style over the other. Overall, these findings highlight a clear preference for socially oriented chatbots in hedonic contexts, while preferences in utilitarian contexts remain neutral. For SO chatbots, the correlations are particularly robust, with values of 0.728 (hedonic) and 0.796 (utilitarian), indicating that users who find SO chatbots easy to use are significantly more likely to intend to use them again. This supports H₃, suggesting that SO chatbots' design enhances user adoption intentions. For TO chatbots, the correlations are weaker but still positive (0.544 for hedonic and 0.561 for utilitarian), implying that while efficiency is valued, it may not drive usage intention as strongly as the social and relational aspects of SO chatbots. These findings underscore the importance of aligning chatbot design with user expectations, particularly in hedonic contexts where SO chatbots excel, while utilitarian contexts show more balanced preferences.

However, the ceiling effect (see Figure 2) likely tempered these correlations, especially at higher levels of perceived ease of use, where further differentiation becomes challenging.

5 Discussion, Limitations and Future Work

Based on the hypotheses tested and the literature referred to, the findings highlight that technology adoption depends on aligning the interaction style of CAs with industry expectations. Firstly, the clear preference for socially oriented (SO) chatbots in hedonic industries suggests that human-like interaction enhances user experience, with interviewees emphasizing the importance of friendly language and emojis in hospitality and entertainment (Ruan & Mezei, 2022). In contrast, utilitarian industries showed no strong preference, indicating that efficiency matters more than social presence (Haugeland et al., 2022). Similarly, interview results stressed professionalism, suggesting that overly casual chatbots may harm credibility in technical or service-driven fields.

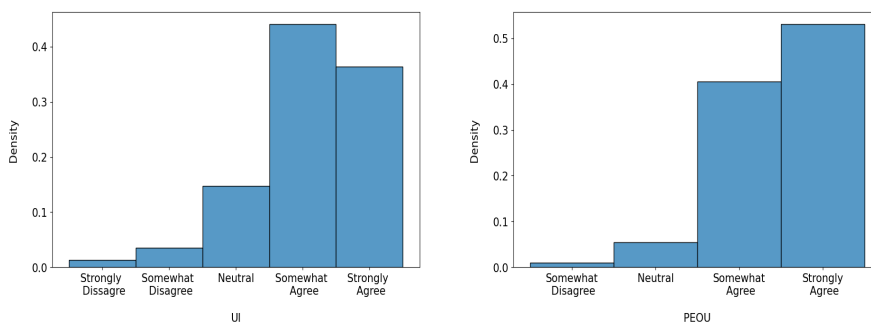


Figure 3: Density Distribution of UI and PEOU

Source: Own

These results extend the Technology Acceptance Model (TAM) by demonstrating that perceived ease of use (PEOU) and usage intention (UI) are moderated by industry context, contributing to the field of context-aware conversational agents (Følstad & Brandtzæg, 2017). Specifically, the preference for SO chatbots in hedonic settings underscores the role of contextual factors in shaping user-centric AI design, offering a theoretical foundation for adaptive conversational systems.

For SMEs, these findings provide actionable guidance: SO chatbots enhance engagement in hedonic industries, while utilitarian contexts benefit from TO or neutral designs. Strategic implementation, including careful management of chatbot disclosure, is critical to optimizing user acceptance and business outcomes.

During the study, it became evident that the responses for the constructs Perceived Ease of Use (PEOU), and Usage Intention (UI) exhibited a ceiling effect. This clustering of responses (shown in Figure 3) complicates the analysis, making it challenging to derive nuanced insights regarding user perceptions and preferences. To add regarding the limitations of this study: a) study sample from Prolific may be biased towards more tech-aware demographic group; b) cultural interpretation of “friendliness” or emojis could differ among populations; c) reliance on hypothetical scenarios with chatbot may not fully represent the real-world interactions. Future research could address these by diversifying samples, examining cultural factors, testing live chatbot interactions, and expanding industry contexts.

In the qualitative analysis, participants revealed distinct preferences based on industry context. In hedonic scenarios, a socially oriented chatbot—characterized by the strategic use of emojis, casual language, and emotional expressions—was consistently perceived as engaging and more likely to encourage future use. In contrast, for utilitarian contexts, respondents favoured a more professional, task-oriented tone, noting that excessive social cues, particularly emojis, could seem unprofessional and off-putting in serious situations. Although one respondent mentioned that removing emojis might make a social approach acceptable, the consensus emphasized that a balanced, restrained communication style is essential to maintain credibility in utilitarian interactions.

6 Conclusion

In conclusion, this research provides targeted insights into how chatbot orientation—social versus task-oriented—aligns with industry-specific user expectations in hedonic and utilitarian contexts for small and medium-sized enterprises (SMEs). The study’s findings reveal a distinct preference for socially oriented chatbots within hedonic industries enhancing engagement and personalized interactions. In contrast, utilitarian industries prioritize functionality over interaction style. Additionally, a notable correlation between perceived ease of use and usage

intention highlights the influence of chatbot design on user acceptance, with stronger correlations observed for socially oriented chatbots across both industry types. This research contributes to the field by filling the knowledge gap concerning industry-specific chatbot preferences and offering actionable recommendations for SMEs aiming to adopt chatbots in alignment with their customer engagement goals. The combined quantitative and qualitative data reinforce these conclusions, underscoring the importance of tailoring chatbot orientation to the unique needs of each industry.

References

- Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 373-383). IFIP Advances in Information and Communication Technology.
- Alford, P., & Page, S. J. (2015). Marketing technology for adoption by small business. *The Service Industries Journal*, 35(11-12: Exploring social media impact within service contexts), 655-669.
- Babin, B., Darden, W., & Griffin, M. (1994). Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value. *Journal of Consumer Research*, 644-655.
- Bedué, P. (2020). Just Fun and Games? Utilitarian and Hedonic Chatbot Perceptions and Their Role for Continuance Intentions. *Business Information Systems, 23rd International Conference* (pp. 291–306). Colorado Springs: Springer.
- Chueh, H.-E., & Huang, D.-H. (2023, February). Usage intention model of digital assessment systems. *Journal of Business Research*, 156.
- Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems : theory and results*. Massachusetts : Massachusetts Institute of Technology.
- Følstad, A., & Brandtzaeg, P. (2017, June 23). Chatbots and the new world of HCI. *Interactions*, 24(4), 38-42.
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 39-50.
- Global retail e-commerce sales 2014-2027 | Statista*. (2024, June). Retrieved from statista: <https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/>
- Hansen, E. B., & Bøgh, S. (2020, 8 12). Artificial intelligence and internet of things in small and medium-sized enterprises: A survey. *Journal of Manufacturing Systems*, 58, 362-372.
- Haugeland, I. K., Følstad, A., Taylor, C., & Bjørkli, C. A. (2022, May). Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction design. *International Journal of Human-Computer Studies*, 161.
- Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019, March). Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*, 38(6), 937-947.
- Ruan, Y., & Mezei, J. (2022). When do AI chatbots lead to higher customer satisfaction than human frontline employees in online shopping assistance? Considering product attribute type. *Journal of Retailing and Consumer Services*.
- Selamat, M. A., & Windasari, N. A. (2021). Chatbot for SMEs: Integrating customer and business owner perspectives. *Technology in Society*.
- Yi, Z. (2018). Chapter One - Introduction to Marketing. *Marketing Services and Resources in Information Organizations*, 1-17.

USING VIDEO SEGMENTATION FOR UNSUPERVISED INDUSTRIAL OPTICAL PRODUCTION CONTROL – CASE STUDY FOR THE SMARTFACTORY@OST

VALMIR BEKIRI, STEFAN STÖCKLER

OST – Eastern Switzerland University of Applied Sciences Institute for Information
and Process Management, Rosenbergstrasse, St. Gallen, Switzerland
valmir.bekiri@ost.ch, stefan.stoeckler@ost.ch

This paper explores the application of video segmentation techniques for unsupervised industrial optical production control. It discusses the limitations of traditional manual inspection methods and the growing need for automated, efficient, and accurate quality control in modern manufacturing. The study focuses on unsupervised methods, particularly Meta's Segment Anything Model 2 (SAM-2), for their adaptability to new product types without extensive labeled training data. One use case from the SmartFactory@OST is presented in detail: video segmentation for monitoring a cobot handling floorballs. The implementation demonstrates the potential of these techniques to enhance efficiency, reduce costs, and improve quality assurance in industrial settings. The paper concludes by highlighting the advantages of unsupervised segmentation methods and suggesting areas for future research and improvement.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.8](https://doi.org/10.18690/um.fov.4.2025.8)

ISBN
978-961-286-998-4

Keywords:
computer vision,
smart factory,
video segmentation,
industry 4.0,
AI



University of Maribor Press

1 Introduction

Quality control is a critical process in industrial manufacturing to ensure that products meet required specifications and standards. Traditionally, quality inspection has relied heavily on manual visual inspection by human operators, which can be time-consuming, labor-intensive, and prone to errors due to fatigue or inconsistency. With the increasing demands for efficiency, accuracy, and automation in modern manufacturing, there is a growing need for more advanced and reliable quality control methods. (Raisul Islam et al., 2024; Vergara-Villegas et al., 2014; Villalba-Diez et al., 2019)

Newly developed quality assurance methods often rely on supervised machine learning techniques, which require extensive labeled datasets and are time-consuming to implement. This has led to a growing interest in unsupervised methods, particularly image and video segmentation, for industrial production quality control. (Ettalibi et al., 2024; Villalba-Diez et al., 2019)

These technologies enable machines to divide visual data into distinct segments, which correspond to different parts or features of an object, allowing for precise identification of defects and irregularities. (Ettalibi et al., 2024; Villalba-Diez et al., 2019)

This paper explores the application of video segmentation techniques for unsupervised industrial production control, examining recent advances, key challenges, and promising future directions in this important area.

1.1 SmartFactory@OST as a production facility

The SmartFactory@OST realizes a manufacturing company that exemplifies a comprehensive learning environment specifically designed to bridge theoretical education with practical implementation of Industry 4.0 concepts across multiple disciplines. Within this innovative framework, the floorball manufacturing process serves as a particularly effective pedagogical model for demonstrating advanced manufacturing concepts and digital integration techniques, see Figure 1. (Stefan Stöckler et al., 2021)



Figure 1: Floorball production cell

Source: <https://www.ost.ch/smartfactory>

The floorball production system represents a sophisticated implementation of configurable product manufacturing that addresses multiple aspects of modern industrial processes. Each ball consists of two independently colored hemispheres, with nine distinct color options available per hemisphere, yielding 45 possible color combinations. This configuration presents an ideal use case for teaching critical concepts such as product variant management, manufacturing execution system (MES) integration, and real-time production planning. The process flow incorporates a collaborative robot (cobot) that handles the ball halves and transfers them to a specialized welding machine, where the hemispheres are melted together under precisely controlled conditions, see Figure 1. (Stefan Stöckler et al., 2021)

The system architecture incorporates multiple data acquisition points, real-time monitoring capabilities, and comprehensive process integration with the central SAP S/4HANA ERP system. Students engage with complex implementation challenges including material master data configuration, production and sales bill of materials differentiation, variant configuration methodology, and integration of order management with manufacturing execution. This interconnected approach facilitates a comprehensive understanding of vertical and horizontal system

integration within manufacturing environments, allowing students to gain hands-on experience with practical industry applications of digital transformation concepts. The floorball production cell thus serves as a microcosm of larger industrial environments, enabling students to develop transferable skills applicable to full-scale manufacturing operations while remaining accessible for educational purposes. (Stefan Stöckler et al., 2021)

Although it is a laboratory environment, the production cell including the cobot of the SmartFactory@OST offers ideal conditions as a case study for optical production control through computer vision.

1.2 Image and Video Segmentation

Computer vision and image processing techniques have emerged as powerful tools for automated visual inspection in industrial settings. In particular image and video segmentation approaches offer significant potential for unsupervised quality control by automatically partitioning images or video frames into meaningful segments that can be analyzed to detect defects or anomalies. Unlike supervised methods that require extensive labeled training data, unsupervised segmentation techniques can adapt to new product types and defect patterns with minimal manual configuration like reference points. (Raisul Islam et al., 2024; Vergara-Villegas et al., 2014)

Image segmentation involves dividing an image into multiple segments or objects, typically to identify boundaries and objects within images. Common segmentation approaches include thresholding, edge detection, region-based methods, and clustering algorithms. Recent advances in artificial intelligence have also enabled more sophisticated segmentation models that can learn complex visual patterns. (Villalba-Diez et al., 2019; Zhou et al., 2021)

In addition, video-based segmentation, which captures temporal information, is becoming increasingly important in real-time quality control applications. Video segmentation allows for continuous monitoring of production processes, enabling the detection of transient defects or handling errors that may not be visible in static images. (Ettalibi et al., 2024; Ravi et al., 2024)

Unsupervised quality control by image segmentation is already well researched in the health care sector see B. Audelan and H. Delingette (Audelan & Delingette, 2019). Villalba-Diez et. al. have also shown the possibilities of computer vision in the printing industry for optical quality control see (Villalba-Diez et al., 2019). Furthermore M. R. Islam et al. is showing advanced methods for quality control in different industries like agriculture, electronics, automotive and more. (Raisul Islam et al., 2024)

Proving that machine learning and computer vision has huge opportunity in quality control.

1.3 Meta's Segment Anything Model 2 (SAM-2)

SAM-2 is an advanced AI model for image and video segmentation, released by Meta AI in 2024. It improves upon the original SAM with enhanced capabilities. It can segment any object in an image without prior training on specific categories, so it works effectively on new, unseen objects, which is crucial for the industrial quality control.

The model accepts various input methods like points and boxes to guide the segmentation and therefore to produce more precise object masks. And most importantly it can technically operate in real-time on consumer hardware. (Ravi et al., 2024)

The Segment Anything Video dataset, which will be used in the prototype implementation, was trained on 50.9K videos. 35.5M masks are trained in the dataset which is 53 times more masks than any other existing video segmentation dataset available in July 2024. One of the most interesting features of SAM-2 is that segmented objects in a video can be occluded by other objects and reappear again and still be tracked by the model. (Ravi et al., 2024)

This makes it an ideal solution for our cobot case study where the robot arm could cover on some occasions the object we are looking for.

2 Design and Implementation

The SmartFactory@OST delivers the perfect base for an proof-of-concept implementation on unsupervised quality control methodologies, with the principal objective of demonstrating the feasibility and efficacy of routine video surveillance integration within industrial manufacturing environments. A specific production case involving floorball manufacturing is the experimental testbed for this prototypical implementation, offering both practical complexity and operational relevance. The error case were manually triggered in the labortory environment to test if the proof-of-concept is working properly.

The technical architecture of this implementation framework employs Python as the primary programming language, using Meta's Segment Anything Model 2 (SAM-2) as the core computer vision foundation. This state-of-the-art model facilitates advanced segmentation capabilities without requiring extensive labeled training datasets, a significant advantage in manufacturing environments where product variations are common and comprehensive data labeling would be prohibitively resource intensive. Only two reference points were necessary for the system to detect the floorball and its two parts correctly. The implementation additionally incorporates OpenCV libraries for fundamental image processing operations, visualization rendering, and supplementary computer vision functionalities.

From a hardware perspective, the implementation deliberately utilizes cost-effective equipment configurations, demonstrating that advanced computer vision applications can be deployed without substantial infrastructure investments. A standard high-definition webcam provides sufficient video capture resolution and frame rate capabilities to support the analytical requirements of the system. This hardware minimalism represents an important consideration for practical industrial adoption, as it substantially reduces implementation barriers and enhances return-on-investment projections for manufacturing facilities considering similar quality control automation.

2.1 Production control with a cobot using video segmentation

In this use case the cobot is grabbing two ball halves and putting them into the welding machine. The welding machine fuses them together and after cooling down the cobot is grabbing the finished ball from the welding machine and transporting it to the other end of the production cell to release it into the dispensing tube. This grabbing and transporting process is an unsafe operation for the cobot. The floorball may slip through the gripper see Figure 1, the ball could be malformed due to issues in the welding process and many different issues could occur during the production. That's exactly where the video segmentation with SAM-2 comes in handy to check if the cobot's handling is failure free. The route that the robot is taking from the welding machine to the dispensing tube could vary based on which dispensing tube is set for the respective order. There are three ways on how the ball can be dispensed. Therefore there are also three different paths the cobot could potentially take. And this is where the charm of our approach comes in: The system does not need to be adapted since it can track basically all current and potential new paths automatically as long as it is somewhere in the viewport visible to the camera.

SAM-2 can technically not process a video file or stream as is. It needs to be converted to images based on the frames of the video. This gives us the additional opportunity to make the system more efficient since we do not need to analyze all frames per second. In the current solution every 15th frame is taken from the video. This setting is tested specifically for this application and needs to be configured and tested separately for every use case, based on how fast and far away the objects are moving in between frames and is also dependent on how many frames the used camera is delivering.



Figure 2: Floorball with different colours

Source: Created by the authors

In the next step an input information needs to be passed to SAM-2 to define which part of the video or respectively the frame is going to be segmented and therefore tracked. Since the floorball's can have a different colour (see Figure 2) per ball half, two points are passed referencing each ball half.

In the video segmentation workflow exemplified by the SAM-2 implementation for floorball detection, a sophisticated processing pipeline is initiated upon point-based reference specification. When the two reference points, each corresponding to a distinct ball half, are supplied to the model, SAM-2 executes a complex series of computational operations. The foundation of this process involves the model's attention mechanism analyzing pixel-level relationships surrounding the specified points, thereby establishing a probabilistic mask boundary delineation. (Ravi et al., 2024)

This segmentation approach is particularly significant since it demonstrates robust object tracking persistence even during partial occlusion by the cobot arm, a critical capability for continuous quality monitoring and therefore it enables completely unsupervised quality control, eliminating the resource-intensive requirement for extensive labeled training datasets.

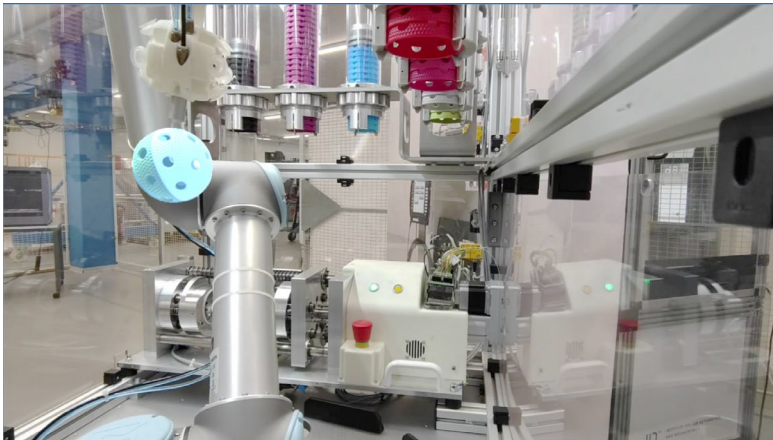


Figure 3: Tracking the floorball parts through the production and visualization through masks

Source: Created by the authors

The resulting segmentation mask, as visualized in Figure 3, provides precise boundary detection of the floorball. With this technique the floorball can be basically tracked within the whole production cell.

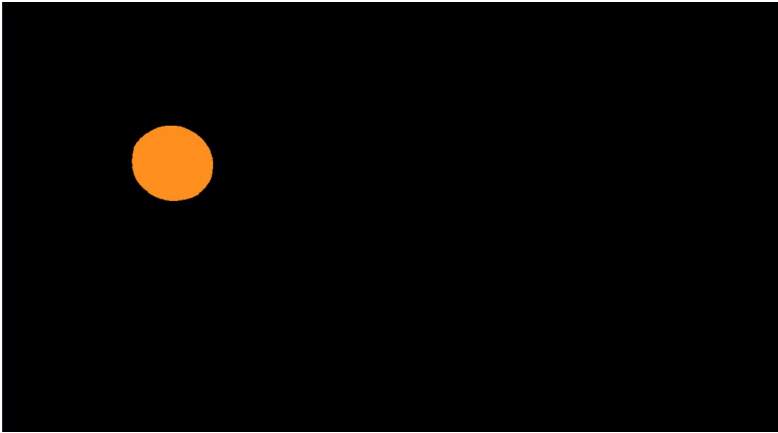


Figure 4: Mask of the detected floorball for additional verification and quality control

Source: Created by the authors

The created mask allows us to perform subsequent comparative analysis against sample images to verify production integrity and cobot handling efficacy throughout the manufacturing process. Using a sample image of a floorball (Figure 2) we can therefore additionally quality control the segmented video and check if the detected part is really a floorball, see Figure 4. By passing this check we can be sure that the floorball is produced correctly, and the production line is working as intended.

As the floorball has a simple shape of an circle it is pretty simple to calculate if the created mask represents the mask of the sample image.

By performing quality assurance during the natural movement of products, you can eliminate the need for dedicated inspection time or additional handling. Quality control becomes seamlessly integrated into the existing production flow, with no separate stations or extra steps required. While the current implementation focuses on basic presence and shape detection for floorballs, this same methodology could extend to more complex products where multiple quality parameters could be verified simultaneously during handling or process steps.

An potential improvement for the current implementation would be to calculate the perimeter, area and aspect ratio of the mask and check if those fit to the expected values. This would further improve the quality assure, since we could technically also know exactly in what distance the floorball should be in each production step and therefore we could not only detect if its an floorball but also if the size of it is genuine and eliminate potential false positives.

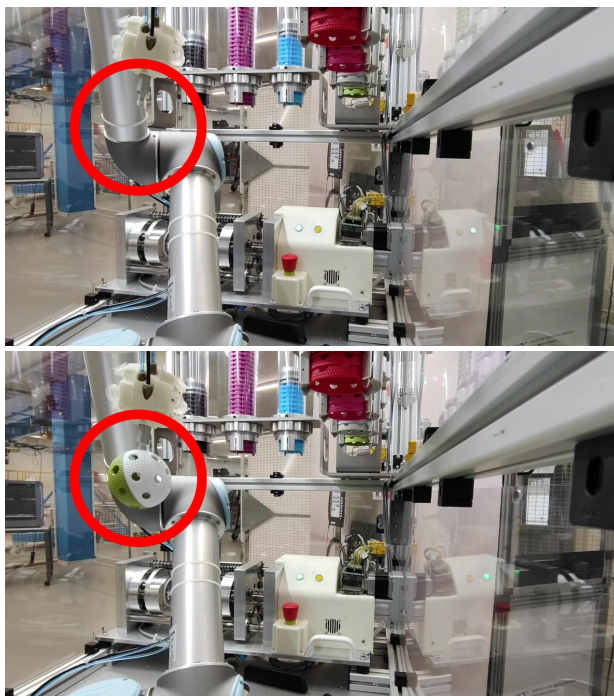


Figure 5: The difference between a wrong and a correct production line

Source: Created by the authors

In Figure 5 two images are shown which describe the difference of an error prone production line and the correct production line. As visible in the above picture there is no floorball in the gripper component. That might happen after the welding process is done and the gripper tries to fetch the now fully cooled down floorball. This state represents a significant production anomaly that traditionally required human supervision for detection or an additional sensor for the specific use case.

Using the implemented proof-of-concept we can now immediately detect that the ball is not in the expected position of the gripper. The implementation therefore demonstrates the efficacy of unsupervised segmentation techniques for industrial quality control applications without requiring extensive labeled training datasets, thereby significantly reducing implementation overhead while maintaining high detection reliability.

3 Conclusion and Outlook

Implementing video segmentation for unsupervised industrial optical process control offers several advantages. Image and video segmentation can adapt to different products and defects without retraining on new labeled data and only using minimal manual input. The automation reduces the time and labor costs associated with manual inspections. Machine-based inspections provide consistent results, minimizing human error without fatigue for 24/7 production lines. Last but not least it allows for immediate detection and correction of defects, reducing waste and downtime. (Ettalibi et al., 2024)

As industries strive for higher efficiency and lower operational costs, the adoption of unsupervised image and video segmentation for quality control is not just advantageous but necessary. It represents a significant step toward fully automated and intelligent manufacturing systems capable of self-monitoring and adaptation. (Ettalibi et al., 2024)

The implementation demonstrates the efficacy of Meta's Segment Anything Model 2 (SAM-2) for unsupervised industrial optical quality control in manufacturing environments. The case study validates several key advantages of this approach within the SmartFactory@OST production facility.

A particularly notable characteristic of the implementation is its path-independence during object tracking. The system successfully maintains continuous object identification regardless of the cobot's movement during transport operations. This intrinsic adaptability eliminates the need for explicit path-specific training, enabling flexible production line reconfiguration without system recalibration.

The model exhibits robust tracking persistence during object manipulation. Product rotation and occlusion by the cobot's gripper or arm presents no significant challenges to the tracking algorithm. Once initial object recognition is established, the system maintains comprehensive product tracking as a cohesive entity despite orientation changes or temporary visual obstruction.

In conclusion, Meta's Segment Anything Model offers a powerful solution for image and video segmentation tasks in industrial quality control. Its ability to operate without extensive supervised training data makes it particularly well-suited for unsupervised applications, where it can enhance efficiency, reduce costs, and improve the overall quality assurance process.

Despite demonstrable efficacy, several technical constraints remain in current implementations. The most significant limitation of the SAM-2 model is the prerequisite for manual specification of initial reference parameters, either points, bounding boxes, or preliminary masks, to effectively segment targeted objects within visual data. This necessitates a priori domain knowledge transfer to the system during configuration phases. A fully autonomous implementation would require supplementary preprocessing algorithms capable of automatic object detection prior to segmentation. So, this knowledge has to be passed first to the system before it can work properly. It would be a big improvement if a processing step could be put in which could autodetect the object.

Therefore, several promising research trajectories emerge from this implementation. The development of automated object detection preprocessing modules would constitute a significant advancement, enabling fully autonomous segmentation without manual reference specification. Additionally, the integration of feedback-driven corrective mechanisms represents an important evolution toward closed-loop quality control systems. For instance, when the system detects a floorball dislocation from the gripper, it could theoretically initiate automated recovery procedures by tracking the object's position and implementing appropriate remediation protocols.

The integration of multimodal imaging technologies, particularly hyperspectral imaging as proposed by De Ketelaere et al. (2022), offers substantial potential for enhancing detection capabilities beyond visible spectrum limitations. Hyperspectral approaches enable material composition analysis and defect identification that

remain imperceptible under standard RGB imaging conditions. (De Ketelaere et al., 2022)

A noteworthy advantage of the proposed methodology is the elimination of specialized sensor arrays traditionally employed in quality control systems. The camera-centric approach enables comprehensive process monitoring which significantly reduces the system complexity and maintenance requirements. This architectural simplification represents an important advancement in industrial monitoring system design, providing both cost efficiency and integration flexibility for existing production environments.

It is also important to mention that the developed use case was build upon an controlled environment within the SmartFactory@OST. The current solution therefore represents an demonstration of an potential professional implementation. Many different egde case would need further tests to gain statistics of the success-rate.

In conclusion, the integration of unsupervised segmentation methodologies, particularly those leveraging foundation models like SAM-2, presents a transformative approach to industrial quality control that aligns with Industry 4.0 paradigms of intelligent, adaptive manufacturing systems. Future work will focus on addressing the identified limitations while expanding the technological capabilities toward increased autonomy.

References

- Audelan, B., & Delingette, H. (2019). *Unsupervised Quality Control of Image Segmentation Based on Bayesian Learning* (S. 21–29). https://doi.org/10.1007/978-3-030-32245-8_3
- De Ketelaere, B., Wouters, N., Kalfas, I., Van Belleghem, R., & Saeyns, W. (2022). A fresh look at computer vision for industrial quality control. *Quality Engineering*, 34(1), 152–158. <https://doi.org/10.1080/08982112.2021.2001828>
- Ettalibi, A., Elouadi, A., & Mansour, A. (2024). AI and Computer Vision-based Real-time Quality Control: A Review of Industrial Applications. *Procedia Computer Science*, 231, 212–220. <https://doi.org/10.1016/j.procs.2023.12.195>
- Raisul Islam, M., Zakir Hossain Zamil, M., Eshmam Rayed, M., Mohsin Kabir, M., Mridha, M. F., Nishimura, S., & Shin, J. (2024). Deep Learning and Computer Vision Techniques for Enhanced Quality Control in Manufacturing Processes. *IEEE Access*, 12, 121449–121479. <https://doi.org/10.1109/ACCESS.2024.3453664>
- Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K. V., Carion, N., Wu, C.-Y., Girshick, R., Dollár,

- P., & Feichtenhofer, C. (2024). *SAM 2: Segment Anything in Images and Videos* (No. arXiv:2408.00714). arXiv. <https://doi.org/10.48550/arXiv.2408.00714>
- Stefan Stöckler, Roman Hänggi, Raphael Bernhardsgrütter, & Christoph Baumgarten. (2021, September 13). (PDF) SAP Academic Community Conference 2021 DACH Industrie 4.0 begreifbar machen -Die SmartFactory@OST. *SAP Academic Community Conference 2021 DACH*. SAP Academic Community Conference 2021 DACH, München. <https://doi.org/10.14459/2021md1622154>
- Vergara-Villegas, O. O., Cruz-Sánchez, V. G., de Jesús Ochoa-Domínguez, H., de Jesús Nandayapa-Alfaro, M., & Flores-Abad, Á. (2014). Automatic Product Quality Inspection Using Computer Vision Systems. In J. L. García-Alcaraz, A. A. Maldonado-Macías, & G. Cortes-Robles (Hrsg.), *Lean Manufacturing in the Developing World* (S. 135–156). Springer International Publishing. https://doi.org/10.1007/978-3-319-04951-9_7
- Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Meré, J., Buchwitz, M., & Wellbrock, W. (2019). Deep Learning for Industrial Computer Vision Quality Control in the Printing Industry 4.0. *Sensors (Basel, Switzerland)*, 19(18). <https://doi.org/10.3390/s19183987>
- Zhou, L., Zhang, L., & Konz, N. (2021). *Computer Vision Techniques in Manufacturing*. <https://doi.org/10.36227/techrxiv.17125652.v1>

ASSESSING THE IMPACT OF CHATBOTS ON CUSTOMER SERVICE EFFICIENCY: A COMPARATIVE STUDY WITH HUMAN AGENTS

FATIMA BILAL, JÓZSEF MEZEI, BAD RE ALAM

Business and Economics, and Law, Åbo Akademi University, Faculty of Social Sciences,
Turku, Finland

fatima.bilal@abo.fi, badre.alam@abo.fi, jozsef.mezei@abo.fi

This research investigates the way AI chatbots influence consumer service efficiency compared to human workers. The proliferation of chatbots across businesses emphasizes the importance of evaluating their impact on productivity, customer satisfaction, and service quality. Although chatbots are effective, they struggle with delicate or complicated interactions. Humans remain essential for providing sympathetic and thorough responses. This study reveals a significant gap in chatbots' ability to meet customer expectations, especially when individualized attention and trust are required. A quantitative research approach was used to collect survey data from customers having experience with both human agents and virtual assistants, i.e., chatbots. According to the findings, while routine inquiries are effectively handled by chatbots, they still struggle with contextual understanding and emotional intelligence. The ideal way to maximize client fulfillment and quality of service is to implement a service structure that combines chatbots for repetitive tasks and human agents for more complicated interactions. The study emphasizes how crucial it is to update chatbot systems frequently to satisfy evolving client demands and boost service effectiveness.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.9](https://doi.org/10.18690/um.fov.4.2025.9)

ISBN
978-961-286-998-4

Keywords:
chatbot,
customer satisfaction,
human agents,
responses,
customer engagement



University of Maribor Press

1 Introduction

In recent years, automated chatbots have become ubiquitous in customer service, handling tasks from answering simple FAQs to managing complex requests. Companies deploy chatbots to improve response times and operational efficiency while aiming to maintain high service quality (Smith, 2021). However, a fundamental question remains: how do chatbot-driven service interactions compare to those handled by human agents in terms of customer satisfaction, service efficiency, and perceived service quality? Existing research has largely examined chatbots in isolation or through limited scenarios, without directly benchmarking them against human agents across multiple performance metrics. De Keyser (2020) suggests that giving chatbots human-like characteristics, such as humor and a more personable conversational style, could make interactions more engaging for customers, but empirical evidence for the impact of these features on customer experience remains limited. It is still unclear whether a humorous, anthropomorphized chatbot improves customer satisfaction or if it might undermine professionalism. Initial studies have begun to explore such design elements, but their effects on key service outcomes have yet to be conclusively demonstrated. This uncertainty has prompted calls for deeper investigation into chatbot personalization, underscoring a second gap in the literature. Moreover, context may significantly influence the success of chatbot deployments. Customer expectations and tolerance for automation vary across industries—what works well for a retail chatbot may not translate to a banking or healthcare setting. For example, banking customers might prioritize accuracy and formality, whereas retail customers could respond more favorably to a friendly, playful tone. Despite these differences, few studies have systematically compared chatbot and human agent performance across different service sectors. As a summary, we have limited insight into whether industry-specific factors moderate the outcomes of chatbot versus human interactions or alter the effectiveness of personalization features. This study contributes to the customer service and AI literature by addressing these interconnected gaps with original empirical evidence. We conduct a comprehensive comparison of chatbot and human agent performance using data collected from three distinct industries: banking, retail, and healthcare. Key performance indicators—including customer satisfaction, service efficiency, and perceived service quality—are measured to evaluate where chatbots can match or even exceed human agents and where they still fall short. In doing so, we extend prior work that examined isolated aspects of chatbot efficacy (Smith, 2021) by

providing a holistic, side-by-side assessment of AI versus human service delivery. Additionally, we incorporate personalized chatbot behaviours, specifically a humorous communication style and other human-like cues, to examine their influence on the customer experience. By testing interactions with and without these personalization features, our study directly responds to calls for further research into chatbot personalization in service encounters (De Keyser, 2020).

1.1 Research Gap

Chatbots are used in customer service extensively to save costs, yet there is limited research on how well these bots deliver a smooth and positive customer experience. There is limited research on what effect that the efficiency gains using chatbots have on overall customer satisfaction and trust (Kaczorowska, 2019). There are significant challenges related to maintaining and establishing the trust of customers during interactions. Customers also like to interact with chatbots while having expectations that they will mimic human-like interactions and understand their emotional contexts (Selamat & Windasari, 2021). Current technology of chatbots is falling short of delivering this kind of interaction and personalized service, which also leads towards dissatisfaction (Nordheim, Følstad & Bjørkli, 2019). Moreover, there is a lack of understanding regarding factors impacting the trust of customers, such as perception of brands, perceived expertise, and responsiveness (Sonntag, Mehmman & Teuteberg, 2023). Furthermore, little study has been done on the use of chatbots amongst medium-sized companies (SMEs) (Gnewuch, 2017). To sum up, these gaps highlight the need for more study on the development, application, and adoption of chatbots within customer support, particularly to improve user experience, trust, as well as the general experience of users.

1.2 Aim of the Research

The main aim of this article is to assess and explore customer satisfaction, trust and perception related to AI-powered chatbots as compared to interaction with human agents throughout the context of customer services. The focus of the study is on how the interaction is experienced by customers with chatbots regarding personalization, effectiveness, and ability to provide empathetic responses. The influence of chatbots is also investigated regarding design features such as response time, anthropomorphism, and contextual understanding on customers' trust level

and their satisfaction (Nordheim, Følstad & Bjørkli, 2019; Sonntag, Mehmman & Teuteberg, 2023).

For these reasons, the following research questions are proposed:

RQ1: To what extent are customers satisfied with the overall service provided by AI-powered chatbots compared to human agents?

RQ2: To what extent do customers trust AI-powered chatbots in handling their inquiries, particularly those requiring personalized and empathetic responses?

RQ3: What factors influence customer preferences for using AI-powered chatbots versus human agents in different stages of the customer service process?

2 Literature Review

AI-powered chatbots have become an essential part of modern customer service operations. They help businesses improve efficiency, cut costs, and enhance customer experiences. Chatbots consist of functions such as consistent responses, routines, and immediate inquiries that meet the expectations of customers for effective and quick services. They are useful in today's globalized and electronically driven businesses since they can work around the clock and manage several client engagements at once (Lee, J. et al, 2019). This understanding is crucial, as it informs the design and functionality of a chatbot, ensuring alignment with customer needs.

2.1 Customer Satisfaction with Chatbots

The evolution of the literature on customer satisfaction with chatbots reflects the increasing integration of these technologies into various sectors, including e-commerce and customer service. The early work by Paula Chaves & Aurelio Gerosa (2019) provides a foundational understanding of human-chatbot interaction design, emphasizing the urgency and applicability of chatbots across diverse domains such as education, health, and business. The importance of technical attributes and social characteristics is highlighted by their surveys while setting the stage for further exploration regarding elements influencing user experience and satisfaction (Cordero, J et al., 2022). For scenarios that require routine services to achieve

customer satisfaction and consistent service, chatbots tend to be more effective. Araujo's (2018) research explores how anthropomorphism affects customer happiness in the chatbot business, especially within the dietary industry. According to their exploratory research, people's opinions of businesses and products are greatly impacted by their interactions with virtual assisting platforms.

It is also argued that customer satisfaction is transcended by the mere functionality of the chatbot and is linked intrinsically to the overall experience of customers online. It is asserted that customer satisfaction is enhanced by chatbots while providing decision support and fostering a culture of loyalty with the use of enhanced communication and mitigating anxiety and uncertainty during the process of purchasing (Chung et al., 2020). The research by Chattaraman, Kwon, and Gilbert (2018) has underscored that customer satisfaction can be enhanced by incorporating elements that are human-like in interaction.

2.2 The Influence of Anthropomorphism on Customer Trust

The exploration of anthropomorphism in marketing and consumer behaviour has garnered increasing attention over the years, particularly in its capacity to influence customer trust. The foundational work by Araujo (2018) outlines the literature surrounding anthropomorphism's effects on consumer perceptions and attitudes, emphasizing its potential in advertising. Their research suggests that the effectiveness of anthropomorphized advertisements hinges on factors such as product involvement and type. This highlights the necessity for marketers to strategically employ anthropomorphism in a manner that aligns with consumer cognitive and emotional responses. A study by Van Pinxteren, Pluymaekers, and Lemmink (2020) has shown that users' trust and satisfaction are significantly increased by employing human-like dialogue strategies such as personalized responses, empathy, and humour. Building on these insights, Paula Chaves & Aurelio Gerosa (2019) further investigate the intricacies of human-chatbot interactions. Their findings underscore the necessity for chatbots to be attuned to individual user quirks and to anticipate needs, which can enhance user engagement.

2.3 Handling Complex Inquiries: A Critical Challenge

There is still a significant challenge regarding the capacity of AI-powered chatbots in terms of handling complicated queries that demand emotional intelligence and contextual understanding. Although there has been an enhancement in chatbots in recent years regarding their ability and processes to respond to context-specific questions and indications, which is still limited (Zion, 2018). The literature on the handling of complex inquiries reveals significant insights into the challenges faced by organizations in managing knowledge and learning processes. Zaman (2016) explores the role of knowledge management systems within enterprise support centres, highlighting their potential to enhance customer service by reducing call times and facilitating problem resolution without the need for in-person visits. The challenges of structuring information either in advance or on the fly further complicate the resolution process, suggesting that while knowledge management systems can provide valuable resources, their effectiveness is contingent upon the ability to navigate the intricacies of information retrieval and application (Dillman, 2014).

2.4 Theoretical foundation:

In addition to user acceptance and trust, the nature of the service task – specifically its complexity or emotional intensity – plays a crucial role in determining whether a chatbot or a human agent is more effective. Here, Media Richness Theory (MRT) provides a useful lens. MRT classifies communication channels by their ability to convey rich information and resolve ambiguity: richer media (those that provide multiple cues, immediate feedback, and personalisation) are better suited for complex, equivocal tasks. Text-based chatbots are a relatively “lean” medium, lacking vocal tone, facial expressions, or deep contextual understanding, whereas human agents (especially via phone or face-to-face) offer richer communication with empathy, adaptive reasoning, and emotional intelligence. According to MRT, we would expect chatbots to perform well for *simple, routine inquiries* that require straightforward question-answer exchanges, but to struggle in *highly complex or emotionally charged scenarios* that demand nuance. The existing customer service literature supports this view: chatbots have demonstrated clear advantages in efficiency for routine tasks – they can instantly answer FAQs, 24/7, and handle high volumes without fatigue. By automating repetitive inquiries, chatbots reduce response times and operational costs, which improves service efficiency for

businesses and often satisfies customers who get quick solutions. However, when an issue is unusual, complex, or requires empathy (for example, a complaint about a sensitive personal matter), purely automated agents often fall short. Users report frustration when chatbots cannot understand context or convey genuine empathy in such cases, leading to lower satisfaction. Lee et al. (2019) and Følstad & Brandtzæg (2020) similarly note that while users find chatbots acceptable for simple tasks, they prefer human agents for complex or emotion-laden issues that need human empathy and flexibility. In essence, human agents remain an important safety net for scenarios where a richer, more adaptive communication style is required. Studies suggest this blended approach can yield higher overall customer satisfaction and service quality than either method alone. Wirtz et al. (2018) predicted that human–AI teamwork would become the optimal service delivery mode for complex services, as customers get immediacy from AI along with the reassurance of human judgment when needed.

2.5 Research Hypotheses

Based on the previous discussion, the following research hypotheses can be formulated:

Hypothesis 1 (H1): Customer satisfaction is enhanced by AI-powered chatbots within routine service interaction by offering more effective and faster responses than human customer care agents.

Hypothesis 2 (H2): Customer trust and engagement are significantly enhanced by incorporating anthropomorphic features (e.g., human-like language, empathy) in AI chatbots.

Hypothesis 3 (H3): Human agents are preferred by Customers over AI-powered chatbots for complex or emotionally sensitive inquiries because of their perceived competence and empathy of human agents.

The subsequent methodology and analysis sections will detail how the data were collected and analyzed to evaluate these hypotheses.

3 Methodology

Quantitative research has been utilized in this paper to analyze customer perceptions and their interaction with human agents versus chatbots. The collection of standardized and measurable data has been facilitated by the structured questionnaire. A quantitative approach has been followed, and the survey data have been analyzed using SPSS software (de Visser, 2018).

3.1 Data Collection

A structured online survey has been conducted to study AI-powered chatbots and how they impact customer satisfaction, service effectiveness, engagement, and trust. The research survey aims to collect diverse experiences of customers with human agents and AI chatbots within different service conditions. The survey has been designed and distributed to ensure diversity, sufficient quality of data, and adherence to established and well-organised research methods. The poll was available for a period of a few weeks to offer enough time to respondents to promote involvement and decrease non-response bias. We sent out several reminders during this time. To increase the completion rate and quality of responses, the survey platform automatically followed up with participants who left incomplete answers. To include respondents from a range of industries, demographics, and customer service scenarios, we chose a variety of distribution strategies (Saunders et al., 2019).

The questionnaire used for this study was designed based on well-established scales from prior research to ensure validity. Items gauging *customer satisfaction* with service interactions were drawn from standardized customer satisfaction surveys in the IT service literature (e.g., adapting wording from Ashfaq et al., 2020, who studied satisfaction with AI service agents), and trust in the chatbot was measured drawing on factors identified by Gursoy, Chi & Lu (2019), and Følstad et al. (2018). New items were developed to address aspects unique to our focus, such as comparing preferences for chatbots vs. humans in complex scenarios, but these were reviewed by an expert panel to ensure they were clear and content valid. We pilot-tested the questionnaire with a small group of users and performed reliability analysis on each multi-item scale. All key constructs (e.g., satisfaction, trust) achieved Cronbach's alpha values above 0.70, indicating good internal consistency. In summary, the study's measures are both theoretically informed and empirically vetted, providing a

solid foundation to test our hypotheses. Each question was structured with a five-point Likert-type scale (ranging from “Strongly Disagree” to “Strongly Agree”), providing a standardised format for respondents to rate their perceptions. In the questionnaire, we collected the following information: general demographic participant information, the use of service platforms, and experience with AI (customer satisfaction, trust, and efficiency of chatbots). A convenience sampling was drawn with individuals who have interacted with customer service chatbots or with human agents recently. Considering also the complexity of our research goals and statistical requirements, the target sample size was chosen to be 160 participants. Power analysis was performed to determine this sample size using a medium effect size of 0.3, a statistical power of 0.8, and a significance level of 0.05 (Cohen, 1988).

3.2 Data Analysis

SPSS software was used for calculating descriptive and inferential statistics. The analysis involved several steps. In descriptive statistics, major key data points such as average satisfaction and trust levels for chatbot and human-agent interactions were. Inferential statistics were employed to compare customer satisfaction ratings between chatbots and human agents in routine and complex service scenarios. Specifically, ANOVA tests were used to assess the differences in customer preference for chatbots vs. human agents handling each of the different types of service inquiries.

4 Results and Analysis

4.1 Demographics analysis of the respondent

Figure 1 demonstrates the age distribution of survey participants. The category of 25-34 showed the highest representation, followed closely by the 18-24 group. It can be claimed that the perspective of younger adults is primarily captured by the study. In general, younger adults (18-34) are more familiar with digital customer service technology. Recognizing this limitation is essential for interpreting the insights of the current study and assessing their relevance to different age groups. Regarding gender distribution, more than 125 respondents are male, while the total sample size is 160. The data indicates that the largest group of respondents uses customer service platforms, "Every day," with over 80 responses.

4.2 Hypothesis 1 Analysis: AI-powered Chatbots vs. Human Agents in Routine Service Interactions

To evaluate H1, we compared customer satisfaction ratings for routine service interactions handled by chatbots versus those handled by human service agents. Because this hypothesis involves comparing mean satisfaction between two service conditions, a t-test was appropriate. We treated the two sets of satisfaction ratings (chatbot-handled vs. human-handled) as related measures, since the survey asked each participant to rate both experiences. The analysis began by reviewing the descriptive statistics (Table 1) for customer satisfaction with AI-powered chatbots and human agents. A sample size of 160 respondents was used to evaluate satisfaction levels for both service types. The mean satisfaction score for AI-powered chatbots was 1.88, while human agents had a higher mean score of 2.32. This suggests that, on average, respondents were more satisfied with the service provided by human agents than with chatbots. A regression analysis (Table 2) explored the relationship between service type (chatbot vs. human agent) and customer satisfaction. The R-value was 0.137, suggesting a weak positive correlation between service type and overall satisfaction. The results of the ANOVA test presented in Table 3 reinforced these findings. With a regression sum of squares of 5.208 and a residual sum of squares of 272.392, the F-value was calculated to be 3.021. The corresponding p-value of 0.084 indicates that the model is not statistically significant at the 0.05 level. We do not reject the null hypothesis, which suggests that AI chatbots do not significantly improve customer satisfaction compared to human agents in routine service interactions. This result shows insufficient statistical evidence to claim that chatbots provide better satisfaction in routine customer service. Finally, a t-test (Table 4) was performed to compare customer satisfaction ratings between chatbots and human agents. The t-test for satisfaction with AI-powered chatbots yielded a one-sided p-value of 0.038, suggesting a marginally significant difference favouring chatbots. Therefore, we **failed to support Hypothesis 1**, instead retaining the null hypothesis that AI-powered chatbots do not significantly enhance routine-service satisfaction compared to human agents. Customers appeared at least as satisfied with human assistance as with chatbot assistance in routine scenarios.

4.3 Hypothesis 2 Analysis: The Impact of Anthropomorphic Features on Customer Trust and Engagement

Hypothesis 2 concerns whether adding human-like traits like empathy and language to AI chatbots boosts customer trust and engagement. The descriptive analysis in Table 5 indicates that adding empathy and human-like traits to AI chatbots does not significantly improve users' trust in the accuracy of the information they provide. This result supports the null hypothesis for Hypothesis 2, suggesting that empathy does not significantly affect the trustworthiness of chatbots in this context. This suggests that other factors beyond empathy might influence users' trust in AI-powered chatbots more. An independent-samples t-test showed a statistically significant difference in trust ratings between the two groups, $t(158) = 4.06, p < .001$. On a 5-point scale, the “empathetic chatbot” group had a higher mean trust score ($M = 3.31, SD = 0.88$, assumed from context) compared to the “non-empathetic chatbot” group ($M = 2.44, SD = 0.93$). Those who experienced the chatbot as empathetic showed a higher intention to keep using chatbot-based service, compared to those who did not perceive empathy in the chatbot. This difference was also significant: $t(158) = 3.01, p = .003$ (two-tailed). Participants in the empathetic-chatbot group reported a greater likelihood of future chatbot use ($M \approx 2.8$ on a 6-point likelihood scale) than the non-empathetic group ($M \approx 1.9$ on the same scale; higher values indicate higher willingness). The magnitude of this effect was medium (Cohen's $d \sim 0.5$). These results indicate that human-like, anthropomorphic design elements in chatbots bolster both user trust and engagement. Thus, **Hypothesis 2 was supported**. We reject the null hypothesis, concluding that adding anthropomorphic features significantly enhances customer trust in the chatbot's information and increases customers' willingness to use the chatbot for future service interactions.

4.4 Hypothesis 3 Analysis: Preference for Human Agents over AI-Powered Chatbots for Complex or Emotionally Sensitive Inquiries

The descriptive statistics point to a prevailing tendency to view human agents as more capable of providing empathy and personalized service, thereby emphasizing the significance of human interaction in customer service. We analyzed several survey items related to this hypothesis: the perceived efficiency of chatbots for routine tasks, the perceived empathy of chatbots, the importance of speaking to a human in service interactions, and whether human agents provide more personalized service

than chatbots. As shown in Table 6, the ANOVA results indicate significant differences in perceptions of personalization, chatbot empathy, and efficiency. The one-way ANOVA is appropriate here to determine if there is significant variance in these ratings that aligns with a general preference for human vs. AI service. Additionally, we examined simple descriptive statistics and frequencies for direct preference questions to see the overall trend. The one-way ANOVA results confirmed that there were statistically significant differences in how participants perceived chatbots and human agents on these key dimensions. For instance, ratings of personalized service differed significantly when comparing responses across individuals (grouped by their level of agreement), $F(4, 155) = 4.48, p = .002$, indicating that the extent to which someone felt humans provide more personalized service than chatbots varied in a non-random way (generally, those who valued personalization highly were the ones preferring humans). Likewise, perceptions of chatbot efficiency showed a significant variance among respondents, $F(4, 155) = 10.90, p < .001$, suggesting some groups of users acknowledge chatbots' efficiency more than others. Notably, chatbot empathy was another area with significant differences in perception, $F(4, 155) = 5.80, p < .001$, implying that virtually all groups rated chatbot empathy low, but with slight variations (overall very low across the board). In contrast, the perceived need for a human option did not differ significantly across groups ($F = 1.29, p = .275$), since almost everyone consistently felt it was important to have the option to escalate to a human agent if needed. Participants attribute higher empathy and personalization to humans, which aligns with the hypothesis. We therefore accept Hypothesis 3: customers showed a clear, statistically supported preference for human customer service agents in complex or emotionally sensitive service situations.

4.5 Discussion

Based on the presented results, Hypothesis 1 aimed to assess whether customers' experiences in routine service interactions were better with AI-powered than human agents. As it was shown, basic inquiries are handled efficiently by chatbots, though slightly better customer satisfaction is provided by human agents. Chandra et al. (2022) have proposed that chatbots require human-like skills if they want to have a relationship with people, such as empathy and cognitive skills. The results of this study suggest, in line with the null hypothesis of H1, that AI chatbots do not noticeably outperform human agents in the realm of routine service interactions regarding user satisfaction. Hypothesis 2 examined the effects of anthropomorphic

traits like empathy on the degree of customer trust and chatbot engagement in the case of an AI chatbot. While Adam et al. (2020) found that using verbal anthropomorphic design in chatbots can increase user compliance and engagement, we did not find that this increases trust directly. Therefore, while there is some feeling of empathy for chatbot recommendations, it may not contribute significantly to chatbot accuracy trust. Hypothesis 3 investigated whether customers prefer human agents over AI-powered chatbots for contacting agents in handling complex or emotionally sensitive calls. The results revealed that customers are willing to interact with human agents as they are more competent, empathetic, and better able to provide personalized service. This is consistent with Beattie et al. (2020), who established that in an emotional context, human interaction is considered more socially attractive, competent, and credible. Overall, AI chatbots demonstrated clear efficiency advantages in handling routine inquiries, often providing instant responses that shorten wait times. This supported our expectation (H1). Participants generally reported positive experiences with chatbot speed and convenience, echoing prior research (Araujo, 2018; Chung et al., 2020). However, the results also confirmed that human agents remain crucial for complex or emotionally sensitive issues. In line with H2, customers achieved higher satisfaction when human support handled nuanced problems, reflecting chatbots' ongoing inability to deliver the empathy and contextual understanding required in such cases. This aligns with findings by Nordheim et al. (2019), who note that purely automated systems often fail to provide the personalized, empathetic service needed to avoid user frustration in complex scenarios. Another key outcome was the role of trust and engagement (H3). The study found that users' trust in chatbot-assisted service was heavily dependent on the chatbot's performance, specifically its accuracy and its ability to handle issues or escalate failures appropriately. Hypothesis 3 was supported: respondents indicated significantly greater trust in chatbots that consistently answered questions correctly and resolved problems efficiently, whereas poor performance quickly eroded confidence. This result is consistent with Gursoy et al. (2019), who emphasize that customers' trust in AI agents hinges on perceived competence and reliability, even if the bot uses anthropomorphic, human-like cues. Our findings reinforce that functional quality is paramount: chatbots must be both **accurate** and **reliable** to maintain user trust. Nevertheless, even the most advanced chatbot in our study did not fully replicate the warmth and empathy of a human agent.

5 Conclusions

This study contributes to the literature and practice by understanding how to improve AI-powered chatbots' effectiveness in customer service. Based on the findings, the chatbots should primarily handle routine tasks while allowing human agents to deal with customers personally, ensuring higher customer satisfaction.

Furthermore, chatbot accuracy and reliability need to be considered to build trust. Ensuring constant updating of the chatbot's knowledge, training it with fundamental customer interactions, and using advanced AI can help make the response reliable and help build customer trust. It is important to note some limitations of this study. The sample was skewed, having very young adults as participants, and an overwhelming leaning toward the male gender. Thus, it restricts the generalizability of the findings. Furthermore, we also recommend incorporating empathy and human-like interaction as they have a favourable effect on user engagement and future usage intentions. The study also restricted itself to certain industries. The effectiveness of a chatbot can differ from industry as customers' expectations and service needs vary depending on the sector. Building on this study, future research should pursue a more diverse and expansive examination of chatbots in customer service. One recommendation is to conduct studies in different organisational contexts, including small and medium-sized enterprises (SMES). Prior literature notes that most chatbot research to date focuses on large organisations, overlooking the unique challenges and resource constraints of SME.

References

- Adam, M., Wessel, M. and Benlian, A. (2020). AI-based Chatbots in Customer Service and Their Effects on User Compliance. *Electronic Markets*, 31(2). doi: <https://doi.org/10.1007/s12525-020-00414/7>.
- Andrade, I.M.D. and Tumelero, C. (2022). Increasing customer service efficiency through an artificial intelligence chatbot. *Revista de Gestão*, 29(3). doi: <https://doi.org/10.1108/rege-07-2021-0120>.
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, 85, pp.183–189. doi: <https://doi.org/10.1016/j.chb.2018.03.051>.
- Behera, R.K., Bala, P.K. and Ray, A. (2021). Cognitive Chatbot for Personalised Contextual Customer Service: Behind the Scene and Beyond the Hype. *Information Systems Frontiers*. doi: <https://doi.org/10.1007/s10796-021-10168-y>.
- Belanche, D., Casaló, L.V., Flavián, C. and Schepers, J. (2020). Service Robot implementation: a Theoretical Framework and Research Agenda. *The Service Industries Journal*, 40(3-4), pp.203–225. doi: <https://doi.org/10.1080/02642069.2019.1672666>.

- Bickmore, T.W., Puskar, K., Schlenk, E.A., Pfeifer, L.M. and Sereika, S.M. (2010). Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers*, 22(4), pp.276–288. doi: <https://doi.org/10.1016/j.intcom.2010.02.001>.
- BRATTIANU, C. (2013). Exploring Knowledge Entropy in Organizations. *Management Dynamics in the Knowledge Economy*, 7(3), pp.353–366. doi: <https://doi.org/10.25019/mdke/7.3.05>.
- Bryman, A. and Bell, E. (2015). *Business Research Methods*. Oxford University Press, Oxford.
- Chattaraman, V., Kwon, W. S., & Gilbert, J. E. (2012). Virtual agents in retail websites: Benefits of simulated social interaction for older users. *Computers in Human Behavior*, 28(6), 2055–2066.
- Chaves, A.P. and Gerosa, M.A. (2020). How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design. *International Journal of Human-Computer Interaction*, 37(8), pp.1–30. doi: <https://doi.org/10.1080/10447318.2020.1841438>.
- Chung, M., Ko, E., Joung, H. and Kim, S.J. (2020). Chatbot e-service and Customer Satisfaction regarding Luxury Brands. *Journal of Business Research*, 117, pp.587–595. doi: <https://doi.org/10.1016/j.jbusres.2018.10.004>.
- Ciechanowski, L., Przegalinska, A., Magnuski, M. and Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92, pp.539–548. doi: <https://doi.org/10.1016/j.future.2018.01.055>.
- Cordero, J., Barba-Guaman, L. and Guaman, F. (2022). Use of chatbots for customer service in MSMEs. *Applied Computing and Informatics*. doi: <https://doi.org/10.1108/aci-06-2022-0148>.
- Creswell, J.W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th ed. Thousand Oaks, California: Sage Publications.
- Davenport, T.H. (2018). Artificial Intelligence for the Real World. *Harvard Business Review*. Available at: <https://hbr.org/webinar/2018/02/artificial-intelligence-for-the-real-world>.
- De Keyser, A., Köcher, S., Alkire (née Nasr), L., Verbeeck, C. and Kandampully, J. (2019). Frontline Service Technology infusion: conceptual archetypes and future research directions. *Journal of Service Management*, 30(1), pp.156–183. doi: <https://doi.org/10.1108/josm-03-2018-0082>.
- De Keyser, A., Verleye, K., Lemon, K.N., Keiningham, T.L. and Klaus, P. (2020). Moving the Customer Experience Field Forward: Introducing the Touchpoints, Context, Qualities (TCQ) Nomenclature. *Journal of Service Research*, 23(4), p.109467052092839. doi: <https://journals.sagepub.com/doi/10.1177/1094670520928390>.
- de Visser, E.J., Pak, R. and Shaw, T.H. (2018). From ‘automation’ to ‘autonomy’: the importance of trust repair in human-machine interaction. *Ergonomics*, 61(10), pp.1409–1427. doi: <https://doi.org/10.1080/00140139.2018.1457725>.
- DeSimone, J.A., Harms, P.D. and DeSimone, A.J. (2014). Best practice recommendations for data screening. *Journal of Organizational Behavior*, 36(2), pp.171–181. doi: <https://doi.org/10.1002/job.1962>.
- Dillman, D.A., Smyth, J.D. and Christian, L.M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4th ed. Hoboken, New Jersey: Wiley.
- Følstad, A. and Brandtzæg, P.B. (2017). Chatbots and the new world of HCI. *Interactions*, 24(4), pp.38–42. doi: <https://doi.org/10.1145/3085558>.
- Følstad, A. and Skjuve, M. (2019). Chatbots for customer service. *Proceedings of the 1st International Conference on Conversational User Interfaces - CUI '19*, pp.1–9. doi: <https://doi.org/10.1145/3342775.3342784>.
- Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. *Proceedings of the 38th International Conference on Information Systems (ICIS)*. Available at: <https://aisel.aisnet.org/icis2017/Implementation/Presentations/7/>.
- Go, E. and Sundar, S.S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, pp.304–316. doi: <https://doi.org/10.1016/j.chb.2019.01.020>.
- Gombolay, M., Jensen, R., Stigile, J., Golen, T., Shah, N., Son, S.-H. and Shah, J. (2018). Human-Machine Collaborative Optimization via Apprenticeship Scheduling. *Journal of Artificial Intelligence Research*, 63, pp.1–49. doi: <https://doi.org/10.1613/jair.1.11233>.

- Grand View Research (2014). *Chatbot Market Size And Share Analysis | Industry Report, 2014 - 2025*. Available at: <https://www.grandviewresearch.com/industry-analysis/chatbot-market>.
- Gursoy, D., Chi, O.H., Lu, L. and Nunkoo, R. (2019). Consumers' acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management*, 49, pp.157–169. doi: <https://doi.org/10.1016/j.ijinfomgt.2019.03.008>.
- Hill, J., Randolph Ford, W. and Farreras, I.G. (2015). Real conversations with artificial intelligence: A comparison between human-human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49, pp.245–250. doi: <https://doi.org/10.1016/j.chb.2015.02.026>.
- Huang, M.-H. and Rust, R.T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, 21(2), pp.155–172. doi: <https://doi.org/10.1177/1094670517752459>.
- Huang, M.-H., Rust, R. and Maksimovic, V. (2019). The Feeling Economy: Managing in the Next Generation of Artificial Intelligence (AI). *California Management Review*, 61(4), pp.43–65. doi: <https://doi.org/10.1177/0008125619863436>.
- Huang, X. et al. (2020). The Role of AI in Customer Service: A Case Study of Chatbot Implementation. *Journal of Business Research*, 115, pp. 357–366.
- Jo, H. & Park, J. (2024). The Impact of Chatbot-Human Interaction on Customer Satisfaction. *Journal of Service Research*, forthcoming.
- Kaczorowska-Spychalska, D. (2019). How chatbots influence marketing. *Management Dynamics in the Knowledge Economy*, 7(3), pp. 341–354. doi: 10.25019/mdke/7.3.05.
- Kaushal, V. and Yadav, R. (2022). Learning successful implementation of Chatbots in businesses from a B2B customer experience perspective. *Concurrency and Computation: Practice and Experience*. doi: <https://doi.org/10.1002/cpe.7450>.
- Khana, A., Abdul Hamid, A.B., Saad, N.M. and Arif, A.R. (2023). Effectiveness of Artificial Intelligence in Building Customer Loyalty: Investigating the Mediating Role of Chatbot in the Tourism Sector of Pakistan. *International Journal of Academic Research in Business and Social Sciences*, 13(9). doi: <https://doi.org/10.6007/ijarbss/v13-i9/18422>.
- Lee, J. et al. (2019). Customer Satisfaction and Chatbot Performance: Insights from the Banking Sector. *Computers in Human Behavior*, 97, pp. 120–130. doi: 10.1016/j.chb.2019.03.032.
- Nordheim, C. B., Følstad, A. & Bjørkli, C. (2019). An Initial Model of Trust in Chatbots for Customer Service - Findings from a Questionnaire Study. *Interact. Comput.*, 31, pp. 317–335. doi: <https://doi.org/10.1093/iwc/iwz022>.
- Radziwill, N. M., & Benton, M. C. (2017). Evaluating the quality of chatbots and intelligent conversational agents. *Software Quality Professional*, 19(3), pp. 25–36.
- Smith, A. & Anderson, M. (2021). Efficiency vs. Empathy: A Comparative Study of Chatbots and Human Agents in Customer Service. *International Journal of Marketing Studies*, 13(3), pp. 45–60. doi: 10.5539/ijms.v13n3p45.
- Webster, Craig and Ivanov, Stanislav Hristov (2020). Robots in Travel, Tourism and Hospitality: Key Findings from a Global Study. Varna: Zangador. Available at SSRN: <https://ssrn.com/abstract=3542208>.
- Zion Market Research. (2018). Size of the chatbot market worldwide in 2017 and 2024 (in million U.S. dollars). *Statista*. Statista Inc. <https://www-statista-com.ezproxy.mdx.ac.uk/statistics/966893/worldwide-chatbot-market-value/>.

Appendix

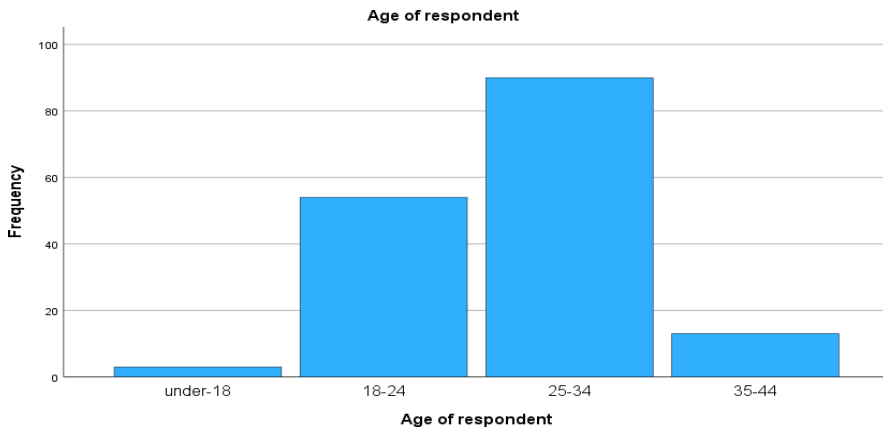


Figure 1: Analysis of the age of the respondent

Source: Own

Table 1: Descriptive Statistics(customer satisfaction with chatbot and human agent while handling routine inquiries)

Descriptive Statistics (customer satisfaction with chatbot and human agent while handling routine inquiries)					
	N	Minimum	Maximum	Mean	Std. Deviation
How satisfied are you with the responses provided by AI-powered chatbots	160	1.00	3.00	1.8812	.74709
How satisfied are you with the service provided by human customer service agents?	160	1.00	5.00	2.3187	.92737
Which type of service do you prefer for handling routine inquiries?	160	1.00	3.00	1.6562	.61477
Valid N (listwise)	160				

Table 2: ANOVA Test(customer satisfaction with chatbot and human agent while handling routine inquiries)

ANOVA						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.208	1	5.208	3.021	.084 ^b
	Residual	272.392	158	1.724		
	Total	277.600	159			
a. Dependent Variable: overall satisfaction						
b. Predictors: (Constant), agent service type						

Table 3: T-Test(customer satisfaction with chatbot and human agent while handling routine inquiries)

T-Test(customer satisfaction with chatbot and human agent while handling routine inquiries)											
	Levene's Test for Equality of Variances			t-test for Equality of Means							
	F	Sig.		t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided P	Two-Sided P			Lower	Upper
How satisfied are you with the responses provided by AI-powered chatbots	Equal variances assumed	.008	.931	- 1.792	146	.038	.075	-.20951	.11691	-.44055	.02154
	Equal variances not assumed			- 1.809	144.709	.036	.073	-.20951	.11584	-.43846	.01944
How satisfied are you with the service provided by human customer service agents?	Equal variances assumed	1.193	.277	.520	146	.302	.604	.07886	.15155	-.22065	.37838
	Equal variances not assumed			.535	144.404	.297	.594	.07886	.14743	-.21253	.37026

Table 4: Descriptive Statistics of Efficiency, Empathy, Personalization, and Human Interaction Preferences in Customer Service

ANOVA Test Results for Customer Preferences, Perceived Personalization, Efficiency, and Empathy in Chatbot and Human-Agent Interactions						
		Sum of Squares	df	Mean Square	F	Sig.
How important is it for you to have the option to speak to a human agent during customer service interactions?	Between Groups	2.972	4	.743	1.295	.275
	Within Groups	88.971	155	.574		
	Total	91.944	159			
Do you feel that human agents provide more personalized service than chatbots?	Between Groups	10.649	4	2.662	4.479	.002
	Within Groups	92.126	155	.594		
	Total	102.775	159			
How would you rate the efficiency of chatbots in handling routine inquiries (e.g., checking account balance, order status)?	Between Groups	42.326	4	10.581	10.904	<.001
	Within Groups	150.418	155	.970		
	Total	192.744	159			
How often do you find chatbots capable of providing empathetic responses to your queries?	Between Groups	17.183	4	4.296	5.799	<.001
	Within Groups	114.817	155	.741		
	Total	132.000	159			

Table 5: Regression – Empathy Predicting Future Chatbot Use

Predictor	B	Std. Error	Beta	t	Sig. (p-value)
Constant	0.99	0.22	—	4.45	< 0.001
Empathy Score	0.61	0.09	0.47	6.64	< 0.001

Table 6: ANOVA – Preference for Human Agents (Complex Inquiries)

Factor	F	df	Sig. (p-value)	Outcome
Perceived Personalization	4.48	(4,155)	0.002	Significant difference
Chatbot Efficiency	10.90	(4,155)	< 0.001	Significant difference
Chatbot Empathy	5.80	(4,155)	< 0.001	Significant difference
Need for Human Option	1.29	(4,155)	0.275	Not statistically significant

EXPLAINABLE DIGITAL TWINS OF PATIENTS: TOWARDS PRECISION AND PERSONALISATION THROUGH COHORT MATCHING

NILMINI WICKRAMASINGHE,¹ NALIKA ULAPANE²

¹ La Trobe University, Optus Chair for Digital Health, Melbourne, Victoria, Australia
N.Wickramasinghe@latrobe.edu.au, nilmini.work@gmail.com

² La Trobe University, Melbourne, Victoria, Australia
n.ulapane@latrobe.edu.au, nalika.ulapane@gmail.com

Modern healthcare services have advanced greatly due to rapid improvements in technology. The next generation of advancements requires precise and personalised treatments, especially for chronic diseases. Computational means are an effective way to achieve this through intelligent decision support assisted by superior data collection and analytics. An emerging concept to facilitate this is digital twins (DTs)—digital replicas of physical entities. DTs have evolved over the years across various industries including aerospace, control engineering, manufacturing, design optimization, and more. DTs in healthcare though, have been explored only relatively recently. One of the most interesting questions lies in creating DTs of humans to model healthcare aspects to enable intelligent decision support. Working towards this quest, this paper attempts to answer the research question: How might precise and personalised treatments for chronic diseases be planned in real-time through explainable digital twins? We attempt to answer this question in the context of breast cancer.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.10](https://doi.org/10.18690/um.fov.4.2025.10)

ISBN

978-961-286-998-4

Keywords:

artificial intelligence,
breast cancer,
cancer,
chronic diseases,
digital health,
digital twin,
explainable AI,
personalisation,
precision medicine



University of Maribor Press

1 Introduction

Digital Twins (DTs), i.e., digital replicas of physical entities, have evolved over the past several decades across various industries (Barricelli et al., 2019; Tao et al., 2018) including aerospace, control engineering, manufacturing, design optimization, and more. The parallel advancements in computation capabilities and internet connectivity that have led to 5G and beyond in telecommunication, and Industry 4.0 and beyond in industrialization, have enabled Artificial Intelligence (AI) to be coupled with DTs, thereby enhancing the capabilities and the applicability of DTs. This evolution encouraged the exploration of DTs in more human-centric sectors like healthcare as well (Katsoulakis et al., 2024).

One of the most potent questions that arose along this exploration is, “Can we create digital twins of human beings?” Why? Because they may be able to capture and help us make sense of the enormous amounts of data being generated in healthcare, and maybe this can help deliver superior healthcare through more precise and personalized medicine derived through data-driven decision support offered to clinicians and other stakeholders including patients. This thought has catalyzed several research attempts in modern times (Björnsson et al., 2020; Liu et al., 2019; Ștefăniță et al., 2024; Wickramasinghe et al., 2021; Wickramasinghe & Ulapane, 2024, 2025; Wickramasinghe, Ulapane, Andargoli, et al., 2022; Wickramasinghe et al., 2023; Wickramasinghe, Ulapane, Nguyen, et al., 2022; Wickramasinghe, Ulapane, Sloane, et al., 2024; Wickramasinghe, Ulapane, Zelcer, et al., 2024), which have attempted to conceptualize and model different aspects of humans, more specifically, patients in healthcare, and healthcare processes, with the objective of deriving data-driven intelligent decision support to help make superior clinical decisions and better manage healthcare processes.

With the involvement of all these data and AI, one of the challenges that emerges is computational complexity (Andargoli et al., 2024) which demands the likes of cloud computing (Dang et al., 2019; Liu et al., 2019). With that comes a whole lot of other questions such as privacy, security, ownership of data, intellectual property, and so on. In the pursuits towards the future, we might find solutions for such questions, but for the time being, it is fair to say that we have not yet found optimal solutions for these challenges. Therefore, whilst maintaining the hype and being ambitious, it

is also important to wind back and be practical in this pursuit for a digitally enabled healthcare.

To do this, it is still important be realistic and to think how might we make use of concepts like DTs, in a computationally feasible manner to be tractable through simple and local computers, or ‘edge devices’ (Rancea et al., 2024) as some call them, so that an interested healthcare provider might be able to implement locally. In the interest of this aforesaid computational simplicity, and also the added benefit of explainability, in this paper, we attempt to answer the following research question: How might precise and personalized treatments for chronic diseases be planned in real-time through explainable digital twins? We attempt to answer this question through using breast cancer as a case study for a chronic disease. More specifically, we target at achieving more precision and personalization in planning a kind of immunotherapy for Triple-Negative Breast Cancer (TNBC) (Foulkes et al., 2010).

2 Review of Related Literature

2.1 Digital Twins (DTs) in Healthcare

A recent scoping review in 2024 by (Katsoulakis et al., 2024) has identified eight main applications of DTs for health; namely, personalized medicine, clinical trials, biomarker and drug discovery, bio-manufacturing, device design, surgical planning, hospital management design & care coordination, and wellness. Several years before, in 2021, (Wickramasinghe et al., 2021) too identified similar applications for DTs in healthcare and viewed DTs as mathematical models irrespective of the application. (Wickramasinghe et al., 2021) identified DTs to function as grey box, surrogate, or black box mathematical models. Grey box DTs are partially or fully governed by well-known principles, such as physics, or statistics. Surrogate DTs need not necessarily have underlain sophisticated analytics capabilities but can be useful for display purposes—such as a dashboard display of an emergency department workflow. Black box digital twins are often underpinned by sophisticated machine learning models, such as neural networks. They may not be apparently explainable but usually have strong data handling and analytics utility. Building on this knowledge, in this paper as an attempt to answer our target research question, we attempt to combine the grey box and surrogate model attributes to realize DTs targeted at personalized medicine. Our choice of grey box and surrogate attributes

helps our DTs function in real-time on edge devices while being computationally simple and explainable.

2.2 Immunotherapy for Triple-Negative Breast Cancer (TNBC)

TNBC is an aggressive type of breast cancer. It is linked with the lack of three receptors: namely, estrogen receptors, progesterone receptors, and human epidermal growth factor receptor 2 (HER2) (Dass et al., 2021; Foulkes et al., 2010). As a result, TNBC is typically more aggressive. It is also faster in growth rate and has higher risk of metastasis. Moreover, it has limited treatment options and poor prognosis (Kesireddy et al., 2024; Obidiro et al., 2023).

In this backdrop, certain immunotherapies, such as certain immune checkpoint inhibitors have been identified as potential treatment options for TNBC. Such immune checkpoint inhibitors typically make cancer cells more vulnerable to our body's own immune system, thereby making our own immune system work against cancer. However, when such treatment is carried out, it is not only the cancer cells that become vulnerable to the immune system, because such immunotherapies typically cannot be delivered locally. When delivered, such treatments affect the whole body, and thus healthy cells too become more vulnerable to the immune system, leading to certain complications and side-effects. That is a trade-off of this type of treatment. However, immune checkpoint inhibitors such as PD-1 and PD-L1 inhibitors have been approved in certain countries, including Australia, as a form of immunotherapy less than a decade ago to treat TNBC. The outcomes and side effects, though, have remained variable. It is therefore beneficial if it is possible to predict likely prognosis prior to treatment, hence we have chosen TNBC as a case study for our DTs. Through this case study we explore as to how DTs might be used for precise and personalized planning of PD-1/PD-L1 inhibitor treatment for TNBC candidates by trying to find beforehand, who is statistically congruent to best respond to which PD-1/PD-L1 inhibitor treatments.

3 Methodology

We used an approach inspired by Design Science Research Methodology (DSRM) (Hevner et al., 2010) to design a DT platform to help precise and personalized planning of PD-1/PD-L1 inhibitor treatments for TNBC. The approach followed

is depicted in Figure 1. Details about each step are provided in the following subsections.

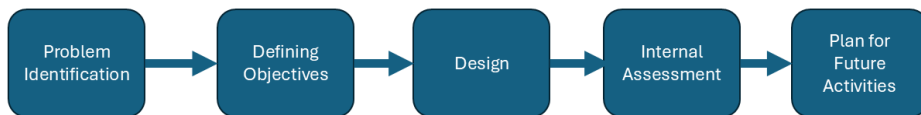


Figure 1: The DSRM-inspired methodology followed in this study

3.1 Problem Identification Phase

This work started through a collaboration between the authors and a team of cancer biologists who study the effects of immunotherapies on breast cancer at a cellular biomarker level. This collaboration started in a backdrop of immune checkpoint inhibitor treatments such as PD-1 and PD-L1 inhibitors have been approved in Australia recently, especially for TNBC patients. The state of matters was that such treatments were quite expensive, and a limited number of candidates were eligible to receive them in Australia. Even among the limited recipients, the outcomes and the side effects were variable. Therefore, we asked the pertinent question as to whether we can predict the prognosis of PD-1 and PD-L1 inhibitor treatments on TNBC patients, as this would enable the identification of best receptive patients, prior to receiving immunotherapy treatment, and in turn this would reduce the incidences of adverse outcomes and side-effects. This line of thought led the authors, and the collaborators to secure the Victorian Medical Research Acceleration Fund (VMRAF) grant GA-F4669635-1352, to explore avenues for predicting prognosis of PD-1 and PD-L1 inhibitor treatments, making use of genetic information obtained from TNBC patients. The view from the collaborating cancer biologists was that they have a wealth of genetic information, but it was difficult to find direct patterns. This opened the avenue to explore the use of machine learning and AI to address this problem.

Moreover, as part of the ‘problem identification’ phase, a rapid review of literature was conducted with the objective of mapping out the diversity of immunotherapy treatment options and outcomes pertaining to TNBC. The protocol followed for this review was published in Open Science Framework at: <https://osf.io/2bzds/>.

The PRISMA guidelines (Stevens et al., 2018) and the Cochrane methodology (Garritty et al., 2024) were followed for reviewing the literature and reporting. The outcome of this review was a set of themes that summarized the diversity of immunotherapy treatment options for TNBC, along with associated outcome measures and side-effects. These findings are summarized in Figure 2 in the Results section.

3.2 Defining Objectives Phase

Following the ‘problem identification’ phase, we defined objectives for a potential solution. Given the nascency of this area, we decided to be guided by the themes identified from the conducted rapid review to design a computationally simple web-based DT platform that can provide clinical decision support on precise and personalized planning of PD-1/PD-L1 treatments for TNBC. The web-based function was chosen to ensure easy accessibility over the internet, while computational simplicity was weighed in to ensure ability to implement on edge devices and local servers. To achieve computational simplicity, for this design we chose offering decision support purely based on statistical congruence—meaning assessing how statistically similar a present patient is to cohorts of past patients. Thereby we derive decision support based on DTs constructed from a cohort matching approach.

3.3 Design Phase

At the design phase we planned three key aspects for a solution. The first was the patient journey with the DT platform integrated within. The planned patient journey is depicted in Figure 3 in the Results section. This DT platform was planned to perform as a clinical decision support tool. Therefore, it is designed primarily for the use of clinicians. However, it can be used by clinicians in consultation with patients for shared decision-making. The second aspect we designed was a web-based frontend to display results that would assist clinical and shared decision-making. The designed web-based frontend is depicted in Figure 4 in the Results section. Thirdly, we designed graphical means to display results more elaborately along with DTs of patients identified based on statistical congruence, or cohort matching—meaning, finding cohorts of past patients that statistically best match a present patient. Details about the functionality of the DT dashboard are provided in the Results section.

3.4 Internal Assessment Phase

Following the 'design phase,' the designs were informally discussed with the collaborating cancer biologists. Their feedback, thoughts and recommendations were recorded. Their recommendations are summarized in the Results section.

3.5 Plan for Future Work

As future work it was planned to refine the web-based frontend as well as the statistical rules for cohort matching based on the feedback received from the cancer biologists and implement the platform. Initially, due to the lack of data from TNBC patients, it was decided to run the platform based on synthetic data produced mimicking real-world patients.

4 Results and Discussion

4.1 Diversity of Immunotherapy Treatments on TNBC

The findings from the rapid review that was conducted as part of the 'problem identification' phase were summarized highlighting the diversity of immunotherapy treatment options for TNBC, along with associated outcome measures and side-effects. These findings are presented in Figure 2.

4.2 The Planned Patient Journey

The patient journey that was planned as part of the 'design phase' is depicted in Figure 3. The DT platform is integrated within the patient journey as a clinical decision support tool plus a decision aid for shared decision making through the collaboration of both clinicians and patients.

4.3 The Designed Web-Based DT Frontend

The designed web-based DT frontend is depicted in Figure 4 and it is followed by a description about how decision support is derived through our proposed cohort matching approach.

Immunotherapy Type	Treatment Drug	Objective response rate (ORR) (%)	Median Progress free survival (PFS) (months)	Median Overall survival (OS) (months)	Specific Biomarkers	Significant Findings	Adverse Events / Safety
PD-1 inhibitors	Pembrolizumab	21 - 43%	1.8 - 11.7	6.3 - 26.0	PD-L1 CPS ≥ 1, TNBC, CD8+ T-cells	PD-L1 CPS ≥ 1 correlated with improved responses and survival.	Fatigue, nausea, diarrhea, immune-related AEs (e.g., rash, thyroid dysfunction).
	Nivolumab	12.5 - 20%	1.4 - 3.7	5.6 - 8.1	PD-L1+, TNBC, stTILs, immune gene expression	Predict responses, higher benefit in biomarker-selected populations (e.g., TNBC).	Fatigue, anaemia, immune-related toxicities (e.g., colitis, pneumonitis), TILs).
	Camrelizumab	43.3 - 81.3%	3.7 - 13.6	8.1 - 12	CD8+ T-cells, PD-L1+, IL-8, VEGFR2	Continuous dosing improved outcomes; biomarkers CD8+ T-cells showed influence.	Anaemia, neutropenia, fatigue, increased toxicity with continuous dosing.
	Toripalimab	66%	8.4	32.8 (PD-L1+)	CDNA, PD-L1+	HIGH response observed in PD-L1+ subgroups; CDNA detection associated with relapse.	Immune-related AEs consistent with PD-1 inhibitors (e.g., rash, thyroid dysfunction), reductions minimized severe toxicities.
	SHR-1210	No data	3.7	Not reported	TGF-β, OPN, PD-1/PD-L1 expression	Combination with low-dose Apatinib improved outcomes; biomarkers influenced response.	Manageable safety, mild side effects; dose reductions minimized severe toxicities.
PD-L1 inhibitors	Atezolizumab	20 - 56%	4.3 - 7.5	14.7 - 28.0	PD-L1+, CD8+ T-cells, TILs	PD-L1+ tumours benefited most; strong association with CD8+ T-cells.	Fatigue, neutropenia, nausea, immune-related AEs (e.g., hypothyroidism, colitis).
	Durvalumab	36 - 53.4%	2.7 - 6.1	18.3 - 21.2	BRCA mutations, PD-L1+, TILs	Survival benefits observed in early-stage and combination therapies.	Nausea, anaemia, fatigue, immune-related AEs observed, including maculopigmentation.
	Avelumab	18.2 - 34.8%	3.6 - 5.3	Limited data	BRCA mutations, DNA damage repair	Durable responses in DNA damage repair-positive tumours and BRCA-attenuated cancers.	Haematological toxicities (e.g., anaemia, neutropenia, thrombocytopenia).
	Atezolizumab	90% (PCR)	No data	No data	PD-L1+	HIGH PCR rates with combination therapy; manageable safety profile.	Grade 3-4 AEs common (e.g., neutropenia, anaemia, leukopenia); manageable overall profile.
PD-1 + VEGFR inhibitors	Apatinib, Farnitinib	43.3 - 81.3% (continuous dosing)	3.7 - 13.6	8.1 - 12	CD8+ T-cells, PD-L1+, IL-8, VEGFR2	Continuous dosing improved outcomes; biomarkers like CD8+ T-cells showed influence.	Haematological toxicities (anaemia, neutropenia), hand-foot syndrome, hypertension.
PD-1/PD-L1 + PARP inhibitors	Olaparib, Talazoparib, Fucloparib	6.3 - 54.8%	3.7 - 6.1	Limited data	BRCA mutations, PD-L1+, DNA damage	Biomarkers such as BRCA mutations and PD-L1 trends influenced efficacy.	Haematological toxicities (e.g., anaemia, neutropenia), fatigue, hypertension.
Radiotherapy + immunotherapy	Various PD-1/PD-L1 inhibitors	17.6 - 90%	No data	No data	PD-L1+, TNBC, CD8+ T-cells	Enhanced pathological complete response (pCR) rates and objective responses rate.	Lymphopenia, dermatitis, fatigue, immune-related AEs observed with combination therapy.
Combination with Chemotherapy	Nab-paclitaxel, Paclitaxel, Carboplatin, Etoposide, Capecitabine, Eribulin	29 - 81.3%	4.2 - 13.6	16.1 - 28.9	PD-L1 CPS ≥ 1, TILs, CD8+ T-cells	Combination therapies consistently outperformed monotherapy, particularly in PD-L1+.	Neutropenia, anaemia, nausea, immune-related AEs with PD-1/PD-L1 inhibitors.

Figure 2: Key findings from the rapid review (<https://osf.io/2bzds/>) conducted on immunotherapy for TNBC

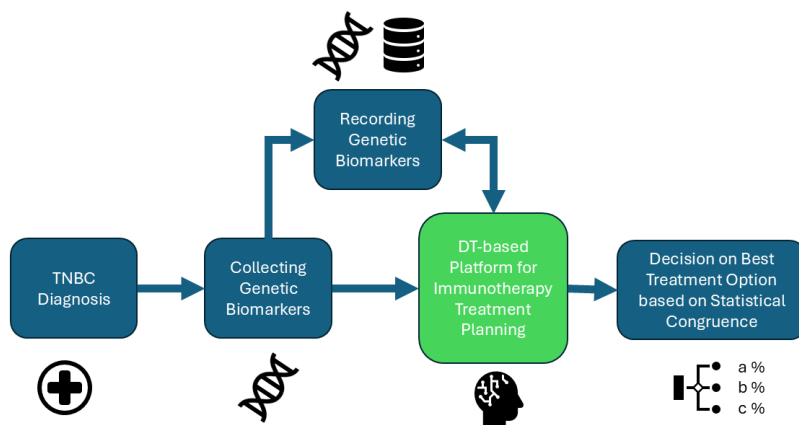


Figure 3: Proposed patient journey with DT platform integrated

Cancer Type: TNBC

Load Patient

Disease Detail

Cancer Stage: IIA

Biomarkers:

PD-L1 CPS Score = aaa ±10%

CD8+ T-cells = bbb ±10%

Treatment

PD-L1 I: Atezolizumab

Demographic

Age: 45 ±5 years

Sex: Female

Ethnicity: European

Match Cohort

Outcome Measures

Population: 10,000

Cohort: 855

ORR: 46% - 56%

PFS (months): 5.5 - 7.5

OS (months): 17.7 - 25.0

Figure 4: Snapshot of the proposed DT frontend dashboard for precise and personalized immunotherapy treatment planning for TNBC (please zoom in if required to read texts in this image)

The DT frontend in Figure 4 will be accessible for clinicians as a clinical decision support tool. It can be used in collaboration with patients for shared decision making. At the top right, it has a button named ‘Load Patient.’ That is where the process starts. When a patient is in front of the clinician, the process can be started by the clinician pressing on the ‘Load Patient’ button. This will provide an option

like scanning a unique QR code for the patient. Once the unique identifier (QR code or similar) is scanned, the patient's genetic biomarkers along with other relevant information will be fetched from a dedicated server. The fetched information will be displayed onscreen as shown in Figure 4. To confirm the fetched data matches the present patient, a preceding popup window will appear, showing some patient personal identifiers like name and date of birth. That information will appear only on that popup window but not on the DT screen shown in Figure 4 to preserve patient's privacy. This popup window can be cross-checked by the clinician and the patient to confirm the right information has been fetched. Once confirmed, the popup window can be closed, and then the relevant information will appear on the DT screen shown in Figure 4.

As can be seen in Figure 4, patient information is stratified into two groups, i.e., demographic information and disease detail. Demographic information includes patient's age (in years, inclusive of a range plus or minus five years for cohort matching), sex (Male or Female), and ethnicity (e.g., African, East Asian, South Asian, Southeast Asian, Middle Eastern and North African (MENA), European, Hispanic or Latino, Indigenous Peoples, Pacific Islander, Mixed or Multi-Ethnic, etc.). All such demographic details are eventually used to select a cohort of past patients to which the present patient will match on a statistically congruent basis—thereby achieving a degree of personalization.

Then, as disease details, the dashboard is specifically designed for TNBC. The other main disease information considered is the cancer stage (e.g., Stage IA, IB, IIA, etc.). Next, relevant genetic biomarkers are considered. In the context of TNBC and PD-1/PD-L1 inhibitors, the literature review we conducted as part of the 'problem identification' phase revealed that biomarkers such as PD-L1 CPS score, CD8+ T-cell infiltration and Tumor Mutational Burden amongst others are relevant. These biomarkers are designed to be considered as quantitative values, i.e., a real-valued number. For cohort matching, these numbers are considered with a margin of $\pm 10\%$. That means, past patients who have had biomarker values within $\pm 10\%$ of the present patient will form a matching cohort to the present patient.

Next, the clinician can select available immunotherapy treatment options (e.g., PD-1 inhibitors (e.g., Pembrolizumab, Nivolumab, Camrelizumab, Toripalimab, or SHR-1210), PD-L1 inhibitors (e.g., Atezolizumab, Durvalumab, Avelumab, or

Adebrelimab), and click on the button named ‘Match Cohort.’ This will then fetch the outcome data of the past patients within the matching cohort to the present patient, and display certain outcome measures, and possibly side-effects also as shown in the ‘Outcome Measures’ panel in Figure 4. Displayed outcome measures include the likes of Objective Response Rate (ORR), Problem Free Survival (PFS) range in months, and Overall Survival (OS) range in months, and likely side-effects (e.g., fatigue, nausea, list of immune-related adverse events (AEs), etc.). This display can be used by clinicians to make informed decisions about the best treatment options that are statistically congruent pertaining to the present patient, thereby achieving a degree of personalization. These decisions can be made collectively with patients, thereby achieving patient empowerment through shared decision making. When presenting outcome measures, data of a present patient is interfaced and used in an intelligent digital model of a patient constructed for the purpose of clinical decision support, thereby the DT paradigm being realized. Furthermore, we have planned more elaborate displays of outcome measures such as the one in Figure 5, which illustrates the population of past patients along with a discovered DT (i.e., the most statistically congruent cohort to the present patient). Such displays can better express the degree of personalization that has taken place.

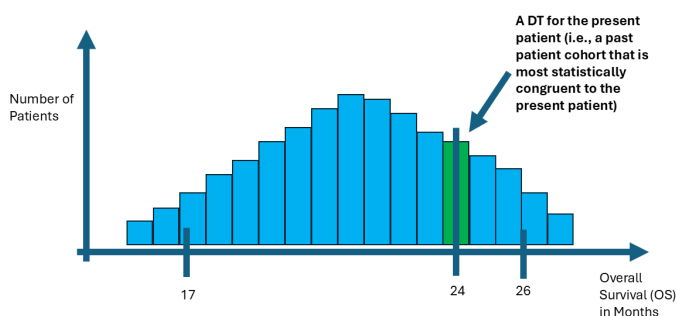


Figure 5: Population of past patients alongside a discovered DT (i.e., the most statistically congruent cohort to the present patient) depicting personalization of outcome measures

4.4 Outcomes from the Internal Assessment Phase

From the discussions that were carried out with the collaborating cancer biologists, their main suggestion was about the biomarkers. They suggested using biomarker quantities normalized per square area and normalized per number of cancer cells

present. In an intuitive mathematical sense, normalizing biomarkers as such makes great sense, and therefore we decided to use normalized values in our future implementations of this platform.

5 Conclusions

Answering the research question “How might precise and personalized treatments for chronic diseases be planned in real-time through explainable digital twins?” was attempted. A DSRM-inspired approach was followed to answer the research question. A computationally simple, real-time, and explainable DT platform that is based on statistical congruence was proposed. This is designed to provide clinical decision support to achieve precise and personalized planning of immunotherapy treatments for TNBC.

This paper contributed to theory by proposing an approach of matching cohorts of past patients based in statistical congruence to present patients as a form DTs capable of deriving decision support and insights based on statistics, thereby preserving computational simplicity, explainability and the ability to implement on edge devices and local servers to offer real-time clinical decision support.

As a contribution to practice, this paper proposed a DT-incorporated patient journey, frontend designs for a DT platform, and certain statistical and logical steps for deriving DTs and insights through cohort matching in a computationally simple and explainable manner.

TNBC and immunotherapy, the target healthcare context of this work, is new and we are restricted by the lack of data from TNBC patients. Therefore, we had to resort to proceeding with synthetic data to develop our algorithms. The lack of real data and having to rely on synthetic data is a limitation that underpins our work.

DTs in healthcare are at their infancy but there is the potential for significant benefits to realize simultaneously more personalized and precise clinical decision support. Future work will focus on refining our algorithms and DT platforms based on feedback received from our collaborating cancer biologists and the latest findings in cancer treatment, all whilst being condescended to value-based healthcare principles.

Acknowledgements

This work is an outcome of VMRAF grant GA-F4669635-1352 secured by the authors. The support by Optus and La Trobe University, Australia, are acknowledged.

References

- Andargoli, A. E., Ulapane, N., Nguyen, T. A., Shuakat, N., Zelcer, J., & Wickramasinghe, N. (2024). Intelligent decision support systems for dementia care: A scoping review. *Artificial Intelligence in Medicine*, 102815.
- Barricelli, B. R., Casiraghi, E., & Fogli, D. (2019). A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE access*, 7, 167653-167671.
- Björnsson, B., Borrebaeck, C., Elander, N., Gasslander, T., Gawel, D. R., Gustafsson, M., Jörnsten, R., Lee, E. J., Li, X., & Lilja, S. (2020). Digital twins to personalize medicine. *Genome medicine*, 12, 1-4.
- Dang, L. M., Piran, M. J., Han, D., Min, K., & Moon, H. (2019). A survey on internet of things and cloud computing for healthcare. *Electronics*, 8(7), 768.
- Dass, S. A., Tan, K. L., Selva Rajan, R., Mokhtar, N. F., Mohd Adzmi, E. R., Wan Abdul Rahman, W. F., Tengku Din, T. A. D. A.-A., & Balakrishnan, V. (2021). Triple negative breast cancer: a review of present and future diagnostic modalities. *Medicina*, 57(1), 62.
- Foulkes, W. D., Smith, I. E., & Reis-Filho, J. S. (2010). Triple-negative breast cancer. *New England journal of medicine*, 363(20), 1938-1948.
- Garrity, C., Hamel, C., Trivella, M., Gartlehner, G., Nussbaumer-Streit, B., Devane, D., Kamel, C., Griebler, U., & King, V. J. (2024). Updated recommendations for the Cochrane rapid review methods guidance for rapid reviews of effectiveness. *bmj*, 384.
- Hevner, A., Chatterjee, S., Hevner, A., & Chatterjee, S. (2010). Design science research in information systems. *Design research in information systems: theory and practice*, 9-22.
- Katsoulakis, E., Wang, Q., Wu, H., Shahriyari, L., Fletcher, R., Liu, J., Achenie, L., Liu, H., Jackson, P., & Xiao, Y. (2024). Digital twins for health: a scoping review. *NPJ digital medicine*, 7(1), 77.
- Kesireddy, M., Elsayed, L., Shostrom, V. K., Agarwal, P., Asif, S., Yellala, A., & Krishnamurthy, J. (2024). Overall survival and prognostic factors in metastatic triple-negative breast cancer: a national cancer database analysis. *Cancers*, 16(10), 1791.
- Liu, Y., Zhang, L., Yang, Y., Zhou, L., Ren, L., Wang, F., Liu, R., Pang, Z., & Deen, M. J. (2019). A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE access*, 7, 49088-49101.
- Obidiro, O., Battogtokh, G., & Akala, E. O. (2023). Triple negative breast cancer treatment options and limitations: future outlook. *Pharmaceutics*, 15(7), 1796.
- Rancea, A., Anghel, I., & Cioara, T. (2024). Edge computing in healthcare: Innovations, opportunities, and challenges. *Future internet*, 16(9), 329.
- Ștefăniță, S. A., Cordoș, A. A., Ivascu, T., Feier, C. V. I., Muntean, C., Stupinean, C. V., Călinici, T., Aluș, M., & Bolboacă, S. D. (2024). Advancing Precision Oncology with Digital and Virtual Twins: A Scoping Review. *Cancers*, 16(22), 3817.
- Stevens, A., Garrity, C., Hersi, M., & Moher, D. (2018). Developing PRISMA-RR, a reporting guideline for rapid reviews of primary studies (Protocol). *Equator Network*, 1-12.
- Tao, F., Zhang, H., Liu, A., & Nee, A. Y. (2018). Digital twin in industry: State-of-the-art. *IEEE Transactions on industrial informatics*, 15(4), 2405-2415.
- Wickramasinghe, N., Jayaraman, P. P., Forkan, A. R. M., Ulapane, N., Kaul, R., Vaughan, S., & Zelcer, J. (2021). A vision for leveraging the concept of digital twins to support the provision of personalized cancer care. *IEEE Internet Computing*, 26(5), 17-24.
- Wickramasinghe, N., & Ulapane, N. (2024). Digital Twins of Patients: A Liquid Neural Network Interpretation. *ACIS 2024 Proceedings*. 32. <https://aisel.aisnet.org/acis2024/32>

- Wickramasinghe, N., & Ulapane, N. (2025). A Solution for the Health Data Sharing Dilemma: Data-Less and Identity-Less Model Sharing Through Federated Learning and Digital Twin-Assisted Clinical Decision Making. *Electronics*, 14(4), 682.
- Wickramasinghe, N., Ulapane, N., Andargoli, A., Ossai, C., Shuakat, N., Nguyen, T., & Zelcer, J. (2022). Digital twins to enable better precision and personalized dementia care. *JAMIA open*, 5(3), ooac072.
- Wickramasinghe, N., Ulapane, N., Andargoli, A., Shuakat, N., Nguyen, T., Zelcer, J., & Vaughan, S. (2023). Digital twin of patient in clinical workflow. *Proceedings of the Royal Society of Victoria*, 135(2), 72-80.
- Wickramasinghe, N., Ulapane, N., Nguyen, T. A., Andargoli, A., Ossai, C., Shuakat, N., & Zelcer, J. (2022). Towards discovering digital twins of dementia patients: matching the phases of cognitive decline. *Alzheimer's & Dementia*, 18, e066336.
- Wickramasinghe, N., Ulapane, N., Sloane, E. B., & Gehlot, V. (2024). Digital Twins for More Precise and Personalized Treatment. In *MEDINFO 2023—The Future Is Accessible* (pp. 229-233). IOS Press.
- Wickramasinghe, N., Ulapane, N., Zelcer, J., & Saffery, R. (2024). 'Omics-Based Digital Twins for Personalised Paediatric Healthcare. *Studies in Health Technology and Informatics*, 318, 180-181.

TECHNOLOGICAL FRAMES AND LEARNING EXPERIENCES: STUDENTS' INTERPRETATIONS OF RULE-BASED CHATBOTS IN HIGHER EDUCATION

MERJA ALANKO-TURUNEN, OLLI-JAAKKO KUPIAINEN,
MARI RAASKA

Haaga-Helia University of Applied Sciences, Helsinki, Finland
merja.alanko-turunen@haaga-helia.fi, olli.kupiainen@haaga-helia.fi,
mari.raaska@haaga-helia.fi

This study examines how students construct technological frames of rule-based chatbots and how these frames relate to their learning experiences, particularly information processing and reflection. This research is grounded in the social construction of technology and technological frames theory. Semi-structured interviews were conducted with 45 undergraduate students at a University of Applied Sciences. Through thematic analysis, three distinct technological frames were identified: collaborative, completion-driven, and unhelpful. Each frame was found to shape how chatbot technology was engaged with and integrated into students' learning processes. This study extends technological framing theory to the educational field, linking technological frames to students' learning experiences. We also provide an empirically grounded framework that helps understand student learning experiences within chatbot-supported learning environments. This research contributes to an understanding of how interpretations of technology construct educational experiences and emphasizes the critical relationship between pedagogy and technology.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.11](https://doi.org/10.18690/um.fov.4.2025.11)

ISBN
978-961-286-998-4

Keywords:
rule-based chatbot,
technological frame,
educational technology,
learning experience,
higher education



University of Maribor Press

1 Introduction

Chatbots have become an increasingly common feature in higher education, offering students structured learning support and facilitating interactions in online environments (Følstad et al., 2019; Kuhail et al., 2023; Labadze et al., 2023; Okonkwo & Ade-Ibijola, 2021; Wollny et al., 2021). While AI-driven chatbots are now widely available in higher education (Almogren et al., 2024; Ayanwale & Ndlovu, 2024; Stöhr et al., 2024; Tlili et al., 2023) rule-based chatbots are still relevant because their structured design efficiently guides students through pre-defined learning paths, while allowing learners also to take ownership of their educational journey. Rule-based chatbots empower students to engage with course material at their own pace, in their preferred location, and at times that best suit their individual learning, offering the opportunity for a more personalized learning experience (e.g. Vanichvasin, 2021; Winkler & Söllner, 2018).

Rule-based chatbots have been utilized in higher education in various ways. They support students' goal setting and social presence in online learning (Hew et al., 2023) and provide personalized guidance tailored to individual learning needs (Papakostas et al., 2024). Research shows that rule-based chatbots have been used in language teaching (Fryer et al., 2018; Lee et al., 2020) and they have had a positive impact on learning outcomes in science education (Yin et al., 2020). However, it is important to note that rule-based chatbots are typically rigid communication systems operating strictly through predefined scripts. As a result, they are not particularly "intelligent" and cannot answer questions they were not programmed to answer (Adamopoulou & Moussiades, 2020; Kohnke et al., 2023).

Integration of new technologies into educational settings inevitably transforms students' learning experiences; it represents an interconnected, co-evolving phenomenon arising from complex social, technological, and pedagogical interactions (Anderson, 2016; Marell-Olsson & Jahnke, 2019). The rule-based chatbots in this study were designed to support the constructivist learning approach by scaffolding learners' active knowledge construction and integration of new concepts with existing knowledge. More importantly, the role of the rule-based chatbots was to promote reflection during learning (Kirkwood & Price, 2013).

While the widely used Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT) explain why students adopt chatbots (e.g. Al-Abdullatif, 2023; Güldal & Dinçer, 2024; Malik et al., 2021; Strzelecki, 2024), they do not capture how students interpret and integrate them into their actual study practices. These models fail to address how students' framing of chatbot technology influences their theoretical understanding and practical way of utilizing these applications. To explore this, our research question *was how students construct technological frames of rule-based chatbots, and how these frames relate to their learning experiences, especially information processing and reflection.*

Data was collected through semi-structured interviews with 45 undergraduate students. We investigated how students actively constructed and framed (Orlikowski & Gash, 1994) rule-based chatbots. This approach allows us to highlight the constructive dynamics that shape technology adoption in educational settings.

Our findings highlight three technological frames for chatbot interaction: collaborative, completion-driven, and unhelpful. While some students engaged collaboratively, others approached chatbots superficially or found them technologically limited. This study extends technological framing theory to the educational field, linking technological frames to students' learning experiences. We also provide an empirically grounded framework that helps understand student learning experiences within chatbot-supported learning environments. This suggests that effective educational chatbots must align with student expectations and learning experiences, while incorporating more adaptive features. The study emphasizes student's active role in educational technology adoption.

2 Theoretical Background

2.1 Technological frames

The social construction of technology (SCOT) is one of the constructivist ways of studying science and technology. It provides a theoretical framework for understanding the relationship between society and technology (Bijker, 2001). The significance of SCOT for our research is important as it challenges the assumption of technological determinism by offering an alternative perspective: technology alone does not determine human behaviour. Instead, technology is culturally shaped

and interpreted through processes involving multiple actors with their own interests and agendas (e.g. Silverstone, 2006). This perspective is complemented by the concept of technological framing, which emphasizes that interpretive frames of technology can change as individuals engage with applications in particular contexts, rather than being predetermined by the features of the technology (Davidson & Pai, 2004). These two perspectives enable our research to analyse the social construction of technology and the users' interpretive framings.

Individuals' experiences, resources, expectations and beliefs about technology can lead to different technological interpretations (Treem et al., 2015). Technological framings are central to structuring individuals' experiences, interpreting ambiguous situations, and reducing uncertainty. They provide a basis for action in changing circumstances (Lin & Silva, 2005).

Technological applications shape learning experiences through their affordances (Conole & Dyke, 2004; Van Osch & Mendelson, 2011). Affordance as a concept originates from ecological psychology, where it refers to the possibilities for action that the environment offers to the agent (Gibson, 2014). The concept has been extended to the field of design, distinguishing between real and perceived affordances (Norman, 1999). According to Gaver (1991), technological affordances are both social and material constructions (see also Selwyn, 2012), which relates directly to the framework of our study on the social construction of technology. Students actively interpret and adapt these affordances to serve their learning needs and contexts (Oliver, 2013; Squire & Dijkers, 2012).

Technological frames (Orlikovski & Gash, 1994; Kaplan & Tripsas, 2008; Spieth et al., 2021) are cognitive structures that individuals and groups use to understand and make sense of technology. Their studies emphasized the importance of making sense of new technologies in business organizations. Our study builds on these perspectives on technological framing, examining how students apply these frames in higher education. For example, when examining how students frame technological applications, focus can be placed on how they interpret the nature of technology (Orlikowski & Gash, 1994). Technology can be perceived as a learning enabler when students recognize its role in achieving learning outcomes. Additionally, students' assumptions, expectations, and prior user experiences shape their technology-in-use, influencing how different technological applications are

actively utilized for learning (Orlikowski & Gash, 1994). For example, a student may use an application as a ready-made stimulus for learning or simply as a content provider. Students rely on their technological frames to interpret technology's functionality and assess its implications (cf. Treem et al., 2015).

A recent study examined how students framed a mobile interaction application designed to engage them during online lectures. They largely rejected the application, describing it as inappropriate for the context, socially uncomfortable and an unacademic way of learning (Drew & Mann, 2018). In contrast, another study found that students valued mobile devices as learning aids for quickly accessing information and engaging on social networks. Mobile devices empowered students to work efficiently and achieve their goals through collaboration with peers, teachers, and employers (Squire & Dikkers, 2012). Social and technical factors influence how students frame technology, and these frames evolve through experience and reflection, particularly in group work (Bjørn et al., 2006). Thus, prior research highlights the diversity and contextual dependency of technology framing.

2.3 Integrating educational technology and learning experience in this study

Learning experiences, inherently intertwined with technology in contemporary education, have been examined through diverse theoretical lenses (Anderson, 2016; Jahnke, 2023; Shrerer et al., 2019). In this study, we define a learning experience as the multifaceted technology-mediated interactions between learners, content, educators, and peers that facilitate information processing, skill development, and reflection while promoting attitudinal and personal growth (Liu et al., 2023). Furthermore, Chen et al. (2018) emphasize that the perception of a learning experience is influenced by learners' personal epistemologies, their sense of competency, self-efficacy beliefs, and attitudes toward learning.

We assigned three pedagogical roles to the rule-based chatbots: they served as a learning assistant to enhance information processing and reflection (Wollny et al., 2021), a teaching assistant to engage students (Kuhail et al., 2023) and a provider of feedback to students. On the technological side, the chatbot relied on chatbot-driven conversations, meaning that the chatbot controls the flow of the conversations (Kuhail et al., 2023). We built the chatbots on *Landbot.pro* which is an intuitive no-

code conversational chatbot builder. External technological expertise was required to develop a pedagogical script to ensure a balanced learning experience encompassing content, user experience, and social interaction (e.g. Huhtanen, 2020). Subsequently, the chatbots were piloted with students.

Designing for learning involves creating scripts and scaffolds that shape learners' activity. These scaffolds influence how activity unfolds, while the activity itself continuously modifies and reconstructs the learning environment (Goodyear & Dimitriadis, 2013). While designing the content, we focused on practically applicable material. This was achieved by incorporating brief videos in which an educator explained the core content and then encouraged students to apply and reflect on its meaning in relation to their own experiences. We designed the content and user experience to feature interactive chatbot conversations with both open-ended and multiple-choice questions (cf. Kuhail et al., 2023). We created learning activities for the students to do (Figure 1) during the session with the chatbot (e.g. students reflected their project planning competences after watching the educator's video of topic). However, what the students do with the activities may diverge from what we intended—since students often engage with tasks in ways that reflect their own interpretations and preferences (e.g. Goodyear et al., 2021; Ellis & Goodyear, 2010).

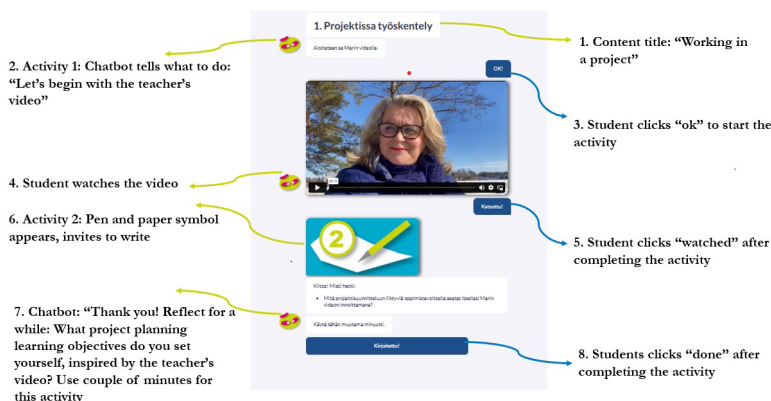


Figure 1: Joors Truli ("Yours truly" in English) chatbot

When designing the social experience, we aimed to enhance student motivation and engagement. To achieve this, we used a conversational style and everyday language to guide students, while the chatbot also asked clarifying questions (Følstad et al.,

2019; Kim et al., 2019). Additionally, the chatbot facilitated participation experiences, providing students with opportunities to engage in meaningful conversations while receiving guidance (Chang et al., 2023; Lin & Chang, 2023). For example, in one chatbot conversation, students reflected on their project work priorities and connected their insights to previously studied Belbin team roles

3 Methodology

3.1 Research context

Our interpretative study explores how undergraduate business students of a University of Applied Sciences frame educational technology and how the framing relates to their learning experiences, especially their information processing and reflection. The courses (5 ECTS each) were a mandatory *Teamwork and Project Skills* and an elective *Leadership Skills*. Students engaged in online interactions with a rule-based chatbot on the Moodle learning platform as part of their required course work; these interactions formed the basis for their subsequent reflective essays. The chatbots addressed one specific theme in each course – developing teamwork and project planning skills in the former, and the latter introduced team leader's role in advancing team members' work well-being (cf. Lindebaum, 2024). These themes are important because as modern workplaces require understanding of teamwork components (Salas et al., 2015) and higher education students need leadership practice (Siewiorek et al., 2012).

3.1 Participants and data collection

We interviewed 45 voluntary students who had used chatbots in their courses in the spring of 2023. Participant recruitment was conducted via announcements on the Moodle platform. The semi-structured interview questionnaire was inspired by the Metacognitive Awareness Inventory (MAI) (Schraw & Dennison, 1994) which builds on Flavell's (1979) metacognition theory. The operationalization through MAI examined cognitive dimensions by exploring how students constructed understanding of chatbot effects on their information processing. Skill-related dimensions focused on students' goal setting, planning, and strategic integration of chatbots into personalized learning. Attitudinal dimensions interpreted how students

reflected on chatbot utility and constructed their perspectives on these tools as learning support.

Interviews were conducted using MS Teams application, which allowed us to take an advantage of automatic transcription feature. Interviews were conducted by a course educator from the research team and an experienced external educator. All the interviews lasted between 15 and 25 minutes and followed a semi-structured format. To ensure anonymity, informants' personal information was removed from the transcripts. A short summary was added at the beginning of each transcript to guide the subsequent analysis of the interviews.

3.2 Analysis process

Our thematic analysis was informed by Braun and Clarke (2006). The analysis was iterative and collaborative, involving regular reflective discussions throughout the process. In the initial discussions about the research material, we observed that students' interpretations of technology varied. Some students criticized its functionality, others highlighted the importance of peer interaction in classrooms, and many felt that educational technology supported their learning. Our discussions were critical in understanding the data from a theoretical perspective (Carlsson, 2023) eventually formalizing our understanding of students' technological frames (Orlikowski & Gash, 1994) and their learning experiences. We identified key themes, e.g., a positive learning experience students had with the chatbots. We then consulted literature and decided to focus on our analysis especially on the nature of technology and technology-in-use (Orlikowski & Gash, 1994) as well as to students learning experiences. Table 1 illustrates the coding process which we explain below.

We reconvened to discuss our insights and findings, which allowed us to identify three overarching themes of technological frames. Those were collaborative thinking frame, completion-driven frame and unhelpful frame. Subsequently, we proceeded with independent coding for the remaining data. Finally, we integrated our coding, using Atlas.ti co-occurrence analysis. It helped us discuss to form a joint understanding of different technological frames and what kind of learning experiences they entailed.

Table 1: Illustrative example of analytical process

Quote (Student # 14)	Initial coding	Themes	Technological framing
<i>It guided well; At first there was some theory and then you had to write down and really think about and you really had to do it yourself [...]. It was nice that there were tasks throughout the discussion. It was for yourself, and you had to be able to understand the most important competencies. Then (using pen and paper) you had to write it down.</i>	Chatbot guides; chatbot is a content provider; learning as an active process; learning as a reflective process; self-paced learning; relationship with technology is positive	Relationship with chatbot is constructed positively and it enables learning by providing guidance and content that a learner must actively process	Collaborative thinking frame

4 Findings and discussion

The collaborative thinking frame refers to the way students interpret their interactions with the chatbot, leading to the development of a positive relationship. This relationship emerged as students actively constructed new understandings through dialogue with the chatbot, which provided prompts that facilitated their reflective thinking processes. For example, students felt as though the chatbot was thinking alongside them, motivating them to engage more deeply with the learning material. One student described this experience as follows: *Well, it helped me a lot because I am bad at concentrating [...] it always stopped me, and I just could not run through the material [...] especially when it asked me what I was thinking about some things (the course material) [...] it encouraged me to process it much deeper* (Student #8).

The chatbots served as scaffolding tools that enabled students to actively reconstruct their learning objectives beyond the interaction with the chatbot: *I refined my learning objectives during the chatbot and after that as well. Perhaps I had too superficial learning objectives originally for the course. Chatbot made to sharpen my personal goals (of the course).* (Student #26). More importantly, the chatbot's pedagogically formulated open questions facilitated a collaborative thinking frame by guiding discussions that included course content. This combination of questions and guidance helped students learn and encouraged them to actively reflect on their experiences in relation to the course material. One student explained this experience as follows: *When the chatbot asked the right kind of questions (about project management and teamwork) that you really had to think about—not just yes and no questions (closed questions)—it made me reflect more deeply. The*

chatbot also provided information before I had to answer those questions, which helped me understand the material better. (Student #5). Therefore, in educational contexts, students actively construct technological interpretations, which both shape and are shaped by their engagement with intended learning outcomes. A student's refinement of learning goals following chatbot engagement suggests the evolution of technological frames beyond static mental models. As Kaplan and Tripsas (2008) observe, these frames initially emerge from user assumptions but undergo modification through continued interaction. The chatbot's pedagogically structured questions provided scaffolding that enabled students to reconstruct their technological frames in ways that enhanced their educational experience. *I think it was a good system [...]so it kind of helped in that way, also in terms of learning* (Student #3).

Students actively constructed interpretations of the chatbot as a collaborative thinking partner through their technological frames (Orlikovski & Gash, 1994). The chatbot's interventions promoted engagement with learning materials. *I started to think somethings I've learned on this course [...] then I realized that maybe I want to read about (course materials) before I do the report* (Student #22). This collaborative thinking frame enabled the students to externalize their thinking processes. The chatbot's overall questioning strategy was able to facilitate reflection, demonstrating how this technological frame support metacognitive processes.

Possible factors contributing to the collaborative thinking frame could include that students with strong self-regulatory behaviour are more likely to engage in collaborative activities (Scager et al., 2017) even when using technological applications (Sharma et al., 2024). Such students may find chatbot interaction intellectually stimulating rather than disruptive to their learning process. Students with positive prior technological experiences may also be more inclined to perceive chatbot interactions as beneficial (e.g., Pesonen, 2021).

The completion driven frame characterizes how students develop an instrumental relationship with chatbots, using them primarily as information retrieval tools rather than learning partners. This frame emerged when students approached chatbots as content delivery mechanisms, focusing on extracting course-related information rather than engaging in learning processes. For example, some students recalled videos that were embedded in the chatbots, without further reflection: *There were*

videos, which told some of the course content. It told me facts throughout the discussion. (Student #21)

When another student was asked how the chatbot helped to think of what one should learn about project management, initially s/he responded and later emphasized the utility function of the chatbot: *How projects work from the beginning to the end. How and when goals should be set [...] Why budget is important, and you need to know your own skills and groups' skills [...] That piece of information (course content) was useful. At least I can concentrate on things (skills) that require most developing.* (Student #16)

In the completion driven frame, clicking to the next step was import: *(Through questions) I was able to think what I should learn about project work [...] when you clicked one option, it explained it.* (Student #18). Although the chatbots provided valuable information, the students typically approached their tasks with a focus on completion rather than genuine learning. The following quote demonstrates this: *I kind of felt as we still must write one more report for the course, where we need to comment on the chatbot we reviewed [...] It seemed something like that affected the assignment. It was clear that this might be something to keep in mind and made me definitely feel maybe this is something that matters.* (Student #7)

The completion driven frame illustrates a different kind relationship between technological frames (Orlikowski & Gash, 1994). Our analysis shows these students typically approached the chatbots with a practical, utilitarian mindset. They saw these rule-based chatbots mainly as ways to access information rather than as partners in the learning process. This completion-driven frame heavily influenced how they understood and used the technology in their studies.

Their framing revealed a perception of chatbots as offering fixed, linear paths through pre-set content—a rigid tool with set procedures rather than an adaptive system capable of responsive engagement. References to "clicking" through options highlight their mechanical understanding of the learning.

Students might adopt this completion-driven approach due to practical pressures. Time constraints and course structures encourage quick answers over deep engagement. When educational technology is framed by students as merely a content delivery system rather than a collaborative medium, their engagement naturally

follows this limiting perception. *At the beginning, I set myself the goal of passing the course, and it (chatbot) certainly helped* (Student # 4). The background in traditional education, combined with assessment methods that reward specific knowledge rather than exploration, could reinforce this task-oriented behaviour.

The unhelpful frame refers to the way students interpret their interactions with the chatbot, contributing to the development of a somewhat negative relationship. This relationship emerged when students paid attention to the technological shortcomings, and not realizing the meaningful course content or the provided learning affordances (Van Osch & Mendelson, 2011). Some students compared the chatbots with more advanced technologies, expecting more guidance as illustrated by the following quote: *I'd say it didn't help. [...] Other chatbots were much better, for example ChatGPT, which I've used quite actively when writing essays and other things. It didn't really give any feedback, which would have been useful [...] It just matches definitions to the text. So, in my opinion, you did not get any particularly good feedback and constructive criticism from it.* (Student # 33)

When students focused on the technological features, such as using “stars” to identify the most meaningful content for themselves in each section of the chatbot, they missed opportunities to reflect on the course material. The following example illustrates this: *I feel that I did not reflect much what I should do. By using stars, it was done badly. I went those sections through quickly, and you don't think about them.* (Student #19)

The unhelpful frame of chatbots was marginal in our study. In this frame, students constructed technological frames around the technological user experience, focusing on its shortcomings and comparing it to more advanced technologies. This reveals a clear incongruence (Orlikowski & Gash, 1994) between our learning design intentions and students' interpretations. While the chatbots were designed as tools for reflection to deepen engagement with course material, students primarily focused on the technological interface itself rather than the affordances (Gaver, 1991; Norman, 1999) it provided. Arguably, this attention to technological limitations overshadowed educational benefits, with students evaluating the technology rather than engaging with course concepts, leading to limited perceived learning value. These findings suggest that effective educational technology implementation requires careful alignment between technological design, pedagogical learning outcomes, and student expectations to avoid superficial engagement.

5 Conclusions

This study shows how students' technological frames (Orlikowski & Gash, 1994) function as active interpretive structures through which they process information and reflect on their learning in individualized ways (Chen et al., 2018; Li & Xiu, 2023), thereby forming learning experiences within social and technological contexts (Huhtanen, 2020). Educational technology affordances—the action possibilities they offer (Norman, 1999)—mediate how students' technological frames translate into their learning, either amplifying or limiting engagement based on alignment with their framings (Treem et al., 2015).

From a theoretical perspective, the study advances the understanding of technological frames (Orlikowski & Gash, 1994) in higher education by illustrating how these frames mediate students' learning experiences. It reinforces the idea that technology is not a neutral medium but is shaped by pedagogical contexts and user interactions (Fawns, 2022; Oliver, 2011). This study extends technological framing theory to the educational field, linking technological frames to students' learning experiences. We also provide an empirically grounded framework that helps understand student learning experiences within chatbot-supported learning environment. These findings suggest that effective educational technology implementation requires careful alignment between technological design, pedagogical learning outcomes, and student expectations to foster meaningful engagement (Li & Xue, 2023; Liu et al., 2023.) rather than superficial interaction.

In practical terms, chatbot design should aim to align closely with students' expectations and learning experiences. AI-driven chatbots need enhanced adaptability and context-awareness to foster meaningful engagement. As Treem et al. (2015) have demonstrated, individuals' experiences, resources, expectations, and beliefs about technology lead to different technological interpretations. Therefore, educational technologies should not merely serve as content delivery mechanisms but actively support collaborative and engaging learning experiences.

Despite its contributions into rule-based chatbot integration into business education, our study has generalizability limitations due to its social science focus and single-institution sample (Polit & Beck, 2010). While statistical generalization is unattainable (e.g. Lee & Baskerville, 2003), the research offers theoretical

transferability through its identified frames (Orlikowski & Gash, 1994). Future research should explore diverse institutional settings, academic disciplines, and technological contexts (including comparative analyses of rule-based versus AI-driven chatbots), employ longitudinal tracking to assess frame evolution, and address these limitations to enhance theoretical robustness (Williams, 2000) and pedagogical understanding of rule-based chatbot integration in higher education.

Acknowledgements

We want to sincerely thank our colleague Heli Bergström for her invaluable contributions and active engagement throughout the development and research process of the chatbots. Her insights and collaboration were important in shaping this paper.

References

- Adamopoulou, E., Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2, 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Al-Abdullatif, A. M. (2023). Modeling students' perceptions of chatbots in learning: Integrating technology acceptance with the value-based adoption model. *Education Sciences*, 13(11), 1151. <https://doi.org/10.3390/educsci13111151>
- Almogren, A. S., Al-Rahmi, W. M., & Dahri, N. A. (2024). Exploring factors influencing the acceptance of ChatGPT in higher education: A smart education perspective. *Heliyon*, 10(11), e31887. <https://doi.org/10.1016/j.heliyon.2024.e31887>
- Anderson, T. (2016). Theories for learning with emerging technologies. In G. Veletsianos (Ed.), *Emergence and Innovation in Digital Learning: Foundations and Applications* (pp. 35-50). Edmonton, AB: Athabasca University Press.
- Ayanwale, M. A., Ndlovu, M. (2024). Investigating factors of students' behavioral intentions to adopt chatbot technologies in higher education: Perspective from expanded diffusion theory of innovation. *Computers in Human Behavior Reports*, 14, 100396. <https://doi.org/10.1016/j.chbr.2024.100396>
- Bijker, W. E. (2001). Understanding technological culture through a constructivist view of science, technology, and society. In S. H. Cutcliffe & C. Mitcham (Eds.), *Visions of STS. Counterpoints in science, technology, and society studies* (pp. 19–34).
- Bjørn, P., Scupola, A. & Fitzgerald, B. (2007) Expanded technological frames towards mediated collaboration: groupware adoption in virtual learning teams, *Scandinavian Journal of Information Systems*, Vol 18, No, 2, pp. 28-47
- Braun, V., Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Carlsson, L. (2023). Strategizing organizational capabilities for industrial digitalization – exploring managers' technological frames. *Journal of Manufacturing Technology Management*, 34(9), 20–39. <https://doi.org/10.1108/JMTM-07-2022-0252>
- Chang, D.H., Lin, M.P.-C., Hajian, S., & Wang, Q.Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability*, 15, 12921. <https://doi.org/10.3390/su151712921>
- Chen, J., Wang, M., Kirschner, P. A., & Tsai, C.-C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research*, 88(6), 799–843. <https://doi.org/10.3102/0034654318791584>

- Conole, G., Dyke, M. (2004). What are the affordances of information and communication technologies? 12(2). <https://doi.org/10.1080/0968776042000216183>
- Davidson, E., Pai, D. (2004). Making sense of technological frames: Promise, progress, and potential. In B. Kaplan, D. P. Truex, D. Wastell, A. T., Wood-Harper, & J. I. DeGross (Eds.), *Information Systems Research*. IFIP International Federation for Information Processing, https://doi.org/https://doi.org/10.1007/1-4020-8095-6_26
- Drew, C., Mann, A. (2018). Unfitting, uncomfortable, unacademic: a sociological reading of an interactive mobile phone app in university lectures. *International Journal of Educational Technology in Higher Education*, 15(43). <https://doi.org/10.1186/s41239-018-0125-y>
- Ellis, R., Goodyear, P. (2010). Students' experiences of e-learning in higher education: The ecology of sustainable innovation. Routledge.
- Fawns, T. (2022). An entangled pedagogy: Looking beyond the pedagogy-technology dichotomy introduction: Moving past the technology-pedagogy dichotomy. *Postdigital Science and Education*. <https://doi.org/10.1007/s42438-022-00302-7>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fryer, L. K., Nakao, K., & Thompson, A. (2018). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, 93, 279–289. <https://doi.org/10.1016/j.chb.2018.12.023>
- Følstad, A., Skjuve, M., & Brandtzaeg, P. B. (2019). Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11551 LNCS, 145–156. https://doi.org/10.1007/978-3-030-17705-8_13
- Gaver, W. W. (1991). Technology affordances. In *Conference on Human Factors in Computing Systems - Proceedings* (pp. 79-84) <https://doi.org/10.1145/108844.108856>
- Gibson, J. J. (2014). *The ecological approach to visual perception*. Houghton Mifflin.
- Goodyear, P., Carvalho, L., & Yeoman, P. (2021). Activity-centred analysis and design (ACAD): Core purposes, distinctive qualities and current developments. *Educational Technology Research and Development*, 69(2), 445-464. <https://doi.org/10.1007/s11423-020-09926-7>
- Goodyear, P., Dimitriadis, Y. (2013). In medias res: Reframing design for learning. In *Research in Learning Technology* Vol. 21, 1. Association for Learning Technology. <https://doi.org/10.3402/rlt.v21i0.19909>
- Güldal, H., Dinçer, E. O. (2024). Can rule-based educational chatbots be an acceptable alternative for students in higher education? *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12977-5>
- Hew, K. F., Huang, W., Du, J., & Jia, C. (2023). Using chatbots to support student goal setting and social presence in fully online activities: Learner engagement and perceptions. *Journal of Computing in Higher Education*, 35(1), 40-68. <https://doi.org/10.1007/s12528-022-09338-x>
- Huhtanen, A. (2019) Verkko-oppimisen muotoilukirja & Oppimismuotoilun työkalupakki [The design book for online learning]. Avointen oppimateriaalien kirjasto. https://urn.fi/urn:nbn:fi:oerfi-202209_00022860_4
- Jahnke, I. (2023). Quality of digital learning experiences – effective, efficient, and appealing designs? *The International Journal of Information and Learning Technology*, 40(1), 17–30. <https://doi.org/10.1108/IJILT-05-2022-0105>
- Kaplan, S., Tripsas, M. (2008). Thinking about technology: Applying a cognitive lens to technical change. *Research Policy*, 37, 790–805. <https://doi.org/10.1016/j.respol.2008.02.002>
- Kim, S., Lee, J., & Gweon, G. (2019). Comparing data from chatbot and web surveys: Effects of platform and conversational style on survey response quality. In *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Paper No. 86, pp. 1-12). Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300316>

- Kirkwood, A., & Price, L. (2013). Technology-enhanced learning and teaching in higher education: What is 'enhanced' and how do we know? A critical literature review. *Learning, Media and Technology*, 39(1). <https://doi.org/10.1080/17439884.2013.770404>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1), 56. <https://doi.org/10.1186/s41239-023-00426-1>
- Lee, A. S., Baskerville, R. L. (2003). Generalizing generalizability in information systems research. *Information Systems Research*, 14(3), 221–243. <https://doi.org/10.1287/isre.14.3.221.16560>
- Lee, J. H., Yang, H., Shin, D., & Kim, H. (2020). Chatbots. *ELT Journal*, 74(3), 338–344. <https://doi.org/10.1093/elt/ccaa035>
- Li, J., Xue, E. (2023). Dynamic interaction between student learning behaviour and learning environment: Meta-analysis of student engagement and its influencing factors. *Behavioural Sciences*, 13(1), 59. <https://doi.org/10.3390/bs13010059>
- Lin, M. P.-C., Chang, D. (2023). CHAT-ACTS: A pedagogical framework for personalized chatbot to enhance active learning and self-regulated learning. *Computers and Education: Artificial Intelligence*, 5, Article 100167. <https://doi.org/10.1016/j.caeai.2023.100167>
- Lin, A., Silva, L. (2005). The social and political construction of technological frames. *European Journal of Information Systems*, 14(1), 49–59. <https://doi.org/10.1057/palgrave.ejis.3000521>
- Lindebaum, D. (2024). Management learning and education as “big picture” social science. *Academy of Management Learning & Education*, 23(1), 1–7. <https://doi.org/10.5465/amle.2023.0173>
- Liu, Y., Morard, S., Adinda, D., Sanchez, E., Trestini, M., Trestini, M. A., & Liu, Y.-D. (2023). A Systematic Review: Criteria and Dimensions of Learning Experience. <https://hal.science/hal-04264325v1>
- Malik, R., Shrama, A., Trivedi, S., & Mishra, R. (2021). Adoption of chatbots for learning among university students: Role of perceived convenience and enhanced performance. *International Journal of Emerging Technologies in Learning (IJET)*, 16(18), 200. <https://doi.org/10.3991/ijet.v16i18.24315>
- Marell-Olsson, E., Jahnke, I. (2019). Wearable technology in a dentistry study program: Potential and challenges of smart glasses for learning at the workplace. In I. Buchem, R. Klamma, & F. Wild (Eds.), *Perspectives on wearable enhanced learning (WELL)* (pp. 433–451). Springer.
- Norman, D. A. (1999). Affordance, conventions, and design. *interactions*, 6(3). <https://doi.org/10.1145/301153.301168>
- Okonkwo, C. W., Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Oliver, M. (2011). Technological determinism in educational technology research: Some alternative ways of thinking about the relationship between learning and technology. *Journal of Computer Assisted Learning*, 27(5), 373–384. <https://doi.org/10.1111/j.1365-2729.2011.00406.x>
- Oliver, M. (2013). Learning technology: Theorising the tools we study. *British Journal of Educational Technology*, 44(1), 31–43. <https://doi.org/10.1111/j.1467-8535.2011.01283.x>
- Orlikowski, W. J. & Gash, D. C. (1994). Technological frames: Making sense of information technology in organizations. *ACM Transactions on Information Systems*, 12(2), 174–207. <https://doi.org/10.1145/196734.196745>
- Papakostas, C., Troussas, C., Krouska, A., & Sgouropoulou, C. (2024). A rule-based chatbot offering personalized guidance in computer programming education. In A. Sifaleras & F. Lin (Eds.), *Generative intelligence and intelligent tutoring systems. ITS 2024. Lecture Notes in*

- Computer Science (Vol. 14799). Springer, Cham. https://doi.org/10.1007/978-3-031-63031-6_22
- Pesonen, J.A. (2021). 'Are you OK?' Students' trust in a chatbot providing support opportunities. In Zaphiris, P., Ioannou, A. (Eds.), *Learning and collaboration technologies: Games and virtual environments for learning. HCII 2021. Lecture Notes in Computer Science* vol 12785. Springer, Cham. https://doi.org/10.1007/978-3-030-77943-6_13
- Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies*, 47(11), 1451-1458. <https://doi.org/10.1016/j.ijnurstu.2010.06.00>
- Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., & Lazzara, E. H. (2015). Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human Resource Management*, 54(4), 599-622. <https://doi.org/10.1002/hrm.21628>
- Scager, K., Boonstra, J., Peeters, T., Vulperhorst, J., & Wiegant, F. (2017). Collaborative learning in higher education: Evoking positive interdependence. *CBE—Life Sciences Education*, 15(4). <https://doi.org/10.1187/cbe.16-07-0219>
- Schraw, G., Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460-475. <https://doi.org/10.1006/CEPS.1994.1033>
- Selwyn, N. (2012). Making sense of young people, education and digital technology: The role of sociological theory. *Oxford Review of Education*, 38(1), 81-96. <https://doi.org/10.1080/03054985.2011.577949>
- Sharma, K., Nguyen, A., & Hong, Y. (2024). Self-regulation and shared regulation in collaborative learning in adaptive digital learning environments: A systematic review of empirical studies. *British Journal of Educational Technology*, 55(4), 1398-1436. <https://doi.org/10.1111/bjet.13459>
- Shearer, R. L., Aldemir, T., Hitchcock, J., Resig, J., Driver, J., & Kohler, M. (2019). What students want: A vision of a future online learning experience grounded in distance education theory. *American Journal of Distance Education*. <https://doi.org/10.1080/08923647.2019.1706019>
- Siewiorek, A., Saarinen, E., Lainema, T., & Lehtinen, E. (2012). Learning leadership skills in a simulated business environment. *Computers and Education*, 58(1), 121-135. <https://doi.org/10.1016/j.compedu.2011.08.016>
- Silverstone, R. (2006). Domesticating domestication. Reflections on the life of a concept. In T. Berker, M. Hartmann, Y. Punie, & K. J. Ward (Eds.), *Domestication of Media and Technology* (pp. 229-248). Open University Press.
- Spieth, P., Röth, T., Clauss, T., & Klos, C. (2021). Technological frames in the digital age: Theory, measurement instrument, and future research areas. *Journal of Management Studies*, 58(7), 1962-1993. <https://doi.org/10.1111/joms.12720>
- Squire, K., Dijkers, S. (2012). Amplifications of learning: Use of mobile media devices among youth. *Convergence*, 18(4), 445-464. <https://doi.org/10.1177/1354856511429646>
- Stöhr, C., Ou, A. W., & Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers and Education: Artificial Intelligence*, 7, 100259. <https://doi.org/10.1016/j.caeai.2024.100259>
- Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*, 49(2), 223-245. <https://doi.org/10.1007/s10755-023-09686-1>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Treem, J. W., Dailey, S. L., Pierce, C. S., & Leonardi, P. M. (2015). Bringing technological frames to work: How previous experience with social media shapes the technology's meaning in an

- organization. *Journal of Communication*, 65(2), 396–422.
<https://doi.org/10.1111/jcom.12149>
- Vanichvasin, P. (2021). Chatbot development as a digital learning tool to increase students' research knowledge. *International Education Studies*, 14(2), 44–53.
<https://doi.org/10.5539/ies.v14n2p44>
- Van Osch, W., Mendelson, O. (2011). A. Typology of affordances: Untangling sociomaterial interactions through video analysis. *Proceedings of the International Conference on Information Systems (ICIS)*, Shanghai, China, December.
- Williams, M. (2000). Interpretivism and generalisation. *Sociology*, 34(2), 209–224.
<https://doi.org/10.1177/S0038038500000146>
- Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. *Academy of Management Proceedings*, 2018(1).
<https://doi.org/10.5465/AMBPP.2018.15903abstract>
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet? A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, 4, 654924. <https://doi.org/10.25656/01:22886>
- Yin, J., Goh, T.-T., Yang, B., & Xiaobin, Y. (2020). Conversation Technology With Micro-Learning: The Impact of Chatbot-Based Learning on Students' Learning Motivation and Performance. *Journal of Educational Computing Research*, 59(1), 154–177.

VALUE CO-CREATION THROUGH DIGITALISATION: FROM A MICRO- LEVEL TO A MESO-LEVEL APPROACH

YIRAN CHEN ZHOU, ANASTASIA TSVETKOVA

Åbo Akademi University, Faculty of Science and Engineering, Laboratory of Industrial
Management, Turku, Finland
yiran.chen@abo.fi, anastasia.tsvetkova@abo.fi

This study identifies micro-to-meso level hindrances for value co-creation through digitalisation at port operations. Digitalisation in ports faces significant challenges due to complex, interdependent actor relationships. This study explores how micro-level hindrances affect meso-level value co-creation in a port digitalisation context. Based on an explorative case study of a RoPax port digitalisation project, we identified hindrances across levels could constrain and interdepend on each other. Micro-level hindrances within an organisation, such as – lack of digitalisation leads, limited operational data, outdated digital infrastructure and security concerns – could reinforce six hindrances on meso-level: 1) unaware co-creators, 2) lack of collective leadership for digitalisation, 3) limited data accessibility, 4) low inter-organisational resource compatibility, 5) misaligned or misunderstood value co-creation targets, 6) resistant to change.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.12](https://doi.org/10.18690/um.fov.4.2025.12)

ISBN

978-961-286-998-4

Keywords:

value co-creation,
digitalisation,
service-dominant logic,
meso-level approach,
micro-level hindrances



University of Maribor Press

1 Introduction

With technological advancements, inter-organisational value co-creation has become more complex. Research on value co-creation enabled by information technology calls for going beyond dyadic relationships; co-creators range from customer and user, to digital solution provider, smart objects and partners (Zhang et al., 2024). This study follows Burton et al. (2024), focusing on two levels for value co-creation: micro- and meso-levels. Micro-level refers to the context where an actor, such as an organisation or company, can fully control the context and resources within its organisation, co-creating value focusing on its customers' needs. Meso-level goes beyond organisational boundaries, involving numerous direct value co-creating dyads between two actors, because an actor can be in several dyadic relationships simultaneously. Given the distinction, we address the research question: what micro-level hindrances affect value co-creation through digitalisation at the meso-level?

Digital technologies decentralise business roles (Vial, 2019), and value co-creation requires the collective coordination of multiple actors. However, studies on digitalisation's impact on value co-creation remained on specific activities, capabilities and resources (Sjödin et al., 2020; Tuunanen et al., 2023), or a single actor (Wieczerzycki et al., 2024). Eggert et al. (2019) encouraged broader examination at the collective level to create superior value. Service-dominant logic (SDL) positions co-created value as the research centre rather than specific actors or exchanged services (Vargo et al., 2020), and value co-creation could happen on diverse levels and for several actors simultaneously.

Transportation drives business activity concentration and significantly influences global trading (Rodrigue, 2024, p. 14). Maritime transportation, in contrast, has been slow in adopting digital technologies and innovations, such as shipping companies and ports are still working in silos within their organisations, remaining on micro-level value co-creation and slow in decision-making (Raza et al., 2023; Sanchez-Gonzalez et al., 2019). As transportation hubs, ports can significantly contribute to maritime transport efficiency and decarbonisation (Zis et al., 2020). However, port digitalisation involves multiple co-creators, resulting in a complex value co-creation process (Haraldson et al., 2021). Advancing port digitalisation requires examining inter-organisational relationships, deepening meso-level understandings and considering further value co-creators in addition to shipping companies. Other

actors, such as the municipality, road transportation companies, and digital solution providers, are also critical for port development. Hence, SDL provides a suitable theoretical lens, as it decentralises inter-organisational service-for-service exchange, where value emerges through multiple-actor interaction (Vargo et al., 2024).

Port digitalisation research explored large container ports, and smaller ports with different traffic profiles may uncover deeper insights into value co-creation mechanisms (Chen et al., 2024). These ports may also reveal unique value co-creation dynamics due to their distinctive organisational setups and localised challenges. Shifting beyond the dyadic firm-customer relationship in value co-creation could enable further digitalisation in ports. Meso-level value co-creation in transport is also becoming recognised as essential for sustainable development (Wieczerzycki et al., 2024). Therefore, we argue that a meso-level approach to value co-creation is necessary to understand critical activities for efficient digitalisation.

An exploratory case study on a RoPax port digitalisation project was designed to address the research question. RoPax ports handle roll-on and roll-off vessels that carry both vehicles and passengers. Our analysis focuses on shifting from micro-level to meso-level value co-creation through participating companies in the focal project, analysing their value co-creation processes and challenges. In total, we identified six meso-level hindrances affected by four micro-level hindrances.

The next section reviews the literature on value co-creation and its relationship with digitalisation. Section 3 details the research design. Section 4 presents the findings and discussion. Research limitations and future research suggestions are discussed in the conclusion.

2 Literature Review

2.1 Contexts for Value

Value is a subjective, relational, and contextual concept (Chandler & Vargo, 2011), formed by perceivers and their relationship with other actors (Eggert et al., 2019). In service-dominant logic (SDL), consumers are also relevant value co-creators (Vargo & Lusch, 2008). This study focuses on the value-in-use due to the nature of the research question, as we aim to understand the phenomenon from a meso-level

and multi-actor perspective. Value-in-use stands for the outcomes of an activity that either facilitates or hinders achieving a goal (Kleinaltenkamp et al., 2022; Macdonald et al., 2016). Context understanding is essential for value-in-use, which consists of "*a set of unique actors with unique reciprocal links among them*" (Chandler & Vargo, 2011, p. 40). Kleinaltenkamp et al. (2022) remarked that value-in-use could serve an individual company or organisation (micro-level) or a collective of them (meso-level).

SDL emphasises that value is co-created by multiple actors rather than a single actor. Value co-creation has its own narrative and process, value co-creators may also be unaware of each other when multiple co-creators are involved (Vargo & Lusch, 2016). Non-participative co-creators may emerge due to unclear expectations, scarce information, or lack of trust (Järvi et al., 2018). Hence, actor engagement is critical for value co-creation, which starts with awareness of co-creators and resource integration (Storbacka et al., 2012). Value co-creation occurs at various aggregation levels (Vargo et al., 2020), defined by the number of involved actors (dyadic or triadic) or the nature of relations, direct or indirect (Burton et al., 2024; Chandler & Vargo, 2011). Understanding the role of other actors in multi-actor configuration is another SDL application (Vargo et al., 2020). Storbacka et al. (2012) proposed a three-phase process for meso-level value co-creation: origination, mobilisation and stabilisation. They also strived that a focal actor should consider micro- and meso-level configurations for purposeful value co-creation.

SDL scholars defined the "micro-level" of value co-creation differently. Chandler & Vargo (2011) referred to the actors involved in exchange activities, where they could control the context and resources for a specific activity. In contrast, Burton et al. (2024) framed their definition within digital service innovation. They conceptualised micro-level within an organisation's departments or business units, where a focal firm has full control over activities and resources. Despite the micro-level within an organisation, we consider the customer or digital solution user to be another value co-creator on this level, and only the provider-customer relationship is focused on this level. Based on the research question, which is related to digitalisation hindrances across micro- and meso-levels. The current study follows Burton et al (2024)'s definition for a clear distinction.

Due to the complex actor relationships in ports, hindrances to value co-creation are diverse. Actor engagement is challenging when interdependent operations require collaboration among multiple, sometimes competing actors (Raza et al., 2023). Hindrances can reinforce each other, creating compounding effects that continuously slow port digitalisation. Thus, identifying and analysing individual hindrances to digitalisation is insufficient for our research question. Instead, hindrances must be contextualised within the co-creation processes.

2.2 Digitalisation and Its Impacts

Digitalisation is the application of digital technologies to society or work (Vial, 2019), which serves as a basis for enabling new ways of working and collaboration (Gong & Ribiere, 2021). The ultimate driver for digitalisation is business performance enhancement and value co-creation. Hence, digitalisation should evolve alongside business development (Porter & Heppelmann, 2014).

Value (co-)creation through digitalisation in an organisation or among two organisations was previously studied (Martinez, 2019; Wiczerzycki et al., 2024). Nevertheless, research on the multilevel impacts of digitalisation-enabled value co-creation is scarce (Zhang et al., 2024). Organisations introduce digital technologies at different paces (Martinez, 2019). By scaling digitalisation across multiple organisations, inter-organisational coordination becomes critical. Empirical studies on digitalisation's impacts on value co-creation are available across diverse fields, including manufacturing (Martinez, 2019), logistics (Michel et al., 2023), and highly traditional food production (Hauke-Lopes et al., 2023).

RoPax shipping gained limited attention compared to other ports, such as container ports (Du et al., 2023). Digitalisation hindrances and drivers were studied in ports (Brunila et al., 2021; Inkinen et al., 2019, 2021) and maritime transport (Sanchez-Gonzalez et al., 2019; Tijan et al., 2021). Common hindrances are: 1) organisational resistance to change (Gerlitz & Meyer, 2021); 2) security concerns (de Langen, 2021; Tijan et al., 2021); 3) outdated management perspectives (Tsvetkova et al., 2021); 4) technical limitations; and 5) resource scarcity. Each port is a unique self-organising system (Watson et al., 2021). Nevertheless, value co-creation through port digitalisation remains unclear, which is the goal of this study.

2.3 Theoretical Framework

This paper positions on the micro-to-meso-level value co-creation, focusing on two aspects: 1) hindrances on both levels, and 2) their interdependences. Drawing from the literature, we categorise hindrances based on the three-phase process of meso-level value co-creation: origination, mobilisation and stabilisation (Storbacka et al., 2012). The definitions of these phases are presented in the following table.

Table 1: Three-phase Meso-level Value Co-creation Process
Adopted from Storbacka et al. (2012)

Origination	Mobilisation	Stabilisation
Invention or introduction of a new value co-creation process, which would alter the meso-level value co-creation structure.	Change management process, where value co-creators would communicate, prove and understand potential barriers to change, and resources would be mobilised for the target value co-creation.	Establishment and acceptance of new co-creation practices would become the dominant logic on a meso-level.

Figure 1 illustrates the research focus on interdependent hindrances. In the figure, actors A and D are active and engaged co-creators with resource contributions. Actors B and C are identified as co-creators but non-participative due to specific hindrances. These hindrances may be interdependent and reinforcing, further preventing these actors from participating in value co-creation.

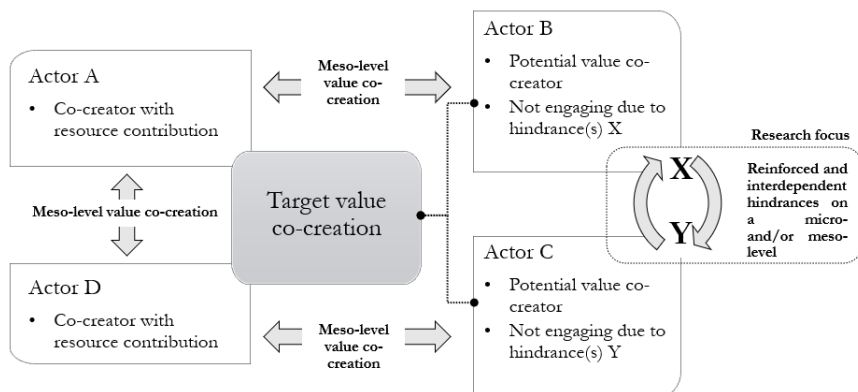


Figure 1: Reinforced and interdependent hindrances

Source: Own

3 Methodology

This study investigates hindrances and their interdependencies for port digitalisation from a value co-creation lens. Given the limited theoretical and empirical foundations, an exploratory case study was designed. Previous studies on port business changes due to technological drivers also applied this research design (Henríquez et al., 2022). An exploratory study provides first-hand insights into emerging empirical phenomena (Swedberg, 2020). Furthermore, we adopted an abductive approach, iterating between theoretical frameworks and empirical observations to refine our understanding (Dubois & Gadde, 2014). An ongoing, dynamic interaction between theory and fieldwork characterises this research.

A RoPax port upgrading project in the Baltic Region was selected for the case study. The port is situated near the urban area, where RoPax ships are the primary means for the locality to connect with international markets. Low traffic intermodality and high operational sensitivity result in a slow digitalisation pace. The project aimed to improve operational efficiency, reducing environmental and social impacts through digitalisation. Project participants included both incumbent port operations actors in the focal port, such as the port authority, the company that manages the port area, and new entrants (digital solution providers from other industries). Adopting the research question, the units of analysis are value co-creation targets that involve more than two actors, and each actor is at least active in one dyadic value co-creation relationship.

Data collection and analysis were continuously refined and developed in conjunction with fieldwork and theoretical framework development. The process spanned from April 2021 to April 2023, involving semi-structured interviews with project participants and field visits as primary data. Secondary data was sourced from project documentation (project proposals, reports, and presentations). Secondary data provided insight into project participants' businesses and interview structure formulation.

Interview questions included digitalisation opportunities and challenges, and at least two managerial-level representatives from each project participating company were interviewed. Interviewees were selected based on the relevance of their expertise and their involvement within their organisation in the port upgrading project. 13

interviews were conducted, reaching out to 20 individual interviewees. Table 2 presents detailed interviewees' roles. The interviews lasted 60 to 90 minutes and were recorded, transcribed, and documented. Three field visits to the terminals, led by shipping companies, resulted in observation documents.

First, we identified value co-creation targets, such as digitalisation benefits for port operations and individual firms. Next, we pinpointed hindrances that participants faced in advancing digitalisation, which were classified based on the theoretical framework, by answering: 1) to what level does this hindrance belong? , and 2) to what phase of meso-level value (Table 1) does it correspond? Finally, we examined interdependencies between hindrances. To ensure data validity, all project researchers reviewed and cross-checked interview transcriptions. Interviewees also reviewed synthesised memos to prevent misinterpretation. Findings, particularly on port digitalisation hindrances, were compiled into project reports, which participants reviewed, serving as a validation instrument to ensure an unbiased analysis.

Table 2: Interview Data

Company type	Interviewees' role	N1*	N2**
Port authority	Technical Director, COO, and IT Manager	3	1
	Sales and deputy managing director	1	1
Shipping companies	IT Manager A (Interview and site visit)	1	1
	IT Manager A, Captain, and Cargo planner (Interview and site visit)	3	1
	Terminal manager	1	1
	Sales manager	1	1
	Operation manager (Interview and site visit)	1	1
Digital solution providers	Research leaders	2	1
	CEO and project manager	2	1
	CEO	1	1
	CEO and Research leader	2	1
	Director and Research leader	2	1
	Research leader	1	1
Total of different interviewees		20	13

N1*: Interviewee number per session; N2**: Interview session number.

4 Findings and Discussion

Following the theoretical framework and based on the research data, we identified six meso-level hindrances that are related to four micro-level hindrances for value co-creation in port digitalisation. A description of each meso-level digitalisation hindrance can be found in Table 3.

During the interview, micro-level hindrances were often mentioned together with meso-level hindrances: some are interdependent, while others remain on a micro-level. Hence, to avoid misleading, we present micro-level hindrances relevant for meso-level value co-creation in the current section. These are:

1. Lack of digitalisation leads: digitalisation often being treated as short-term projects without dedicated personnel overseeing long-term development. This was mentioned by incumbent port operation-related actors, such as port and shipping companies.
2. Limited operational data: some port operations remain analogue, and critical environment data for port digitalisation has never been collected.
3. Outdated digital infrastructure: each organisation has their own digitalisation pace. Some are reluctant or do not perceive sufficient value in further digital infrastructure investment.
4. Digital security concerns: fears of cybersecurity threats or risks introduced by digitalisation or new digital solutions.

The cross-analysis of interviews provided the tensions between the micro-level and meso-level hindrances. For instance, a new entrant digital solution provider remarked:

"We do not have a clear picture of this business yet and have joined this project to understand better what we can offer and to whom in particular. [...] New data sources are needed as some data is not available (e.g. parking situation at the port area)."

In this case, the meso-level hindrances are limited data accessibility (H3) and low inter-organisational resource compatibility (H4) because they understand the technical part of the solutions but lack knowledge of the meso-level setup. The

micro-level hindrances would be a lack of operational data (at the port), which was confirmed by analysing the port authority's interview.

Table 3: Meso-level Value Co-creation Hindrances

Phases	Hindrances and description	Description
Origination	– Unaware co-creators [H1]	Co-creators are unaware of value co-creation opportunities. This may be caused due a lack of digital knowledge in an organisation.
	– Lack of collective leadership for digitalisation [H2]	Multi-actor collaboration requires collective leadership from several actors, but they are often reluctant and unmotivated.
Mobilisation	– Limited data accessibility [H3]	Data may exist but not be sharable or usable due to security concerns or incompatible inter-organisational systems.
	– Low inter-organisational resource compatibility [H4]	Digitalisation requires new resources, such as knowledge, business intelligence or enabling technology. The resource combination between different co-creators may not be compatible or result in an efficient resource combination.
	– Misaligned or misunderstood value co-creation targets [H5]	Co-creators do not have aligned value targets or are unable to communicate efficiently.
Stabilisation	– Resistant to change [H6]	The co-created value could be implemented, but larger-scaled co-creators resist change. This could be that they demand further social value co-creation or more aligned interests.

To understand micro-to-meso level hindrance relationships, we further mapped the tension in Table 4. The first column presents meso-level hindrances, and the first row includes the micro-level hindrances. The findings revealed two micro-level hindrances that impact almost all the meso-level aspects: lack of digitalisation leads and security concerns.

As RoPax port digitalisation is still emerging, many co-creators focus on short-term digitalisation projects, working without a strategic digitalisation plan. This was common at the port and shipping companies. Some started to create new positions to adapt to the technological changes – a digitalisation manager – who strategically oversees all the digitalisation-related collaborations. Digital security concerns were also commonly mentioned in the literature (Tijan et al., 2021). Because RoPax ports involve both passenger and vehicle traffic, concerns regarding privacy regulations such as GDPR and image protection of the passengers were mentioned. A balanced

consideration of customer privacy, safety, and efficiency should be carefully designed.

Table 4: Micro-to-meso Level Hindrance Relationships

Meso-level hindrances	Micro-level hindrances				
		Lack of digitalisation leads	Limited operational data	Outdated digitalisation infrastructure	Digital security concerns
	Unaware co-creators [H1]	X			
	Lack of collective leadership for digitalisation [H2]	X		X	X
	Low data accessibility [H3]		X		X
	Low inter-organisational resource compatibility [H4]		X	X	X
	Misaligned or misunderstood value co-creation targets [H5]	X		X	X
	Resistant to change [H6]	X			X

The most independent micro-level hindrance would be limited operational data. Since many may already have established data collection systems, the more challenging aspect is how to exploit value through the collected data and enable data treatment for business intelligence. Nevertheless, if the digital infrastructure is outdated, the actor may have to upgrade first, then start co-creating with other actors. This observation seconds with Storbacka et al. (2012), that the focal actor should consider hindrances from both levels for co-creating value.

Micro-level hindrances, such as lack of digitalisation leads, outdated digital infrastructure and security concerns, were broadly mentioned in the literature (see Section 2.2). These are closely linked to institutional setups, where actors should collaborate and jointly overcome problems like institutional inertia and insufficient leadership support (Tijan et al., 2021).

5 Conclusion

This study explores the interplay among hindrances in value co-creation through digitalisation and their implications for theory and management. Theoretically, the tensions between various port digitalisation hindrances were revealed (Brunila et al., 2021; Inkinen et al., 2019, 2021). By adopting an SDL perspective (Vargo & Lusch,

2016) and phases of meso-level value co-creation (Storbacka et al., 2012), we could understand these hindrances contextually and decentralisedly. Our findings indicate the need for continued research on effective strategies to foster digitalisation in traditionally slow-to-digitalise industries. Managerial implications include addressing these hindrances contextually, for example, by encouraging digital leadership coalitions across the port development community, engaging governments and other relevant stakeholders to incentivise data sharing. Promoting shared platforms and standardised protocols can also aid in this effort. Overall, clear communication and coordinated efforts are essential for overcoming obstacles and facilitating accelerated digitalisation.

This study has several limitations. First, the ongoing nature of the port upgrade in this case means that the value co-creation outcomes are unknown. Further research considering the realised value of co-creation outcomes would be valuable. Second, difficulties in engaging organisations outside the project limit the completeness of the case study analysis, suggesting a need for further investigation into non-participative co-creators. Last, we question the generalisability of the study due to the uniqueness of each port. We encourage future research to consider these aspects and further study the research topic within the context of port operation.

References

- Brunila, O. P., Kunnaala-Hyrkki, V., & Inkinen, T. (2021). Hindrances in port digitalization? Identifying problems in adoption and implementation. *European Transport Research Review*, 13(1).
- Burton, J., Story, V. M., Zolkiewski, J., & Nisha, N. (2024). Digital service innovation challenges faced during servitization: a multi-level perspective. In *Journal of Service Management* (Vol. 35, Issue 2, pp. 202–226). Emerald Publishing.
- Chandler, J. D., & Vargo, S. L. (2011). Contextualization and value-in-context: How context frames exchange. *Marketing Theory*, 11(1), 35–49.
- Chen, Y., Tsvetkova, A., Edelman, K., Wahlström, I., Heikkilä, M., & Hellström, M. (2024). How does Digitalisation Transform Business Models in Ropax Ports? A Multi-Site Study of Port Authorities. *Journal of International Technology and Information Management*, 32(1), 162–189.
- de Langen, P. W. (2021). Seaports as Clusters of Economic Activities. In *International Encyclopedia of Transportation* (pp. 310–315). Elsevier.
- Du, Y., Li, C., Wang, T., & Xu, Y. (2023). Special issue on "Smart port and shipping operations" in Maritime Policy & Management. In *Maritime Policy and Management* (Vol. 50, Issue 4, pp. 413–414). Routledge.
- Dubois, A., & Gadde, L. E. (2014). "Systematic combining"-A decade later. *Journal of Business Research*, 67(6), 1277–1284.
- Eggert, A., Kleinaltenkamp, M., & Kashyap, V. (2019). Mapping value in business markets: An integrative framework. *Industrial Marketing Management*, 79, 13–20.

- Gerlitz, L., & Meyer, C. (2021). Small and Medium-Sized Ports in the TEN-T Network and Nexus of Europe's Twin Transition: The Way towards Sustainable and Digital Port Service Ecosystems. *Sustainability (Switzerland)*, 13(8), 4386.
- Gong, C., & Ribiere, V. (2021). Developing a unified definition of digital transformation. *Technovation*, 102.
- Haraldson, S., Lind, M., Breitenbach, S., Croston, J. C., Karlsson, M., & Hirt, G. (2021). The Port as a Set of Socio-technical Systems: A Multi-organisational View. In *Maritime Informatics* (pp. 47–63).
- Hauke-Lopes, A., Ratajczak-Mrozek, M., & Wiczerzycki, M. (2023). Value co-creation and co-destruction in the digital transformation of highly traditional companies. *Journal of Business and Industrial Marketing*, 38(6), 1316–1331.
- Henríquez, R., Martínez de Osés, F. X., & Martínez Marín, J. E. (2022). Technological drivers of seaports' business model innovation: An exploratory case study on the port of Barcelona. *Research in Transportation Business & Management*, 43, 100803.
- Inkinen, T., Helminen, R., & Saarikoski, J. (2019). Port digitalization with open data: Challenges, opportunities, and integrations. *Journal of Open Innovation: Technology, Market, and Complexity*, 5(2).
- Inkinen, T., Helminen, R., & Saarikoski, J. (2021). Technological trajectories and scenarios in seaport digitalization. *Research in Transportation Business and Management*, 41, 2210–5395.
- Järvi, H., Kähkönen, A. K., & Torvinen, H. (2018). When value co-creation fails: Reasons that lead to value co-destruction. *Scandinavian Journal of Management*, 34(1), 63–77.
- Kleinaltenkamp, M., Eggert, A., Kashyap, V., & Ulaga, W. (2022). Rethinking customer-perceived value in business markets from an organizational perspective. *Journal of Inter-Organizational Relationships*, 28(1–2), 1–18.
- Macdonald, E. K., Kleinaltenkamp, M., & Wilson, H. N. (2016). How business customers judge solutions: Solution quality and value in use. *Journal of Marketing*, 80(3), 96–120.
- Martinez, F. (2019). Process excellence the key for digitalisation. *Business Process Management Journal*, 25(7), 1716–1733.
- Michel, S., Bootz, J. P., & Bessouat, J. (2023). Possible futures of crowd logistics for manufacturers: results of a strategic foresight study. *Journal of Business and Industrial Marketing*, 38(10), 2019–2029.
- Porter, M. E., & Heppelmann, J. E. (2014). How Smart, Connected Products Are Transforming Competition. *Harvard Business Review*, 92(11), 64.
- Raza, Z., Woxenius, J., Vural, C. A., & Lind, M. (2023). Digital transformation of maritime logistics: Exploring trends in the liner shipping segment. *Computers in Industry*, 145.
- Rodrigue, J.-P. (2024). *The Geography of Transport Systems* (6th ed.). Routledge.
<http://people.hofstra.edu/geotrans>.
- Sanchez-Gonzalez, P. L., Díaz-Gutiérrez, D., Leo, T. J., & Núñez-Rivas, L. R. (2019). Toward digitalization of maritime transport? *Sensors*, 19(4).
- Sjödin, D., Parida, V., Kohtamäki, M., & Wincent, J. (2020). An agile co-creation process for digital servitization: A micro-service innovation approach. *Journal of Business Research*, 112, 478–491.
- Storbacka, K., Frow, P., Nenonen, S., & Payne, A. (2012). Designing business models for value co-creation. *Review of Marketing Research*, 9, 51–78.
- Swedberg, R. (2020). Exploratory research. *The Production of Knowledge: Enhancing Progress in Social Science*, 2(1), 17–41.
- Tijian, E., Jović, M., Aksentijević, S., & Pucihar, A. (2021). Digital transformation in the maritime transport sector, Technological Forecasting and Social Change, 170, 120879.
- Tsvetkova, A., Gustafsson, M., & Wikström, K. (2021). Digitalizing maritime transport: digital innovation as a catalyzer of sustainable transformation. *A Modern Guide to the Digitalization of Infrastructure*, 123–148.

- Tuunanen, T., Lumivalo, J., Vartiainen, T., Zhang, Y., & Myers, M. M. (2023). Micro-Level Mechanisms to Support Value Co-Creation for Design of Digital Services. *Journal of Service Research*.
- Vargo, S. L., Fehrer, J. A., Wieland, H., & Nariswari, A. (2024). The nature and fundamental elements of digital service innovation. *Journal of Service Management*, 35(2), 227–252.
- Vargo, S. L., Koskela-Huotari, K., & Vink, J. (2020). Service-Dominant Logic: Foundations and Applications. In E. Bridges & K. Fowler (Eds.), *The Routledge Handbook of Service Research Insights and Ideas* (pp. 3–23). Routledge.
- Vargo, S. L., & Lusch, R. F. (2008). Service-dominant logic: Continuing the evolution. *Journal of the Academy of Marketing Science*, 36(1), 1–10.
- Vargo, S. L., & Lusch, R. F. (2016). Institutions and axioms: an extension and update of service-dominant logic. *Journal of the Academy of Marketing Science*, 44(1), 5–23.
- Vial, G. (2019). Understanding digital transformation: A review and a research agenda. In *Journal of Strategic Information Systems* (Vol. 28, Issue 2, pp. 118–144). Elsevier B.V.
- Watson, R. T., Lind, M., Liesa, F., & Delmeire, N. (2021). Shipping: A Self-Organising Ecosystem. In *Maritime Informatics* (Issue 2020, pp. 13–32).
- Wieczerzycki, M., Ratajczak-Mrozek, M., Hauke-Lopes, A., & Colurcio, M. (2024). Value-in-context: co-creation across different context levels in the service ecosystem. *Journal of Business & Industrial Marketing*.
- Zhang, H., Yuan, S., Zhang, F., Wang, B., & Luo, X. (Robert). (2024). A Systematic Literature Review on IT-enabled value Co-creation: Toward an integrative framework. *Computers in Human Behavior*, 152.
- Zis, T. P. V., Psaraftis, H. N., Tillig, F., & Ringsberg, J. W. (2020). Decarbonizing maritime transport: A Ro-Pax case study. *Research in Transportation Business & Management*, 37, 100565.

"DEAR APPLICANT, THE AI WILL SEE YOU NOW!" JOB APPLICANTS' REACTIONS TO AI-ENHANCED SELECTION

FLORIS VERHEIJEN, XANDER LUB, RAN ZHANG

HU University of Applied Sciences Utrecht, the Netherlands
floris.verheijen@student.hu.nl, xander.lub@hu.nl, ran.zhang@hu.nl

As artificial intelligence (AI) reshapes hiring, organizations increasingly rely on AI-enhanced selection methods such as chatbot-led interviews and algorithmic resume screening. While AI offers efficiency and scalability, concerns persist regarding fairness, transparency, and trust. This qualitative study applies the Artificially Intelligent Device Use Acceptance (AIDUA) model to examine how job applicants perceive and respond to AI-driven hiring. Drawing on semi-structured interviews with 15 professionals, the study explores how social influence, anthropomorphism, and performance expectancy shape applicant acceptance, while concerns about transparency and fairness emerge as key barriers. Participants expressed a strong preference for hybrid AI-human hiring models, emphasizing the importance of explainability and human oversight. The study refines the AIDUA model in the recruitment context and offers practical recommendations for organizations seeking to implement AI ethically and effectively in selection processes.

DOI

<https://doi.org/10.18690/um.fov.4.2025.13>

ISBN

978-961-286-998-4

Keywords:

AI-enhanced selection,
applicant perceptions,
AIDUA model,
transparency and fairness,
human-AI collaboration



University of Maribor Press

1 Introduction

“Congratulations, you’ve been selected for an interview! Please record your responses, and our AI will evaluate your answers.” How would you feel if you were to read this in your job application process? Excited? Discouraged? Skeptical?...

As AI-enhanced hiring processes become more common, job applicants increasingly find themselves engaging with chatbots, asynchronous video interviews, and algorithm-driven assessments instead of human recruiters. While these AI-powered selection tools promise efficiency, scalability, and consistency, they also raise concerns about fairness, transparency, and trust (Ochmann & Laumer, 2020; Zhang & Yench, 2022). Applicants often question how AI interprets their responses, whether its decisions are truly objective, and what it means for human judgment in hiring. Research has shown that candidates frequently express skepticism toward AI-enhanced selection methods, particularly when they lack transparency or provide limited opportunities for human interaction (Van Esch et al., 2019; Suen et al., 2019).

This study explores these concerns by applying the Artificially Intelligent Device Use Acceptance (AIDUA) model to examine how candidates perceive and evaluate AI-enhanced hiring processes. The AIDUA framework provides a structured lens to analyze how applicants cognitively appraise AI in hiring, from initial reactions shaped by social influence and anthropomorphism to assessments of AI’s effectiveness, fairness, and usability (Gursoy et al., 2019). Through qualitative insights, this study identifies key psychological mechanisms that determine whether candidates accept or reject AI-enhanced selection methods, while also highlighting critical gaps in transparency and trust that influence AI adoption.

By investigating applicant perceptions across different AI-enhanced hiring scenarios, this study contributes to both theory and practice by addressing three key research gaps. First, while AI-enhanced hiring has gained increasing attention, prior research has primarily focused on organizational adoption rather than candidate perceptions and experiences (Hewage, 2023). Second, the AIDUA model, though widely applied in AI service contexts, has not been extensively tested in HRM and recruitment settings (Gursoy et al., 2019). Third, while fairness in AI-enhanced selection has been widely debated, there is still limited empirical research on how transparency influences applicant trust in AI-enhanced hiring (Suen et al., 2019). This study

addresses these gaps by applying the AIDUA model to hiring, investigating how social influence, anthropomorphism, as well as effort and performance expectancy shape candidate acceptance of AI-enhanced selection, and exploring transparency and fairness as key determinants of trust. In doing so, it refines our understanding of how AI-enhanced selection methods align or clash with applicant expectations, offering actionable insights for organizations seeking to balance automation with human oversight.

2 Research background

AI in hiring and selection processes

AI-enhanced selection is increasingly deployed to boost efficiency and objectivity in recruitment processes. Applications of AI in these processes include AI-assisted interviews, asynchronous video interviews, and chatbot-led interactions, each designed to streamline candidate assessment and selection. AI-assisted interviews utilize algorithms to analyze verbal and non-verbal cues, providing recruiters with data-driven insights (Van Esch et al., 2019). Asynchronous video interviews allow candidates to record responses at their convenience, which are then evaluated by AI systems, offering flexibility and scalability in the hiring process (Van Iddekinge et al., 2023). Chatbot-led interviews involve AI-driven conversations that can handle initial screening and answer candidate queries in real time (Ochmann & Laumer, 2020).

While these AI-enhanced methods offer notable advantages, they also present significant challenges for applicant experience and acceptance. Concerns about transparency arise when candidates are unaware of how AI systems evaluate their responses, leading to potential distrust. Perceived fairness is another critical issue, as AI systems may inadvertently perpetuate existing biases present in their training data, affecting decisions related to hiring and promotions (Açikgoz et al., 2020). Trust deficits can emerge if candidates feel that AI lacks the human judgment necessary to understand contextual nuances, potentially impacting their confidence in the selection process (Zhang & Yench, 2022). These concerns have led scholars to emphasize the need for human oversight in AI-driven recruitment (Suen et al., 2019). A hybrid approach that integrates algorithmic analysis with human decision-making is increasingly seen as a way to balance the efficiency of AI with the empathy, ethical reasoning, and contextual sensitivity of human recruiters (Suen et al., 2019).

Building on this foundation, recent research has expanded our understanding of how applicants perceive and evaluate AI-driven recruitment tools. Studies have shown that while AI can improve efficiency and reduce bias, it may also introduce concerns around algorithmic opacity, perceived dehumanization, and reduced agency in the hiring process. For example, Binns et al. (2018) emphasize how opaque AI decision-making can undermine perceptions of justice and reduce applicant trust. Similarly, Wesche et al. (2024) find that candidates respond more positively to algorithmic decisions when explanations are provided, suggesting that transparency plays a critical role in fostering trust. Other recent work has explored concerns about behavioral control and procedural fairness in AI-based recruitment (Hilliard et al., 2022), while Van Esch et al. (2021) highlight the importance of perceived fairness and trust in shaping applicant reactions to AI-enabled hiring processes. These studies underscore the need for further research into candidate-centered evaluations of AI systems and provide a strong basis for applying models such as AIDUA in the recruitment domain.

Artificially Intelligent Device Use Acceptance Model

To understand how applicants perceive and respond to AI-enhanced hiring, this study applies the Artificially Intelligent Device Use Acceptance (AIDUA) model (Gursoy et al., 2019). This framework evaluates how individuals accept or reject AI-based technologies, structured around three cognitive appraisal stages.

In primary appraisal, users form initial perceptions based on social influence, anthropomorphism, and hedonic motivation (Van Doorn et al., 2017). Social influence reflects societal and peer opinions on AI use, while anthropomorphism refers to how human-like an AI system appears, affecting trust and comfort. Hedonic motivation relates to the enjoyment or engagement derived from AI interactions, influencing acceptance. In secondary appraisal, users assess AI's effectiveness and ease of use. Performance expectancy refers to AI's perceived usefulness in achieving desired outcomes, while effort expectancy concerns how easy or complex AI interaction appears (Venkatesh et al., 2003). The outcome stage implicates users' emotional response towards and behavioral intentions (i.e.: accept or reject) regarding AI-enhanced hiring. Trust in AI's decision-making and concerns about fairness could significantly shape these elements (Wesche et al., 2024).

In sum, this study applies the AIDUA model to examine applicant perceptions of AI-enhanced hiring processes, exploring how acceptance and rejection response varies across different AI-integrated recruitment scenarios.

3 Research method

This study employed a qualitative approach to explore applicant perceptions of AI-enhanced hiring processes. Given the complex and subjective nature of human reactions to AI-enhanced selection, qualitative methods allowed for a deeper understanding of how candidates experienced and interpreted these technologies in recruitment settings.

Participants

Table 1: Sample profile

ID	Interview Date	Duration (hh:mm)	Vignette	Gender	Age	Educational Background	Employment Status
1	16-04-2024	01:16	B	Male	48	MSc. Educational Sciences	Full-time
2	17-04-2024	01:13	A	Male	29	MSc. Applied Ethics	Part-time
3	24-04-2024	00:49	A	Male	46	MSc. Industrial & Organizational Psychology	Full-time
4	25-04-2024	01:15	B	Male	40	MSc. International Development & Postdoc Journalism	Full-time
5	30-04-2024	01:15	A	Female	27	BSc. Economic Development and Globalization	Part-time
6	01-05-2024	01:12	B	Male	36	MSc. Cultural Anthropology	Part-time
7	01-05-2024	01:24	A	Male	58	MSc. Civil Engineering & Management & MSc. Business Admin	Full-time
8	02-05-2024	01:44	B	Female	53	MSc. Business Economics	Full-time
9	06-05-2024	01:20	A	Male	35	MSc. Clinical and Industrial & Organizational Psychology	Full-time
10	16-05-2024	01:10	A	Male	46	MSc. Industrial & Organizational Psychology	Full-time
11	17-05-2024	01:14	C	Female	33	MSc. Human Resources	Full-time
12	22-05-2024	01:32	C	Male	56	MSc. Electrical Engineering	Part-time
13	23-05-2024	01:28	C	Female	48	Executive Master of Security and Defense	Full-time
14	28-05-2024	00:48	C	Female	31	MSc. International Business & MSc. Marketing	Part-time
15	28-05-2024	00:49	C	Female	40	Executive Master of Security and Defense	Full-time

The study sampled employees primarily from the Dutch Ministry of Defense as they provided a unique perspective on AI-enhanced selection within a structured and professional setting. A purposive sampling strategy was used to ensure participants had relevant experience with selection procedures and recruitment technologies. A total of 15 participants took part in the study, offering diverse viewpoints on the use of AI in hiring (see Table 1).

Data collection

The study utilized semi-structured interviews as the primary data collection instrument. This approach enabled participants to share detailed insights while allowing the researcher to probe further into key themes that emerged. A vignette-based method was incorporated to present participants with three distinct AI-enhanced hiring scenarios, encouraging them to reflect on their reactions and decision-making processes in these contexts. The three scenarios were: (A) AI-assisted interview -- a human recruiter conducted the interview, but AI provided analysis and decision support; (B) Asynchronous video interview -- candidates recorded their responses to pre-set questions, which were then analyzed by AI; (C) Chatbot-led interview -- the entire interview process was conducted by an AI-driven chatbot without human intervention. Participants were presented with one of the three vignettes and were asked to reflect on their thoughts, emotions, and concerns regarding AI-enhanced hiring. Each vignette was used on five participants, resulting in an even distribution of the three scenarios across the total 15 participants. The interview guide included open-ended questions designed to elicit detailed responses on fairness, transparency, trust, and perceived usability.

Analytical approach

The collected interview data were analyzed using template analysis, a structured yet flexible method that allowed for coding based on predefined themes while accommodating new insights from participant responses. The analysis was guided by the AIDUA model, ensuring that findings were interpreted within the focal theoretical framework.

Ethical considerations

Ethical approval was obtained prior to data collection and participants provided informed consent before taking part in the study. To protect participant confidentiality, all data were anonymized, and interview transcripts were securely stored. Participants were informed of their right to withdraw at any stage without consequences.

4 Results

The findings of this study provide insights into applicant perceptions of AI-enhanced hiring processes across the three recruitment scenarios: AI-assisted interviews (Vignette A), asynchronous video interviews (Vignette B), and chatbot-led interviews (Vignette C). The results section is structured according to key themes from the final coding template framework (see Table 2), reflecting applicants' experiences, evaluations, and outcomes in relation to AI-enhanced selection methods.

Job application processes

Participants acknowledged that AI-enhanced selection processes could potentially speed up recruitment and eliminate human scheduling conflicts, particularly in AI-assisted interviews (Vignette A). However, some found AI-enhanced selection misaligned with the purpose of a job interview, which they expected to be an opportunity to engage with recruiters, demonstrate interpersonal skills, and learn about company culture. One participant expressed skepticism about AVIs, stating:

"A job interview... is not only meant for the employer to test whether the applicant fits within the organization, but is also meant for the applicant to see whether the organization fits him or her. And I think it is also a bit about the fit or the match that you feel on both sides." (Interviewee 3)

Others echoed concerns about AI-enhanced hiring replacing the human aspect of evaluation, with one participant saying:

"In a sense, it feels less fair, because you are not necessarily judged on the qualities you need for the job, but on the qualities you need to get through the application procedure." (Interviewee 6)

Familiarity with AI

Most participants were familiar with AI through tools such as ChatGPT and recommendation algorithms but had not considered its role in hiring. Some only realized their reliance on AI when asked. While some saw AI as functional, others doubted its ability to assess skills and personality, especially in chatbot-led interviews. As one participant noted:

"If you think about it, we are actually surrounded by Artificial Intelligence without you perhaps being aware of it. What is it? Facial recognition on your phone, Netflix deciding whether I like the movie or not." (Interviewee 13)

Primary appraisal

Applicants formed initial reactions to AI-enhanced selection based on social influence, hedonic motivation, and anthropomorphism. While some saw AI in hiring as inevitable but undesirable, others were curious about its capabilities. One participant viewed AI interviews as an opportunity to explore the technology firsthand, stating:

"I would really love to have an interview with such an advanced chatbot. Just from my personal interest and how well that works and if you ask questions about more on the emotional, on the emotional axis, how you would respond to that, I would find that very interesting." (Interviewee 11)

Perceptions of anthropomorphism played a major role in applicants' acceptance of AI selection. AI-assisted interviews (Vignette A) were seen as more legitimate due to the presence of a human recruiter, offering oversight and fairness. In contrast, chatbot-led interviews (Vignette C) were widely rejected, as participants struggled to engage with an AI lacking human interaction. One interviewee bluntly summarized their reaction:

"This is a nightmare scenario!" (Interviewee 12)

This strong negative response underscores how the removal of human interaction in chatbot interviews felt unsettling, impersonal, and even dehumanizing.

Secondary appraisal

At this stage, participants evaluated AI's usefulness and fairness, with mixed responses. Some appreciated AI's ability to process large volumes of data efficiently, while others questioned whether AI-based decisions were truly fair and explainable. Regarding Performance Expectancy, participants acknowledged AI could enhance consistency in hiring but worried about over-reliance on algorithms. One participant stated:

"Well, I think it's more honest than doing it with a human. Because you completely remove that subjectivity and that cognitive bias." (Interviewee 15)

Another noted:

"If you take the human factor out of it, you take out a whole bunch of potential ways that it could be unfair. I think there are more opportunities for humans to make an unfair decision than there are for an AI to make an unfair decision." (Interviewee 6)

Regarding Effort Expectancy, several participants found AI-driven processes mentally exhausting due to uncertainty about how responses were interpreted. One participant explained:

"I think I would behave a little differently. [...] It seems there's a lot more gravity behind everything you do and say." (Interviewee 2)

The effort needed to prepare for and adapt to AI-based selection criteria was seen as a disadvantage, particularly in AVI and chatbot scenarios.

Outcome stage

Trust in AI-enhanced selection was key in determining acceptance or rejection. Some participants saw AI as improving fairness in hiring by removing human biases. However, most still preferred human involvement in final decisions, citing AI's lack of judgment in assessing complex skills, emotions, and cultural fit. One participant summarized their stance on fully automated hiring:

Table 2. Condensed final coding template framework

Themes	Sub-themes	Descriptive codes	Overarching themes
Familiarity with AI applications	Use of AI	ChatGPT, Regular, daily use, Unconscious use of AI, Functionalities & applications used	Familiarity with AI applications
	Cognitive & algorithmic bias	Computer performs better than humans when processing vast amounts of info, Fast processing of vast information, Computer or algorithm, Lack of meaningful human control	
Understanding of AI	Drawbacks	Mimics human characteristics, Mathematical solution, Automates and accelerates	
	Experiences with interviews within application processes	Limits freedom of choice, Not transparent, Undesirable consequences	Experience with job application processes
		Job application procedures are valuable, Mismatch between procedure and purpose/ function, Mismatch between expectations and execution, Time-consuming	
		Useful for narrowing a large applicant pool	
	Expectations regarding the purpose of a job interview	Measuring applicants' KSAO's, Measuring the fit between applicant, the job and the context, Meet future colleagues, Opportunity to perform	Social influence
		Receive feedback on performance during the procedure, Consistent treatment in regard to other candidates, Procedure should reflect job requirements	
		When I cannot decide, I'll ask relatives or peers, Primary appraisal, With important decisions, I decide myself	
Self-insights	Anthropomorphism	AI is not human, Should be able to assess human-like interaction, Should be able to provide human-like interaction	Primary appraisal
		Importance of a having a job, Reason to apply for a job	Performance expectancy
		AI-Decisions should be based on clear operationalization of function requirements, AI not yet capable of supporting job interviews on a sufficient level	
	Accuracy expectations concerning the AI technology used	Depending on the purpose of a job interview, Human involvement necessary in de procedure, Should be explainable	
		Fair selection decision, Informs applicants on the use of AI in advance, Provide a fair procedure, Respectful treatment, Privacy protection GDPR	Effort expectancy
		Trust in improbability of performance of AI in regard to bias, Risk of automation bias	
	Expected actual use behavior prior to participation	Searching the internet, Wondering what AI takes into account, Prepare participation	
Emotions	Expected effort/drawbacks	No personal connection, Being watched is tiresome, Important to present as good as possible	Fairness perceptions
		Stressful, Weird, Impersonal, Scary	
		Fair decision, Fair procedure	
Signaling		(Un-)professional organization, Innovative organization, Impersonal approach	Outcome stage

"Okay, now you still have that meaningful human control, because it is not yes or no, but it is advice, an extra consideration that you give to the person in the interview. So this would just feel to me like a kind of extra input that the team leader is now getting." (Interviewee 4)

Another explained their strong objection to AI-led hiring:

"What it comes down to is how much I want that job. [...] I wouldn't be jumping for joy, really. It feels impersonal and unbalanced." (Interviewee 8)

Taken together, the results reveal that while AI-enhanced hiring is perceived as efficient and objective, participants remain skeptical about its fairness, trustworthiness, and ability to replace human judgment. AVIs and chatbot interviews were widely criticized for being impersonal and requiring excessive adaptation, while AI-assisted interviews were viewed as more acceptable due to human oversight. The strong preference for hybrid AI-human selection models indicates that organizations should balance technological advancements with human judgment, ensuring AI remains a tool for efficiency rather than a sole decision-maker.

5 Discussion, conclusions, and implications

Theoretical implications

This study advances research on AI-enhanced selection by extending the Artificially Intelligent Device Use Acceptance (AIDUA) model to the domain of hiring. While AIDUA has primarily been applied to consumer and service contexts (Gursoy et al., 2019), our findings demonstrate its relevance for understanding job applicants' nuanced responses to AI-driven selection methods. In doing so, we not only apply the model to a new context but also enrich the dialogue between technology acceptance and recruitment scholarship.

Our findings complement and extend Gilliland's (1993) procedural justice model, which outlines fairness rules such as consistency, opportunity to perform, and feedback. Applicants' responses to AI selection tools reflected deep concern with these principles—particularly transparency, perceived legitimacy, and interpersonal treatment. The AIDUA model's secondary appraisal components, such as

performance and effort expectancy, help explain how these fairness judgments are cognitively constructed. For instance, perceived fairness was tightly linked to performance expectancy (e.g., whether AI could make accurate decisions) and to effort expectancy (e.g., whether the AI process was mentally taxing or confusing). This suggests that AIDUA provides a valuable theoretical bridge between perceptions of fairness and technology-specific appraisals, offering new explanatory depth for applicant reaction research.

In addition, the study highlights how AI selection methods are interpreted through a signaling lens. According to signaling theory (Spence, 1973), applicants infer organizational values from the recruitment process. Our participants frequently viewed fully automated processes—especially chatbot-led interviews—as signaling impersonality, cost-cutting, or a lack of care for applicants. In contrast, hybrid models with human oversight were interpreted as signals of professionalism, accountability, and respect. Integrating AIDUA with signaling theory strengthens our understanding of how applicants evaluate not only the tool’s functionality but also what it reveals about the organization behind it.

Finally, AIDUA offers theoretical value beyond traditional HR frameworks by incorporating constructs such as anthropomorphism and hedonic motivation. These dimensions are not typically addressed in models like procedural justice or signaling theory, yet our findings indicate they play a significant role in shaping applicant acceptance. For example, participants expressed discomfort with AI systems that felt “cold” or dehumanizing—especially when no human contact was present. Emotional reactions such as curiosity, stress, or alienation also influenced how candidates engaged with different hiring formats. These insights suggest that affective and experiential factors—central to AIDUA—are essential in designing AI hiring processes that are not only efficient, but also perceived as fair and human-centered.

In sum, this study refines and extends the AIDUA model by embedding it within core theories of fairness and signaling in recruitment. It positions AIDUA as a useful integrative framework for future research on applicant perceptions of AI, and highlights the need for recruitment scholars to consider both cognitive and emotional dimensions of technology-mediated selection.

Practical implications

The findings of this study offer valuable insights for organizations integrating AI into selection processes, particularly regarding how candidates perceive and respond to AI-enhanced hiring methods. Participants strongly preferred AI-assisted hiring models where human decision-making remained central. Many interviewees raised concerns that fully automated hiring processes, particularly chatbot-led and asynchronous video interviews, lacked the human element necessary for fair and effective candidate evaluation. While AI was seen as useful for pre-screening applications, human involvement was considered essential for making final hiring decisions and assessing soft skills and job fit.

Actionable recommendation: Ensure a human-in-the-loop system, where recruiters review AI-generated insights before making hiring decisions.

Transparency and explainability in AI hiring processes emerged as a key theme. Participants expressed frustration over the opaque nature of AI decision-making, questioning how responses were evaluated and whether AI assessments were fair. A more transparent approach would involve organizations providing candidates with clear explanations of how AI assessments work and the criteria being evaluated. Several participants suggested that receiving feedback on their AI interview performance would improve trust and acceptance of AI-driven selection.

Actionable recommendation: Provide candidates with explainability statements post-interview (e.g., "Your responses were analyzed based on X, Y, Z factors.").

Many participants found AI-driven interviews stressful due to their unfamiliarity with the process. Several interviewees noted that they would have felt more confident if they had been given the opportunity to rehearse before the actual interview. To address this, organizations can implement AI interview practice runs, allowing candidates to familiarize themselves with the format and how AI interprets responses. This would help candidates refine their performance and reduce uncertainty.

Actionable recommendation: Offer candidates a rehearsal or practice-run AI interview before their actual AI-driven application interview.

The candidate experience was another major concern. Many interviewees found chatbot-led and asynchronous video interviews impersonal, negatively impacting

their perception of the hiring process. Several participants expressed a preference for AI-enhanced hiring that included some level of human interaction, such as opportunities to clarify responses or interact with a recruiter at key stages. Designing AI-assisted interviews to include interactive elements could help mitigate candidate stress and improve the overall experience.

Actionable recommendation: Allow candidates to request clarifications, re-record responses, or engage with a human at key decision points.

Participants also highlighted the importance of aligning AI selection processes with organizational values. Several interviewees noted that the use of AI in hiring sent signals about the organization's priorities, and over-reliance on AI risked making companies seem impersonal or overly data-driven. Organizations that emphasize collaboration, innovation, or human-centered leadership should ensure that their hiring practices reflect these values by integrating AI in a way that complements, rather than replaces, meaningful human interaction.

Actionable recommendation: Ensure that AI-enhanced selection methods align with the organization's employer brand and values to enhance the candidate experience.

These insights underscore the need for a thoughtful and candidate-centric approach to AI-enhanced hiring. While AI can enhance efficiency and streamline recruitment, organizations must balance automation with human oversight to ensure fairness, transparency, and trust. By incorporating participant-driven recommendations into AI hiring strategies, companies can foster a more inclusive, engaging, and trustworthy selection process.

Limitations, future research, and conclusions

This study offers valuable insights into applicant perceptions of AI-enhanced selection, but several limitations should be noted. The sample, drawn from the Dutch Ministry of Defense, provides a structured and policy-aware context but may not reflect broader applicant populations. Participants may have heightened concerns around data security and formal procedures. Future research should

examine diverse sectors to explore how organizational culture influences responses to AI in hiring.

The qualitative approach enabled in-depth exploration but limits generalizability. Larger-scale surveys or experiments could test the prevalence of these findings across contexts. Longitudinal research may also clarify whether applicant attitudes shift as AI becomes more integrated into recruitment processes.

Finally, while our practical recommendations aim to improve candidate experience, interventions such as AI rehearsals may lead to unintended effects like over-preparation. Future work should examine how support tools influence fairness and authenticity in AI-driven assessment.

In conclusion, this study extends the AIDUA model to recruitment, highlighting how social influence, anthropomorphism, and performance expectations shape applicant reactions. Concerns about fairness and transparency emerged as central and should be integrated into future models of AI acceptance in hiring.

References

- Açikgoz, Y., Davison, K. H., Compagnone, M., & Laske, M. (2020). Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment*, 28(4), 399-416. <https://doi.org/10.1111/ijsa.12306>
- Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018, April). 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 Chi conference on human factors in computing systems* (pp. 1-14). <https://doi.org/10.1145/3173574.3173951>
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18(4), 694-734. <https://doi.org/10.5465/amr.1993.9402210155>
- Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers' acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management*, 49, 157-169. <https://doi.org/10.1016/j.ijinfomgt.2019.03.008>
- Hewage, A. (2023). Exploring the Applicability of Artificial Intelligence in Recruitment and Selection Processes: A Focus on the Recruitment Phase. *Journal of Human Resource and Sustainability Studies*, 11(3), 603-634. <https://doi.org/10.4236/jhrss.2023.113034>
- Hilliard, A., Guenole, N., & Leutner, F. (2022). Robots are judging me: Perceived fairness of algorithmic recruitment tools. *Frontiers in Psychology*, 13, 940456. <https://doi.org/10.3389/fpsyg.2022.940456>
- Ochmann, J., & Laumer, S. (2020). AI Recruitment: Explaining job seekers' acceptance of automation in human resource management. In M. Heine, K. Poustchi, & K. H., *WT2020*

- Zentrale Tracks* (pp. 1633–1648). GITO Verlag. https://doi.org/10.30844/wi_2020_q1-ochmann
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>
- Suen, H.-Y., Chen, M. Y.-C., & Lu, S.-H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior*, 98, 93–101. <https://doi.org/10.1016/j.chb.2019.04.012>
- Van Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., & Petersen, J. A. (2017). Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences. *Journal of Service Research*, 20(1), 43–58. <https://doi.org/10.1177/1094670516679272>
- Van Esch, P., Black, J. S., & Arli, D. (2021). Job candidates' reactions to AI-enabled job application processes. *AI and Ethics*, 1, 119–130. <https://doi.org/10.1007/s43681-020-00025-0>
- Van Esch, P., Black, J. S., & Ferolie, J. (2019). Marketing AI recruitment: The next phase in job application and selection. *Computers in Human Behavior*, 90, 215–222. <https://doi.org/10.1016/j.chb.2018.09.009>
- Van Iddekinge, C. H., Lievens, F., & Sackett, P. R. (2023). Personnel selection: A review of ways to maximize validity, diversity, and the applicant experience. *Personnel Psychology*, 76(2), 651–686. <https://doi.org/10.1111/peps.12578>
- 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. <https://doi.org/10.2307/30036540>
- Wesche, J. S., Hennig, F., Kollhed, C. S., Quade, J., Kluge, S., & Sonderegger, A. (2024). People's reactions to decisions by human vs. algorithmic decision-makers: The role of explanations and type of selection tests. *European Journal of Work and Organizational Psychology*, 33(2), 146–157. <https://doi.org/10.1080/1359432X.2022.2132940>
- Zhang, L., & Yencha, C. (2022). Examining perceptions towards hiring algorithms. *Technology in Society*, 68, 101848. <https://doi.org/10.1016/j.techsoc.2021.101848>

PREVENTING UNAUTHORIZED USE OF PERSONAL DATA IN GHOSTBOTS POST-MORTEM: AN ACCESS AND AUTHORIZATION MODEL FOR DIGITAL WILLS

JASMIN BÖDECKER, JÜRGEN KARLA

Hochschule Niederrhein University of Applied Sciences, Faculty of Business
Administration and Economics, NRW, Krefeld, Germany
jasmin.boedecker@stud.hn.de, juergen.karla@hs-niederrhein.de

The rise of AI-driven technologies, such as ghostbots, has introduced new challenges in digital legacy management, particularly regarding post-mortem data usage. A digital will can address this challenge. Following the Design Science Research Method, this research proposes a structured access control model that enables Testators to define clear permissions for data access, inheritance, and the creation of ghostbots. The model integrates predefined roles and conditional access policies to ensure the Testators wishes regarding post-mortem data usage are met. The model is assessed by using logical proof and scenario-based evaluation. The findings highlight the necessity of robust access control mechanisms to prevent unauthorized use of personal data to enable ethical AI practices. By addressing these issues, this study contributes to the broader discourse on digital inheritance and ghostbots and provides an access control model framework for managing post-mortem digital identities through a digital will.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.14](https://doi.org/10.18690/um.fov.4.2025.14)

ISBN
978-961-286-998-4

Keywords:
ghostbots,
digital will,
post-mortem privacy,
personal data,
access control model



University of Maribor Press

1 Introduction

The accelerated development of artificial intelligence (AI) has enabled public access to popular systems such as ChatGPT. AI is becoming a part of multiple aspects of modern life, from autonomous driving to improving medical diagnoses (Kitzmann, 2022). The increasing use of technologies in the personal life, such as social media, results in vast amounts of personal, digital data. Such information encompasses personal data and digital identities that persist after a person's death, raising critical questions about how to manage, access, and ensure the security of these digital legacies. So far, there are no concrete regulations for what happens to a person's data when they pass away (Fuchs, 2021). Some companies such as Facebook or Google provide some kind of solution for this problem, like designating a person in charge of the account (Harbinja 2017). In most countries, after a person dies, the autonomy over their data dies with them (Harbinja 2022).

A recent development in AI systems are so-called "ghostbots". Ghostbots are generative AI systems that simulate the personality and behavior of deceased individuals (Figueroa-Torres 2024). There are different forms of ghostbots, ranging from chatbots to virtual avatars. They claim to allow the bereaved to continue interacting with their deceased loved ones. This development is discussed highly controversial and raises numerous ethical, legal, and psychological questions that need to be examined (Harbinja et al. 2023). For instance, the question arises whether an ethical use of ghostbots is even possible when the deceased have not given prior consent for their data to be used in these technologies. Harbinja et al. suggest a platform to manage consent for data usage in ghostbots (Harbinja et al., 2023).

2 Motivation

A digital will allows individuals to determine how their data is used after death, especially regarding ghostbots. It provides consent management, nominates responsible parties, and enforces user wishes across digital platforms. An access and authorization model is essential to govern interactions and prevent unauthorized data use (Harbinja, 2017).

3 Research question

This paper focuses on creating an access and authorization control model for a digital will platform as a concrete response to these aspects. It governs who has the authority to access, manage, or delete data after a user's death. This model must be comprehensive, secure, and adaptable to the various stakeholders involved.

Ghostbots operate by using personal data, including text messages, social media activity, voice recordings, and multimedia content (Lindemann 2022). While these systems hold the promise of comforting bereaved loved ones by allowing them to maintain a sense of connection with the deceased, they also pose significant risks (Jiménez-Alonso and Brescó de Luna 2023). Without adequate oversight and explicit consent from the deceased, the creation of ghostbots can lead to severe privacy violations, unauthorized exploitation of personal identity, and emotional harm to those grieving. Ghostbots can harm the bereaved by being used as a replacement for the deceased, leading to self-deception (Fabry and Alfano 2024) and even leading to addiction (Lindemann 2022). The more realistic the ghostbots, the higher the chance of addiction (Fabry and Alfano 2024) and the more difficult grieving becomes (Reese 2023). A possible risk is the deceased being portrayed inaccurately (Bao and Zeng 2024), creating further harm. Ghostbots allow bereaved to have a fictitious relationship with them (Fabry and Alfano 2024), creating new memories (Lindemann 2022) by enabling bi-directional communication (Jiménez-Alonso and Brescó de Luna 2023). Even though ghostbots allow open communication in private (Puzio 2023; Jiménez-Alonso and Brescó de Luna 2023), the important part of connection and support of others is neglected. Grieving with ghostbots can thus be isolating (Puzio 2023). Disrupting the grieving process can lead to prolonged grief disorder, causing those affected to be restricted in everyday life. This leads to a greater dependency on ghostbots, creating a vicious cycle (Lindemann 2022). To prevent these implications, a digital will can give the data owner the control over their data but also protect their beloved ones from these harms. The proposed access and authorization control model addresses these risks by giving individuals the ability to explicitly regulate the post-mortem use of their data. This allows for a tailored and ethical approach to digital legacy management. This balance between protection and ethical data use reflects the challenges posed by evolving AI technologies in the realm of digital inheritance.

4 Current State of Research

4.1 Digital Legacies and Data Ownership

Due to the increase of personal data, digital legacies have gained more importance (Cook et al. 2019; Cupar et al. 2023). A digital legacy can include digital assets such as social media profiles, photographs, text messages or audio files (Peoples and Hetherington 2015; Cook et al. 2019; Dissanayake and Cook 2019; Harbinja 2022). As mentioned before, there is considerable debate regarding the ownership of data after death. Often people misunderstand the ownership of their data once posted to a platform and assume digital assets are handled like physical assets (Cook et al. 2019). However, unlike physical property, digital assets are often governed by the terms of service of the platforms on which they reside, rather than by established inheritance laws (Fuchs, 2021). Additionally, these platforms each have individual policies, making it difficult to manage one's digital legacy (Cook et al. 2019). Harbinja argues that digital assets should not necessarily be distinguished from physical assets. The lack of regulation leads to an unclear legal situation (Dissanayake and Cook 2019), allowing the bereaved to create ghostbots of loved ones, without prior consent. A universally accepted system is needed, not only for managing digital legacy (Dissanayake and Cook 2019), but for managing the growing popularity of ghostbots (Harbinja 2022).

4.2 Existing Solutions and Their Limitations

Within the European Union (EU) the General Data Protection Regulation (GDPR) sets regulations for data protection (European Parliament & Council of the European Union, 2016). The regulation specifies rights for data, including the right to deletion, restrictions on data processing, and access to personal data, which are particularly relevant to this research (European Parliament & Council of the European Union, 2016). In the context of ghostbots, the focus is on handling personal data of deceased individuals. Currently, many states lack specific laws regulating post-mortem data management, and it remains undetermined whether such data can be inherited or who holds rights over a deceased person's data.

4.3 Emerging Technologies: Ghostbots

A controversial development in digital legacy management is the emergence of ghostbots. To create ghostbots, personal data is used to recreate the voice, look, personality, and behavior of deceased individuals (Puzio 2023). While this technology may offer solace to some grieving individuals, it raises ethical questions (Puzio 2023). If ghostbots are created without the explicit consent of the deceased, this could be considered a serious violation of privacy. Moreover, there is a risk of psychological harm to the bereaved, who might struggle to cope with the simulation's realism (Harbinja et al., 2023).

Compared to other forms of remembrance such as videos or photos, ghostbots allow two-way communication. This allows new memories being formed, which otherwise would not have been possible (Puzio 2023).

4.4 Access and Authorization Control Models

The purpose of access control is to restrict data access to authorized entities only, ensuring the protection of information and preventing any potential misuse (Tsolkas and Schmidt 2017). In the context of an access control model, it is important to differentiate the terms authentication, authorization and access control (Mahalle et al. 2022). Authentication describes the identification of a user, while authorization sets rules which specify which actions a user is allowed to perform (Chen et al. 2009; Boonkrong 2020). Based on the authorization policies, access control then allows access to data or systems (Tsolkas and Schmidt 2017).

Building an effective access control model involves structuring a framework that manages how and when subjects can interact with objects (Tsolkas and Schmidt 2017). Subjects can be users or other systems and object different types of data. Central to this model are roles and permission levels, which define what actions each user can perform based on their responsibilities or identity. For example, in a role-based access control (RBAC) system permissions are assigned to roles rather than to individual users, streamlining the management of access rights (Ferraiolo and Kuhn 1997; Chen et al. 2009; Gupta et al. 2022). A more advanced, dynamic model might incorporate dynamic attributes to adapt permissions in real-time (Atlam et al. 2020; Pal 2021). These attributes can include time, location or events. The success

of an access control model relies on its ability to balance security with usability. Permissions should be clearly defined to prevent unauthorized actions, while still allowing legitimate users efficient access. The core elements must be precisely managed through well-established policies, ensuring that only authorized entities can perform actions.

5 Method

The research methodology chosen for this study is the Design Science Research Method (DSRM) (Hevner et al. 2004; Peffers et al. 2007). This structured approach allows the development of practical and innovative solutions to complex problems.

The first step identifies the problem, highlighting the challenges posed by AI technologies like ghostbots in digital legacy management. A structured literature review established the problem's relevance and the need for a solution. The second step defined objectives, aiming to develop a secure, flexible, and ethical access control model that ensures compliance with the deceased's wishes. The third step involves designing and developing the model, including role definitions and access control policies with a focus on ethical AI use. This is followed by demonstrating the model's functionality, showcasing predefined roles, conditional access enforcement, and consent management. The fifth step evaluates the model's effectiveness through theoretical consistency checks and thought experiments to ensure scalability and precision.

6 Artefact

Ghostbots are being created without the explicit consent of the deceased, raising significant ethical concerns. Additionally, there is a lack of clear directives for handling digital assets post-mortem, including consent to ghostbots. Since digital assets are typically not included in traditional wills, there is no consistent regulation governing the management of personal data after death. Existing platform policies vary widely, are often ambiguous, and lack standardization. A digital will presents a potential solution for individuals to manage their digital data posthumously and prevent ghostbots from being created unethically.

- The objective of this paper is to design an access and authorization control model for a digital will. The model's design requirements (DR) are derived from platform functions and relevant literature as follows: DR1: The model must precisely control which subjects can access specific kinds of data.
- DR2: The model must allow a structured and scalable access management.
- DR3: The access control model must include mechanisms for conditional access.
- DR4: The user should be able to control the consent management.

6.1 Access and Authorization Model

Access control models are frameworks designed to regulate who can access specific resources, what actions they can perform, and under what conditions (Salim et al. 2010; Tsolkas and Schmidt 2017; Boonkrong 2020). These models form the foundation of any secure system, providing mechanisms to enforce permissions and ensure that data, especially personally identifiable information remains protected from unauthorized use or manipulation (Larson 2022). For the digital will platform, a hybrid approach between RBAC and dynamic features is the most suitable. RBAC ensures a clear and scalable role definition for Testators, heirs, and platform administrators, while dynamic elements allow the inclusion of conditional access policies. This combination aligns the platform's needs for clarity and adaptability, ensuring sensitive data is managed securely while accommodating the individual needs of post-mortem data management.

Firstly, the basic parts of an access control model need to be defined. These consist of subjects, objects and actions. Based on the requirements described previously, there are two types of subjects. Subjects refer to both the human users, as well as non-human users. These can be other systems or third-party platforms, that interact with the platform. The primary subjects will have higher privileges and access rights. The data owner controls the rules for the other users. Even when they are no longer actively using the system, their consent and wishes dictate access policies. Followed by the data owner, the primary executor would be the next user in hierarchy. A primary executor is a trusted person, in charge of managing the deceased data according to their digital will. A legal representative has authority over data when legal obligations or disputes need to be handled. A legal representative can also be a

primary executor. These rolls are defined by the data owner. Secondary subjects have limited or specific privileges in comparison to primary subjects. The secondary subjects can be divided into three categories, family members or next of kin, friends and third-party platforms. The third-party platforms can be differentiated into platforms with and without the use of AI. More specifically, platforms which main purpose is to use the deceased data to create ghostbots. Family members or friends have limited access rights based on the deceased preferences while access for third-party platforms is restricted to automated, API-based communication.

Objects are the data or resources that need to be protected and controlled within the platform. It is important to understand the types of data and their sources, to allow appropriate mapping between the objects and subjects. Furthermore, this helps define actions permissible for each relationship. There are different types of data, accumulating from different sources, creating digital identities. Data is not limited to text, audio and visual, but can include metadata, data from digital entertainment, search histories, financial data and any other data that is created using platforms. The different types of objects, or data, can come from different sources. The data owner can upload content directly to the platform. Otherwise, data can be imported through cloud storage, social media platforms or other types of accounts.

It is important to differentiate data by type and source. Being able to give precise permissions for different types of data, allows granular control. It adds scalability, as new data types or sources can be incorporated without restructuring the entire system. This categorization directly influences the design of the access control model. Categorizing data by type and source helps tailor an access control model to ensure that specific roles access only the data relevant to them. Incorporating dynamic features and defining subject-object relationships further refine the model, enabling conditional and precise management of actions allowed for each user role. Actions are operations that subjects can perform on objects. These actions are detect, search/find, compare, show, read, add, change, delete and execute (Tsolkas and Schmidt 2017). These actions can be assigned to the individual roles and be restricted by conditional access. Privileges are the specific permissions granted to subjects to perform actions on objects. These are linked to the subjects' roles.

6.2 The Access Control Model

Predefined roles serve as the foundation of this system, each with default settings aligned with the purpose of the platform and the ethical considerations of post-mortem data management. These roles are categorized into primary roles, which hold the highest levels of responsibility and access, and secondary roles, which are more limited in their permissions. Additionally, third-party platforms are integrated through controlled API connections, ensuring that their access is tightly regulated and specific to predefined tasks.

Primary and Secondary Roles

Primary roles include the Primary Executor and the Legal Representative, both of whom are entrusted with high levels of access but are distinguished by their specific responsibilities.

Secondary roles include the Heir, Next of Kin as well as Additional Individuals. They are provided with restricted access to specific data categories and actions. Table 1 displays the different roles and default actions. The primary and secondary roles do not include system roles, since system roles are not controlled by the Testator. The platform administrator is a system role. This role is responsible for maintaining the platform's functionality and does not have access to the user's personal data. Access is highly restricted and focuses on maintenance, policies and security rather than managing the digital will of people.

The dynamic features enhancing the RBAC framework, allow users to adjust the default permissions to suit their unique circumstances. For instance, the Primary Executor or Heirs may receive conditional access that activates upon the confirmation of death or at specific time intervals. Granular control enables the data owner to define permissions for each role at a detailed level, such as allowing an Heir to view financial records while restricting access to personal emails. This flexibility ensures that the platform can adapt to individual preferences while maintaining the integrity and security of sensitive data.

Table 1: Primary and Secondary Roles

Role	Who	What	When
Testator	User creating their will	Create, Edit and delete digital will, define access control for individuals, view access protocols	Always
Primary Executor	Individual chosen by Testator	Full Access, execute will	After legal confirmation of Testators death, active until completion of will
Legal Representative	Lawyer, notary or legally appointed individual	Validate and confirm will, conditional access only for legal processes	Access triggered by predefined legal needs, expires once needs are completed
Heir	Individuals	Medium Access, view data and perform actions based on the will	Access begins with confirmation of Testators death, expires once inheritance-related tasks are completed
Next of Kin	Family Members	Limited Access, view only	Access triggered by confirmation of Testators death, Access remains indefinitely
Additional Individuals	Friends, Colleagues, Acquaintances	Limited Access, Highly specific, view-only access	Access triggered by confirmation of Testators death, Access remains indefinitely

Third-Party Platforms

Third-party platforms are incorporated into the access control model through APIs. These entities do not interact directly with the platform or its users but receive automated instructions to perform specific tasks. Their access is conditional and highly restricted, ensuring they can only complete the tasks defined by the user's will and cannot access or manipulate unrelated data. The default settings for third-party platforms are show in Table 2.

A critical function of the digital will platform is to provide users with comprehensive control over whether or not a ghostbot be created. Not only should the user be able to make the decision if, but with what data a ghostbot can be created. The access granted in the digital will should reflect the Testators consent and permissions. This control extends beyond approval for creation, but dictate what data can be used,

who can interact with the ghostbot, how long it should exist, and in what form it should be presented.

Table 2: Third-Party Platforms

Type of Third-Party Platform	Who	What	When
Social Media	Social Media Platforms	Delete/deactivate account	Triggered by confirmation of Testator's death
Email Accounts	Email Service Providers	Create and save backup, delete account	Triggered by confirmation of Testator's death
Cloud Storage	Cloud Service Providers	Create Backup, Transfer Ownership	Triggered by confirmation of Testator's death
Financial Accounts	Banking platforms	Transfer ownership	Triggered by confirmation of Testator's death
Ghostbots	Ghostbot Platforms	No Access	Triggered by confirmation of Testator's death

The platform places the decision to permit or deny the creation of a ghostbot entirely in the hands of the data owner. During their lifetime, users can specify whether they wish to allow the generation of a ghostbot after their passing. This decision is reflected in the access, by allowing no access. For those deciding for a ghostbot, the platform allows detailed access policies. The user can define which types of their data are to be utilized in generating the ghostbot's personality, communication style, and responses. Additionally, the platform extends to manage who can interact with the ghostbot. Users can assign specific roles to their chosen heirs, family members, or other individuals, determining who has the privilege to access the ghostbot. For instance, the data owner might decide that only close family members can engage with the ghostbot, while more distant acquaintances have no access. These permissions are managed with the same rules as the broader access control features. This ensures that interactions with the ghostbot align with the user's intent. This can prevent misuse of the ghostbot. Moreover, the platform allows users to establish time-based constraints on the ghostbot's existence. As mentioned previously, data is usually bound to a certain context. By adding time-based constraints to the ghostbot, it can be avoided to create a ghostbot, that is contained in its social context and timeline. This capability helps to decrease ethical concerns regarding the indefinite

representation of individuals who have passed away, which could lead to unintended consequences over time. Equally important is the ability for users to define the scope and format of the ghostbot. For example, some users might want the ghostbot to convey only curated messages for memorial purposes, while others may allow more dynamic and personalized interactions based on their historical data.

All these decisions are enforced by the access control model, which ensures that only authorized actions are carried out. If the data owner wishes to revoke permissions for the creation or use of the ghostbot, the platform must immediately halt any associated processes and restrict access to the relevant data. Likewise, interactions between the ghostbot and third-party platforms, such as social media networks or external applications, are mediated through secure API integrations, adhering strictly to the data owner's established preferences. By providing users with detailed and transparent tools to manage the creation and use of ghostbots, the digital will platform upholds the principles of consent, autonomy, and ethical responsibility. It ensures that these AI representations serve as an extension of the data owner's wishes, rather than a violation of their legacy, fostering trust and accountability in this sensitive area.

7 Logical proof

The basic function of the model is allowing a user precise control over the access to their data (DR1). The model is based on roles with access permissions for specific actions and specific objects. The objects represent the different types of data, subjects refer to the people designated to have access to the objects and actions. Enforcing the permissions through roles ensures that only authorized people can perform actions. The differentiated kinds of objects allow the user to specify actions with a high level of granularity. Additionally, conditional access controls that access is only activated once the user passes away. For example, a friend of the user could receive access to selected photographs, while the heir receives access to all images.

Structured and scalable access management (DR2) is another essential design requirement. The role-based architecture enables a simple structure, reducing ambiguity in access control. This inherently provides structure because roles encapsulate predefined permissions for specific actions on specific data. Additionally, it supports scalability, since new roles and permissions can be added to

the system without redesigning the entire framework. For this requirement, the dynamic features allow further structurization of access model, such as time limits for data access. For example, if a user wants to add more people to the will, they would be added to the role “additional individual” with limited access as a default. The user can then add individual permissions, with time constraints such as “view images for 4 weeks”.

As mentioned before the dynamic features enable the user to add conditions to permissions (DR3). The roles act as baseline for permissions while conditional access mechanisms refine them. Conditional access can include time-related conditions or event-driven conditions. One crucial event-driven condition being the legal confirmation of the user’s death. This ensures that the digital will is only activated when necessary, preventing misuse of data.

The fourth design requirement focuses on consent management, especially in context to ghostbots (DR4). Consent is managed through the access that third-party platforms creating ghostbots receive. The first step is deciding for or against ghostbots. If the user decides against it, access is denied. If creation of ghostbots is allowed, access can be granted granularly for the different types of data or third-party platforms. This way, the user is in control of how he is portrayed. Additionally, the dynamic features can limit the time, the ghostbots exists or who can have access. This could lower the risks for the bereaved that can occur when the creation and access to ghostbots is not controlled.

8 Discussion and limitations

The proposed model gives insight to a small part of the digital will. Future research should explore the security measures that can be implemented, to ensure the safety of personal data against potential threats. Improving user verification through two-factor authentication could significantly improve user verification. Temporary access keys that expire after usage could further strengthen security. To provide transparency and accountability, all user actions should be strictly logged. Access to critical system components by the platform administrator could further be restricted through VPN. The frontend should be intuitive and accessible, both for the Testator and the people who have been given access to the will. It should be evaluated, to

what extend guidance and tips are necessary, both for the Testator to make informed decisions, but also to enable users' easy guidance during grief.

The proposed access control model fulfills the requirements, describes, and demonstrates potential for a digital will platform. It addresses precise control over data, scalability and conditional access, especially in context of ghostbots. However, there are several limitations that must be acknowledged. Differences in legal frameworks, cultural norms and technological adoptions across countries lead to significant issues. This lack of consistency in international regulation complicates the implementation of a universal digital will. Digital assets and posthumous data management need to be regulated on a governmental level, before a digital will can be implemented. Without mandates or regulations requiring platforms to integrate the digital will, gaps in coverage are to be expected. To ensure that a digital will is executed, coordinated legislative action is required. This is currently lacking.

One of the biggest limitations concerns ghostbots specifically. Individuals can still create ghostbots using private data they have, without oversight. Even though the user decision against ghostbots could be grounds for legal action against otherwise created ghostbots, it is difficult to find those ghostbots. Through this loophole, the control mechanisms of the platform can be undermined. A mandatory check between ghostbot platforms and the digital will, could close this gap, however the risk remains. This could lead to ethical issues or even misrepresentation of the deceased. Addressing these issues will require a combination of regulatory interventions, increased public awareness, and platform accountability to ensure the digital will system functions as intended and respects the complexities of diverse user needs and contexts.

The model proposed offers great detail in customizing access. While the model does support scalability, the growing number of users and data may require additional infrastructure. Additionally, conditional access strengthens security, however it also introduces complexity. This could lead to misconfigurations if not implemented carefully. Third-party platforms are connected using APIs. Although it enhances functionality, it can expose vulnerabilities. Especially due to the nature of the data, compromised security is a great risk.

9 Conclusion

There is a growing need for ethical posthumous data management. Developing an access model, gives insight to a digital will platform addressing this issue. The proposed model incorporates roles, different types of media and dynamic feature for conditional mechanisms. However, significant challenges remain. The rapid developments and lack of unified legal regulations create challenges that a digital will cannot overcome. These issues highlight the need for regulatory mandates that compel digital platforms to integrate with digital will systems to ensure comprehensive coverage. There is need for ongoing education about personal data, one's digital footprint and new AI-technologies such as ghostbots. Ghostbots especially introduce ethical, legal and privacy concerns. Possible unauthorized ghostbots violate the deceased and pose risks for the bereaved.

While the proposed access control model outlines a possible foundation for managing digital legacies, the success depends on collaboration with and between governments, providers of platforms and ghostbots. Addressing the risks associated with emerging technologies like ghostbots and ensuring ethical oversight are critical. With continued refinement, adaptation to legal and technological changes, and a focus on accessibility, the platform has the potential to set a standard for ethical and secure posthumous data management in the digital age.

References

- Atlam HF, Azad MA, Alassafi MO, et al (2020) Risk-based access control model: A systematic literature review. *Future Internet* 12:. <https://doi.org/10.3390/fi12060103>
- Bao A, Zeng Y (2024) Embracing grief in the age of deathbots: a temporary tool, not a permanent solution. *Ethics Inf Technol* 26:. <https://doi.org/10.1007/s10676-024-09744-y>
- Boonkrong S (2020) Authentication and access control: Practical cryptography methods and tools. Apress Media LLC
- Chen TY, Chen YM, Wang C Bin, Chu HC (2009) Flexible authorisation in dynamic e-business environments using an organisation structure-based access control model. *Int J Comput Integr Manuf* 22:225–244. <https://doi.org/10.1080/09511920802209041>
- Cook DM, Dissanayake DN, Kaur K (2019) The Usability factors of lost Digital Legacy data from regulatory misconduct: older values and the issue of ownership. *International Conference on Information and Communication Technology (ICICT)* 105
- Cupar D, Ivanović MD, Grgeč A (2023) Personal digital legacy: Findings from an exploratory study among citizens of Croatia. *Education for Information* 39:517–540. <https://doi.org/10.3233/EFI-230057>

- Dissanayake DN, Cook DM (2019) Social computing and older adults: Challenges with data loss and digital legacies. In: *Proceedings - 2019 International Conference on Cyberworlds, CW 2019*. Institute of Electrical and Electronics Engineers Inc., pp 171–174
- Fabry RE, Alfano M (2024) The Affective Scaffolding of Grief in the Digital Age: The Case of Deathbots. *Topoi*. <https://doi.org/10.1007/s11245-023-09995-2>
- Ferraiolo D, Kuhn DR (1997) Role-Based Access Control
- Figuerola-Torres M (2024) Affection as a service: Ghostbots and the changing nature of mourning. *Computer Law & Security Review* 52:105943. <https://doi.org/10.1016/J.CLSR.2024.105943>
- Gupta M, Bhatt S, Alshehri AH, Sandhu R (2022) *Access Control Models and Architectures For IoT and Cyber Physical Systems*. Springer International Publishing
- Harbinja E (2017) Post-mortem privacy 2.0: theory, law, and technology. *International Review of Law, Computers and Technology* 31:26–42. <https://doi.org/10.1080/13600869.2017.1275116>
- Harbinja E, Edwards L, McVey M (2023) Governing ghostbots. *Computer Law and Security Review* 48:. <https://doi.org/10.1016/j.clsr.2023.105791>
- Harbinja Edina (2022) *Digital Death, Digital Assets and Post-Mortem Privacy*. Edinburgh University Press
- Hevner, A.R., March, S.T., Park, J., Ram, S. (2004) Design Science in Information Systems Research. *MISQ*. 28:75–105. <https://doi.org/https://doi.org/10.2307/25148625>
- Jiménez-Alonso B, Brescò de Luna I (2023) Griefbots. A New Way of Communicating With The Dead? *Integr Psychol Behav Sci* 57:466–481. <https://doi.org/10.1007/s12124-022-09679-3>
- Larson JM (2022) *Snowflake Access Control: Mastering the Features for Data Privacy and Regulatory Compliance*. Springer
- Lindemann NF (2022) The Ethics of ‘Deathbots.’ *Sci Eng Ethics* 28. <https://doi.org/10.1007/s11948-022-00417-x>
- Mahalle PN, Bhong SS, Shinde GR (2022) *Authorization and Access Control*. In: *Authorization and Access Control*. CRC Press, pp 19–31
- Pal S (2021) *Internet of Things and Access Control: Sensing, Monitoring and Controlling Access in IoT-Enabled Healthcare Systems*. Springer Nature Switzerland AG, Cham
- Peffer K, Tuunanen T, Rothenberger MA, Chatterjee S (2007) A design science research methodology for information systems research. *Journal of Management Information Systems* 24:45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Peoples C, Hetherington M (2015) The Cloud Afterlife: Managing your Digital Legacy. *IEEE International Symposium on Technology and Society (ISTAS)* 1–7
- Puzio A (2023) When the Digital Continues After Death. *Ethical Perspectives on Death Tech and the Digital Afterlife*. *Communicatio Socialis* 56:427–436. <https://doi.org/10.5771/0010-3497-2023-3-427>
- Reese A (2023) The rise of grief tech. *New Sci* (1956) 41–43
- Salim F, Reid J, Dawson E (2010) *Towards Authorisation Models for Secure Information Sharing: A Survey and Research Agenda*. In Press
- Tsolkas A, Schmidt K (2017) *Rollen und Berechtigungskonzepte Identity- und Access-Management im Unternehmen*

ANALYZING THE DETERMINANTS OF HEALTHCARE TECHNOLOGY ADOPTION USING THE TASK-TECHNOLOGY FIT (TTF) MODEL: A SYSTEMATIC REVIEW AND META-ANALYSIS

AYESHA THANTHRIGE, NILMINI WICKRAMASINGHE

La Trobe University, School of Computing, Engineering and Mathematical Sciences,
Victoria, Australia
thanthrige87@gmail.com, n.wickramasinghe@latrobe.edu.au

This study aims to investigate the determinants of healthcare technology adoption using an extended Task-Technology Fit (TTF) model through a Systematic Literature Review (SLR) and meta-analysis, focusing on healthcare-specific tasks and their alignment with technology characteristics. While TTF has been widely applied across various domains, its application within healthcare is limited, with inconsistent findings. Addressing this research gap, the study provides a clearer understanding of how healthcare-specific tasks align with Technology Characteristics (TechC) to influence adoption among individuals. The extended model includes Behavioral Intention (BI) to assess users' intention to adopt healthcare technologies. The analysis reveals that TTF is a significant predictor of technology use, offering novel insights into the factors that drive successful healthcare technology adoption. The findings contribute to both theoretical advancements in TTF and offer practical implications for improving the design and implementation of digital healthcare solutions. Healthcare solution designers are encouraged to apply the TTF framework when evaluating new technologies to guide technology design and evaluation in real-world healthcare environments.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.15](https://doi.org/10.18690/um.fov.4.2025.15)

ISBN
978-961-286-998-4

Keywords:
task-technology fit (TTF),
technology adoption,
digital health solutions,
systematic literature review
(SLR),
technology characteristics
(TechC)



University of Maribor Press

1 Introduction

This study focuses on healthcare professionals broadly to capture diverse task requirements across various roles such as clinicians, nurses, and administrators, as TTF's applicability varies depending on specific tasks performed. The Task-Technology Fit (TTF) model offers an important framework for understanding how alignment between task requirements and technology functionality influences adoption. Originally introduced by Goodhue & Thompson (1995), TTF suggests that technology is more likely to be adopted when its capabilities match users' specific tasks. The better the fit between task characteristics (TC) and technology features, the more likely users find the technology beneficial, enhancing performance and increasing adoption likelihood. Unlike other adoption models, TTF uniquely focuses on alignment between tasks and technology capabilities. While the Technology Acceptance Model (TAM) emphasizes perceived usefulness and ease of use, and the Unified Theory of Acceptance and Use of Technology (UTAUT) examines performance expectancy, effort expectancy, social influence, and facilitating conditions, TTF specifically addresses how well technology features support task requirements. This distinctive focus makes TTF particularly valuable for understanding adoption in healthcare, where specific functionalities must closely match clinical workflows and patient management requirements. Healthcare heavily relies on technology to improve patient care, enhance clinical decision-making, and streamline efficiency. Technologies like Electronic Health Records (EHR), telemedicine platforms, mobile health applications, wearables, and health information systems are increasingly integrated into healthcare settings (Gu et al., 2021). Despite these advances, adoption and sustained use remain inconsistent (Alkhalifah & Bukar, 2022), suggesting further research is needed to explore determinants of technology adoption in healthcare contexts. TTF is particularly relevant to healthcare due to the diverse and complex tasks performed by healthcare professionals and patients. Clinicians manage patient records, coordinate care, conduct diagnostics, and ensure treatment compliance. Patients with chronic conditions monitor health metrics, adhere to medication schedules, track dietary intake, and manage appointments. A strong fit between these tasks and supporting technologies is critical for effective adoption and use (Janssen et al., 2021). While TTF has been applied in healthcare settings, its application remains limited compared to other domains, leading to inconsistent findings. For example, some studies suggest TTF significantly influences healthcare technology adoption, while

others indicate this relationship may be moderated by factors such as organizational support, user training, or technology complexity (Farivar et al., 2020). Additionally, less attention has been paid to TTF's role in predicting long-term usage. This research gap highlights the need for more comprehensive studies examining how healthcare-specific tasks align with technology features and how this alignment influences both adoption and sustained use (Wang et al., 2023). The application of TTF to healthcare technologies can provide valuable insights into how different stakeholders such as clinicians, patients, and healthcare administrators interact with digital health solutions. For clinicians, the fit between the technology and their tasks might relate to how well the technology supports clinical decision-making, patient monitoring, or data entry. For patients, particularly those managing chronic illnesses, technology needs to align with their daily health tasks, such as tracking blood sugar levels, managing diet, or scheduling medical appointments. Understanding these different dimensions of TTF can help identify the factors that drive successful adoption and long-term use, leading to better patient outcomes and more efficient healthcare delivery (Winckler, 2022). Despite the growing interest in digital health solutions, there remain significant challenges to their widespread adoption in healthcare. Many healthcare professionals and patients are hesitant to use new technologies due to concerns about ease of use, data security, and the perceived benefits of the technology. Additionally, organizational factors such as the availability of technical support, the provision of adequate training, and the compatibility of new technologies with existing systems can influence whether a technology is adopted or rejected (Lambert et al., 2023). Addressing these challenges requires a deeper understanding of how healthcare-specific tasks align with the functionality of the technologies being introduced.

2 Research Model and Hypotheses

This study focuses on understanding how TTF influences Behavioral Intention (BI) in healthcare contexts. The TTF model theorizes that individuals are more likely to use technology when they perceive a strong fit between their task requirements and the technology's functionalities. Our model extends the traditional TTF framework by incorporating BI, which has been widely used in models such as UTAUT. While the original TTF model focuses on how alignment between TC and Technology Characteristics (TechC) impacts performance, our extended model investigates how this alignment influences users' intention to adopt healthcare technologies. Including

BI provides a crucial link between TTF and actual technology adoption, particularly important in healthcare where successful implementation depends on both improved performance and user acceptance. This approach keeps the model focused on adoption behaviors while providing insights into how task-technology alignment influences engagement and usage intentions. Our model brings a unique perspective compared to TAM and UTAUT by emphasizing the fit between healthcare tasks and technology features rather than focusing primarily on perceptions or social factors. This task-technology alignment is particularly critical in clinical environments where workflow integration directly impacts adoption outcomes. By incorporating BI into the TTF framework, we capture both the functional fit aspects and the behavioral aspects of technology adoption in healthcare contexts. Our research examines relationships between TC, TechC, TTF, and BI in healthcare technology adoption (Figure 1). TC refers to healthcare-related tasks users need to perform, while TechC encompasses features and functionalities designed to support these tasks. TTF represents the degree to which technology features align with and support user task requirements, and BI reflects users' intention to adopt and use healthcare technologies. The association between TC and TTF is particularly important in healthcare settings where tasks are often complex, interdependent, and time sensitive. When tasks require significant information processing, coordination across multiple parties, or rapid decision-making, technology's ability to support these specific requirements becomes critical for adoption. Similarly, the alignment between TechC and TTF highlights how technology features directly influence perceived fit with healthcare tasks. This study excludes organizational context to maintain a focus on task-technology alignment, as organizational factors such as culture or support are secondary to the core TTF constructs.

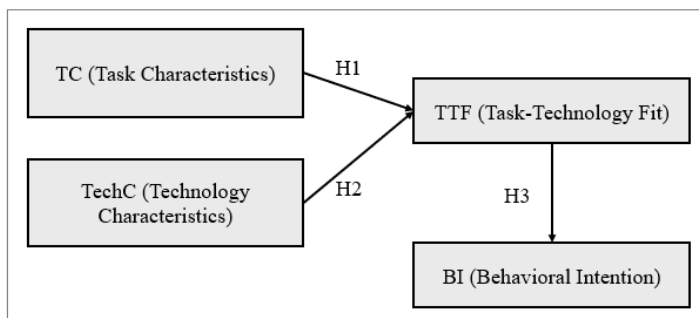


Figure 1: Research model

Based on this framework, we propose the following hypotheses:

H1: TC will be positively associated with TTF.

H2: TechC will be positively associated with TTF.

H3: TTF will be positively associated with BI.

3 Methodology

3.1 Research Question

How effectively does the TTF model predict healthcare technology adoption in the context of healthcare tasks?

3.2 Study Selection

This study follows a systematic literature review (SLR) and meta-analysis approach, adhering to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. Appendix B provides the PRISMA checklist, detailing how each item was addressed. The methodology was structured based on prior meta-analytic practices to ensure comprehensive and unbiased analysis (Dwivedi et al., 2019). We aimed to review studies focusing on the TTF model's application in healthcare settings, including the relationships between TC, TechC, TTF, and BI. The search strategy involved querying multiple databases, including Scopus, Web of Science, IEEE Xplore, ScienceDirect, PubMed, and Google Scholar. These databases were selected based on their comprehensive coverage of technology adoption and healthcare literature. Keywords included combinations of "Task-Technology Fit", "TTF", "healthcare technology adoption", "mHealth", "telehealth", and "technology characteristics" and "chronic disease management" to capture studies relevant to chronic conditions such as diabetes. Controlled vocabularies (e.g., MeSH terms in PubMed) were used to enhance search accuracy. Hand-searching of bibliographies was also conducted to identify additional relevant studies. Search results were imported into the reference management software EndNote for the removal of duplicates, after which they were screened using Covidence software to manage and streamline the review process. The initial

phase involved screening titles and abstracts for relevance against the inclusion and exclusion criteria. To ensure unbiased selection, titles and abstracts were screened independently by two reviewers against inclusion and exclusion criteria, with discrepancies resolved through discussion or by a third reviewer.

Studies were included if they met the following criteria:

- Focused on healthcare technologies, such as EHR, mHealth applications, telehealth, or other digital health innovations, including those supporting chronic disease management.
- Published in English between 2012 and 2025.
- Provided quantitative data, including sample sizes, standardized path coefficients (β), and reliability statistics such as Composite Reliability (CR) or Cronbach's α .
- Examined relationships related to TTF, TC, TechC, BI.

Exclusion criteria were:

- Studies not focused on TTF or healthcare technology adoption.
- Non-quantitative studies or those lacking standardized β coefficients.
- Studies published before 2012 or not in English.

A full-text review of screened articles was conducted by two independent reviewers, with discrepancies resolved by a third reviewer to ensure impartiality. Our analysis identified 15 studies meeting all inclusion criteria.

3.3 Coding Data

For each included study, we extracted publication details, study characteristics, and quantitative data focusing on relationships between independent and dependent variables, with particular attention to the TTF model. Appendix C provides the data extraction template. We harmonized constructs with similar conceptualizations but different labels to maintain consistency across studies. Our analysis included 15 studies providing 43 unique path coefficients. While this sample is smaller than ideal for meta-analysis, it represents the current state of quantitative research explicitly

applying TTF in healthcare contexts. The meta-analysis of these studies offers valuable preliminary insights while highlighting the need for more research in this area.

3.4 Statistical Analysis

This approach is appropriate given the variability in study designs, populations, and healthcare contexts represented. Meta-analysis is particularly valuable for this study as it allows us to systematically combine findings across diverse studies, revealing patterns that might not be apparent in individual studies. By statistically synthesizing results, meta-analysis provides more precise estimates of effect sizes and identifies sources of heterogeneity, highlighting contextual factors that might influence relationships between key constructs. While our sample size is limited, meta-analysis still provides valuable insights by systematically integrating available evidence. The meta-analysis was conducted using R Software.

4 Results

4.1 Study Identification and Screening

The systematic review began by identifying 691 potential studies across various databases, including Scopus (41%), Science Direct (22%), Web of Science (6%), CINAHL (6%), IEEE Xplore (3%), PubMed (19%), and Google Scholar (3%). Following the removal of duplicates (171, 25%), the titles and abstracts of 520 articles were reviewed to assess their relevance to the TTF model in healthcare technology adoption. After this initial screening, 361 full-text articles were reviewed, resulting in 15 studies (4% of the full-text articles assessed) being included in the final meta-analysis. Figure 2 provides the PRISMA flow diagram, summarizing the study selection process.

4.2 Descriptive Analysis

The included studies span several geographical regions and methodologies, offering insights into TTF's role in healthcare technology adoption. The range of technologies studied includes EHR, mHealth apps, telemedicine, and other digital health platforms, with several studies focusing on chronic disease management, such

as diabetes care. Appendix A summarizes key information for each included study, including objectives, study design, sample size, outcomes, key findings, and statistical results, with detailed β values. Age and gender were the most consistently reported demographic factors, with 13 and 14 studies respectively providing data on these categories, while variables such as education, occupation, income, nationality, and usage experience were less frequently reported.

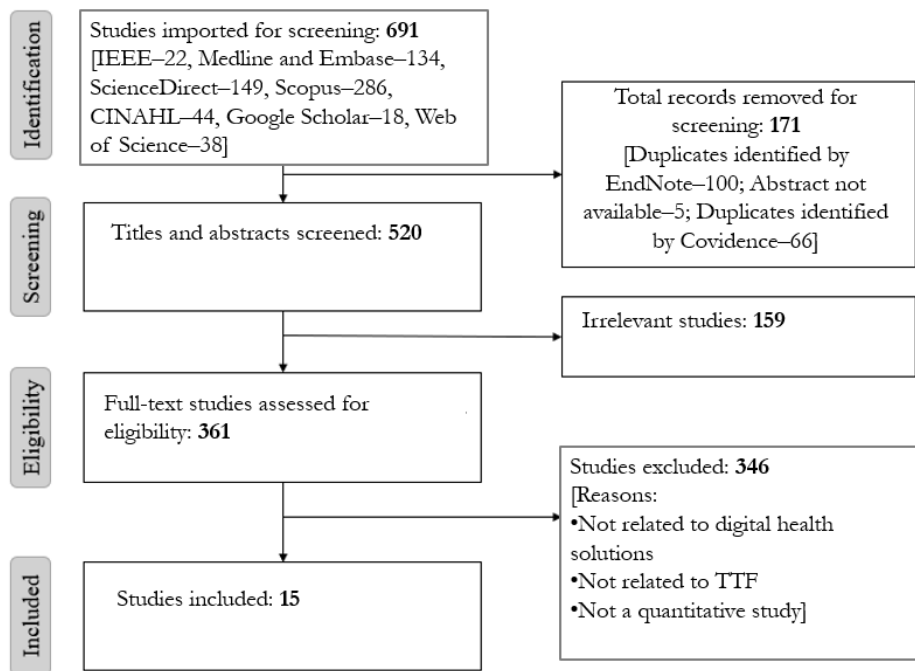


Figure 2: PRISMA Flow Diagram

4.3 Path Coefficients and Statistical Significance

Table 1 summarizes the β reported across the included studies, focusing on the relationships between TC, TechC, TTF, BI. The β exhibited significant variation across the studies. For example, the relationship between TC and TTF (H1) ranged from -0.007 to 0.525, while the relationship between TechC and TTF (H2) ranged from 0.199 to 0.780. Similarly, the relationship between TTF and BI (H3) ranged from -0.209 to 0.712, indicating both positive and negative associations across different contexts. These variations suggest that the impact of task and TechC on

T^{*}TTF, and subsequently on BI, is highly context dependent. The differences in these β highlight how diverse study settings and technology implementations can affect the fit between tasks and technology, as well as the intention to use the technology.

The number of studies examining each relationship showed slight differences. For example, 15 studies analyzed the relationship between TC and TTF (H1). However, 14 studies explored the relationship between TechC and TTF (H2) and an equal number (14) examined the link between T^{*}TTF and BI (H3). Most studies supported the hypothesized relationships, with higher than 90% of the β demonstrating statistical significance at $p < 0.01$. Specifically, 93% of the studies examining the relationship between TC and TTF (H1) found it to be positive and significant, while 100% of the studies analyzing the relationship between TechC and TTF (H2) supported the hypothesis. Additionally, 93% of the studies exploring the relationship between T^{*}TTF and BI (H3) found it statistically significant, reinforcing the importance of T^{*}TTF in predicting BI. Table 1 offers a comprehensive overview of the path coefficients, their significance, and the average effect sizes across the studies.

Table 1: Summary of Path Coefficients, Sample Sizes, Significance, and Weight Analysis

Path	n	Range of β values		Avg β	Sample size				Positive sig β		Negative sig β	
					Min	Max	Avg	Total	No.	%	No.	%
Tech C - TTF	15	0.199	0.780	0.438	102	487	262.3	4041	15	100%	0	0%
TC-TTF	14	-0.007	0.525	0.291	102	487	253.7	3658	13	93%	1	7%
TTF - BI	14	-0.209	0.712	0.327	113	487	273.8	3939	13	93%	1	7%

n = Number of studies; AVG β = Arithmetic mean of β values; MIN = Minimum; MAX = Maximum; AVG = Average values; Total= Total sample size; Sig = Significance

4.4 Meta-Analysis Outcomes

The meta-analysis results (see Table 2) confirmed the hypothesized relationships in the T^{*}TTF model. Figure 3 presents the meta-analytic outcomes, highlighting the relative strength of the relationships within the T^{*}TTF model. The relationship between TechC and TTF (H2) was the strongest, with a meta-analytic effect size of $\beta = 0.445$ ($p < 0.001$), indicating a robust positive relationship across the studies. The relationship between TTF and BI (H3) also demonstrated a significant effect,

with $\beta = 0.271$ ($p < 0.001$). Although the relationship between TC and TTF (H1) was weaker, it remained statistically significant, with $\beta = 0.263$ ($p < 0.001$). The heterogeneity tests revealed high variability across studies, with I^2 values ranging from 86.15% to 94.87%, suggesting that contextual factors may influence the strength of these relationships. Overall, the meta-analysis provides strong support for the hypothesized relationships in the TTF model, reinforcing its relevance in understanding healthcare technology adoption.

Table 2: Meta-Analysis of Path Coefficients, Total Sample Sizes, Significance, and Confidence Intervals

Path	n	TSS	Meta β	p-value β	95% CI β		Heterogeneity test		
					Lower	Upper	Q-value	df (Q)	I^2 (%)
TechC -TTF	15	4041	0.445	0.00	0.414	0.476	155.61	14	91.00
TC -TTF	14	3658	0.263	0.00	0.231	0.296	93.83	13	86.15
TTF - BI	14	3939	0.271	0.00	0.237	0.299	253.29	13	94.87

n = No. of occurrences; TSS = Total sample size; Meta β = Weighted mean effect size; CI = Confidence interval, Q - Total amount of heterogeneity

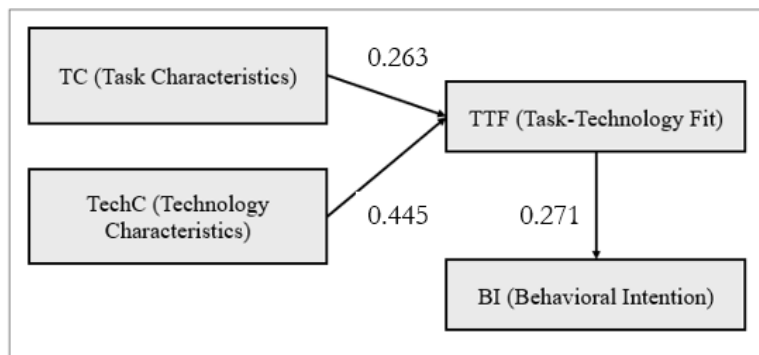


Figure 3: The meta-analytic outcomes

5 Discussion

5.1 Theoretical Contributions

This study contributes to technology adoption theory by validating and extending the TTF model in healthcare contexts. By confirming relationships between TC, TechC, TTF, and BI, we demonstrate TTF's value as a framework for understanding

healthcare technology adoption. The significant influence of TTF on BI shows that task-technology alignment is crucial for adoption decisions, complementing insights from other adoption models such as TAM and UTAUT. Compared to prior studies, our findings align with Wang et al. (2020) and Tao et al. (2023) which emphasize TTF's role in mHealth and wearable adoption, particularly for chronic disease management, but extend these by integrating BI to capture user intentions more explicitly. Our findings highlight the relative importance of technology characteristics compared to task characteristics in determining TTF. In healthcare settings, technology design and functionality appear to play a more influential role in task-technology alignment than the inherent characteristics of healthcare tasks. The high heterogeneity in TechC-TTF, $I^2=91.00\%$ and TTF-BI, $I^2=94.87\%$ relationships suggests that study-specific factors, such as technology type (e.g., EHR vs. mHealth) or user demographics, moderate these effects, warranting further investigation into contextual influences (Howard et al., 2019). This finding carries important implications for technology design and implementation. The substantial heterogeneity observed across studies indicates that TTF's application in healthcare is context dependent. While our sample size is limited, this analysis represents an important first step in systematically examining TTF in healthcare contexts. The consistent patterns observed across our sample suggest that these relationships are robust, though further research with larger samples is needed to strengthen these conclusions.

5.2 Practical Implications

This study provides valuable practical insights for healthcare organizations, decision-makers, and technology providers aiming to improve the adoption and use of healthcare technologies. The research emphasizes the importance of aligning healthcare technologies with the specific tasks of healthcare professionals, such as EHR or telemedicine platforms, to enhance TTF and drive higher adoption rates. For instance, technologies supporting chronic disease management, like diabetes-focused mHealth apps, must align with tasks such as glucose monitoring and medication adherence to improve adoption. User-centered design is also critical; technologies must prioritize ease of use, interoperability, and seamless integration into clinical workflows to reduce cognitive load and increase both BI and actual usage. Comprehensive training and ongoing support are essential to ensure that even well-fitted technologies are adopted successfully. Training should focus on skill

development while providing continuous technical support to address potential resistance due to unfamiliarity with new systems. The study also highlights the need for continuous monitoring and adaptation of technologies to maintain alignment with evolving healthcare tasks, ensuring that technologies remain relevant and useful over time. By incorporating BI into the TTF framework, healthcare organizations can more accurately predict adoption by not only evaluating task-technology alignment but also understanding users' intentions, allowing for more targeted strategies. Decision-makers must also balance feature complexity with usability, ensuring technologies are equipped with essential, task-aligned features without overwhelming users. Addressing these factors can significantly improve technology adoption, enhance patient outcomes, and boost operational efficiency, especially in critical areas like chronic disease management.

5.3 Limitations and Future Research

Despite the valuable insights provided by this meta-analysis, several limitations should be noted. Studies were excluded if they did not focus on TTF, did not relate to healthcare technology adoption, or were not quantitative studies. The relatively small number of research studies found that applying the TTF model in the context of healthcare technology adoption, with only 15 studies meeting the inclusion criteria, highlights a significant gap in literature. This limited number of studies may reflect the emerging application of TTF in healthcare and underscores the need for more quantitative studies to validate these findings. The study protocol was not registered, which may limit transparency. The absence of a formal methodological quality assessment of included studies, due to the focus on quantitative data, may affect the credibility of results. The predominance of studies from Asia, particularly Taiwan, may limit generalizability, as regional differences in infrastructure, policy, or digital literacy could influence adoption outcomes. This finding highlights a significant gap in the literature regarding quantitative applications of TTF in healthcare contexts. The small sample emphasizes the need for more research explicitly applying TTF to the adoption of healthcare technology. The study focused primarily on quantitative studies that reported standardized β and other statistical data related to TTF. As a result, qualitative insights into how task and TechC influence healthcare technology adoption were not included.

6 Conclusion

This systematic review and meta-analysis provide valuable insights into TTF's role in healthcare technology adoption. By extending TTF to incorporate BI, we bridge theoretical perspectives on technology alignment and adoption decisions. Our findings confirm that TTF significantly influences adoption intentions in healthcare settings, with technological characteristics playing a particularly important role in determining the fit. The results highlight the importance of designing healthcare technologies that align with specific task requirements, whether for healthcare professionals or patients. Implementation strategies should emphasize this alignment to enhance adoption. While our study advances understanding of healthcare technology adoption through the TTF lens, it also reveals the limited application of TTF in quantitative healthcare technology research. This gap presents important opportunities for future research to expand the evidence base and further refine our understanding of how task-technology fit influences healthcare technology adoption. As healthcare systems increasingly rely on digital technologies, understanding the determinants of successful adoption becomes increasingly important. The task-technology fit perspective offers valuable insights into how alignment between technology capabilities and healthcare tasks can drive adoption and ultimately improve healthcare delivery.

References

- Al-Emran, M. (2021). Evaluating the Use of Smartwatches for Learning Purposes through the Integration of the Technology Acceptance Model and Task-Technology Fit. *International Journal of Human-Computer Interaction*, 37(19), 1874-1882. <https://doi.org/10.1080/10447318.2021.1921481>
- Alhendawi, K. M. (2022). Task-technology fit model: Modelling and assessing the nurses' satisfaction with health information system using AI prediction models [Article]. *International Journal of Healthcare Management*. <https://doi.org/10.1080/20479700.2022.2136881>
- Aljarboa, S., & Miah, S. J. (2020, 16-18 Dec. 2020). Assessing the Acceptance of Clinical Decision Support Tools using an Integrated Technology Acceptance Model. 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE),
- Alkhalifah, A., & Bakar, U. A. (2022). Examining the Prediction of COVID-19 Contact-Tracing App Adoption Using an Integrated Model and Hybrid Approach Analysis. *Frontiers in public health*, 10, 847184. <https://doi.org/https://dx.doi.org/10.3389/fpubh.2022.847184>
- Alkhwaldi, A. F., & Abdulmuhsin, A. A. (2022). Understanding User Acceptance of IoT Based Healthcare in Jordan: Integration of the TTF and TAM. *Studies in Computational Intelligence*,
- Antoniadis, K., Zafropoulos, K., & Mitsiou, D. (2022). Measuring Distance Learning System Adoption in a Greek University during the Pandemic Using the UTAUT Model, Trust in

- Government, Perceived University Efficiency and Coronavirus Fear [Article]. *Education Sciences*, 12(9), Article 625. <https://doi.org/10.3390/educsci12090625>
- Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the Unified Theory of Acceptance and Use of Technology (UTAUT): Towards a Revised Theoretical Model. *Information Systems Frontiers*, 21(3), 719-734. <https://doi.org/10.1007/s10796-017-9774-y>
- Edo, O. C., Ang, D., Etu, E.-E., Tenebe, I., Edo, S., & Diekola, O. A. (2023). Why do healthcare workers adopt digital health technologies - A cross-sectional study integrating the TAM and UTAUT model in a developing economy. *International Journal of Information Management Data Insights*, 3(2), 100186. <https://doi.org/https://doi.org/10.1016/j.jjimei.2023.100186>
- El-Masri, M., Al-Yafi, K., & Kamal, M. M. (2023). A Task-Technology-Identity Fit Model of Smartwatch Utilisation and User Satisfaction: A Hybrid SEM-Neural Network Approach. *Information systems frontiers : a journal of research and innovation*, 25(2), 835-852. <https://doi.org/https://dx.doi.org/10.1007/s10796-022-10256-7>
- El Kurdi, B., Babar, S., El Iskandarani, M., Bataineh, A., Lerch, M. M., Young, M., & Singh, V. P. (2019). Factors That Affect Prevalence of Small Intestinal Bacterial Overgrowth in Chronic Pancreatitis: A Systematic Review, Meta-Analysis, and Meta-Regression. *Clinical and translational gastroenterology*, 10(9), e00072-e00072. <https://doi.org/10.14309/ctg.0000000000000072>
- 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>
- Gu, D., Khan, S., Khan, I. U., Khan, S. U., Xie, Y., Li, X., & Zhang, G. (2021). Assessing the Adoption of e-Health Technology in a Developing Country: An Extension of the UTAUT Model [Article]. *SAGE Open*, 11(3). <https://doi.org/10.1177/21582440211027565>
- Habeeb, R. J. H., Hilles, S. M. S., & Momani, A. M. (2021). Social Commerce Adoption Based UTAUT Model for Consumer Behavior: Iraq Small and Medium Enterprise. 2nd International Informatics and Software Engineering Conference, IISEC 2021,
- Howard, M. C., & Rose, J. C. (2019). Refining and extending task–technology fit theory: Creation of two task–technology fit scales and empirical clarification of the construct. *Information & Management*, 56(6), 103134. <https://doi.org/10.1016/j.im.2018.12.002>
- Hsiao, J. L., & Chen, R. F. (2012). An investigation on task-technology fit of mobile nursing information systems for nursing performance [Article]. *CIN - Computers Informatics Nursing*, 30(5), 265-273. <https://doi.org/10.1097/NCN.0b013e31823eb82c>
- Janssen, A., Keep, M., Selvadurai, H., Kench, A., Hunt, S., Simonds, S., Marshall, T., Hatton, L., Dalla-Pozza, L., McCullagh, C., & Shaw, T. (2021). Factors That Influence Use of a Patient Portal by Health Professionals. *International Journal of Environmental Research and Public Health*, 18(4). <https://doi.org/https://dx.doi.org/10.3390/ijerph18041877>
- Kamal, M., & Subriadi, A. P. (2021). UTAUT Model of Mobile Application: Literature Review. *Proceedings - IEIT 2021: 1st International Conference on Electrical and Information Technology*,
- Lambert, S. I., Madi, M., Sopka, S., Lenes, A., Stange, H., Buszello, C.-P., & Stephan, A. (2023). An integrative review on the acceptance of artificial intelligence among healthcare professionals in hospitals. *NPJ digital medicine*, 6(1), 111. <https://doi.org/https://dx.doi.org/10.1038/s41746-023-00852-5> (Erratum in: *NPJ Digit Med*. 2023 Jul 11;6(1):125 PMID: 37433964 [<https://www.ncbi.nlm.nih.gov/pubmed/37433964>])
- Semiz, B. B., & Semiz, T. (2021). Examining consumer use of mobile health applications by the extended UTAUT model. *Business & management studies: an international journal*, 9(1), 267-281. <https://doi.org/10.15295/bmij.v9i1.1773>
- Shapira, S., & Cauchard, J. R. (2022). Integrating drones in response to public health emergencies: A combined framework to explore technology acceptance. *Frontiers in public health*, 10, 1019626. <https://doi.org/https://dx.doi.org/10.3389/fpubh.2022.1019626>

- Tao, D., Chen, Z., Qin, M., & Cheng, M. (2023). Modeling Consumer Acceptance and Usage Behaviors of m-Health: An Integrated Model of Self-Determination Theory, Task–Technology Fit, and the Technology Acceptance Model [Article]. *Healthcare (Switzerland)*, 11(11), Article 1550. <https://doi.org/10.3390/healthcare11111550>
- Wang, H., Tao, D., Yu, N., & Qu, X. (2020). Understanding consumer acceptance of healthcare wearable devices: An integrated model of UTAUT and TTF [Article]. *International Journal of Medical Informatics*, 139, Article 104156. <https://doi.org/10.1016/j.ijmedinf.2020.104156>
- Wang, S. L., & Lin, H. I. (2019). Integrating TTF and IDT to evaluate user intention of big data analytics in mobile cloud healthcare system [Article]. *Behaviour and Information Technology*, 38(9), 974-985. <https://doi.org/10.1080/0144929X.2019.1626486>
- Wang, X., Lee, C. F., Jiang, J., Zhang, G., & Wei, Z. (2023). Research on the Factors Affecting the Adoption of Smart Aged-Care Products by the Aged in China: Extension Based on UTAUT Model [Article]. *Behavioral Sciences*, 13(3), Article 277. <https://doi.org/10.3390/bs13030277>
- Williams, M. D., Rana, N. P., & Dwivedi, Y. K. (2015). The unified theory of acceptance and use of technology (UTAUT): a literature review. *Journal of Enterprise Information Management*, 28(3), 443-488. <https://doi.org/10.1108/JEIM-09-2014-0088>
- Winckler, D. (2022). Not another box to check! Using the UTAUT to explore nurses' psychological adaptation to electronic health record usability. *Nursing Forum*, 57(3), 412-420. <https://doi.org/https://dx.doi.org/10.1111/nuf.12686>
- Yamin, M. A. Y., & Alyoubi, B. A. (2020). Adoption of telemedicine applications among Saudi citizens during COVID-19 pandemic: An alternative health delivery system. *Journal of infection and public health*, 13(12), 1845-1855. <https://doi.org/https://doi.org/10.1016/j.jiph.2020.10.017>

Appendix A- Data used for analysis

ID	Author & Year	Aims & Objectives	Sample Size	Demographic Variables	Path Coefficient (Beta)	Study Design	Country	Methodology	Key Findings
1	Hallang Wang et al., 2020	To develop and empirically test an integrated model of UTAUT and TTF to understand consumer acceptance of healthcare wearable devices (HWDs).	406	Gender, Age, Education, Occupation, HWD usage experience	Tech - TTF: 0.726 TC - TTF: 0.118 TTF - BI : 0.219	Cross-sectional survey design	China	Self-administered questionnaire survey	The study shows that performance expectancy, effort expectancy, social influence, facilitating conditions, and task-technology fit positively influence the acceptance of healthcare wearable devices (HWDs). These factors explained 68% of the variance in behavioral intention to use HWDs.
2	Shu Lin Wang & Hsian I Lin, 2019	To integrate Task-Technology Fit (TTF) and Innovation Diffusion Theory (IDT) to evaluate young users' intention to use a mobile cloud healthcare system for diabetes preventive care.	423	Age (15-40 years), Gender, Education Level	Tech - TTF: 0.780 TC - TTF : 0.193 TTF - BI : 0.407	Cross-sectional survey design	Taiwan	Questionnaire survey	Task characteristics, technology characteristics, and relative advantage positively influence task-technology fit, which in turn influences intention to use the mobile healthcare system. Observability had no significant effect.
3	Kamal Mohammed Alhendawi, 2022	To extend the TTF model by incorporating nurses' attitudes and evaluate the predictive power of regression and neural network models for nurse satisfaction with HIS.	164	Age, Gender, Occupation	Tech - TTF: 0.494 TC - TTF : 0.261 TTF - BI : 0.697	Cross-sectional survey	Palestine (Gaza Strip)	Questionnaire survey with AI prediction models	Task and technology characteristics, along with nurses' attitudes and task-technology fit, influence nurses' satisfaction with HIS. ANN model outperforms regression in predictive accuracy.
4	Mohammadiwa AbdeRhoda, Afsaruch Delmad, Javad Zarei, 2022	To identify the key determining factors influencing faculty members' intention to adopt e-learning during COVID-19 by integrating UTAUT and TTF models.	143	Gender, Age, Teaching Experience	Tech - TTF: 0.652 TC - TTF : 0.525 TTF - BI : 0.244	Descriptive-analytical, cross-sectional survey	Iran (Alhwaz University of Medical Sciences)	Questionnaire survey	Technology and task characteristics, task-technology fit, social influences, effort expectancy, performance expectancy, and facilitating conditions positively influence e-learning adoption.
5	Yvonne O'Connor, Pavel Andreev, Philip O'Reilly, 2020	To develop and validate a conceptual model exploring how mHealth affects physicians' perceived quality of care (PQoC) in a hospital setting.	102	Age, Gender, Professional role	Tech - TTF: 0.479 TC - TTF : 0.337	Quasi-experimental posttest-only	Canada	Observational survey	Task-technology fit (TTF) and mHealth utilization positively impact physicians' perceived quality of care (PQoC). Self-efficacy is crucial in mHealth utilization.

ID	Author & Year	Aims & Objectives	Sample Size	Demographic Variables	Path Coefficient (Beta)	Study Design	Country	Methodology	Key Findings
6	Ali Alkhalifah, Umar Ali Bakar, 2022	To predict the factors influencing adoption of the Tawakkalna COVID-19 contact-tracing app in Saudi Arabia using TAM, PCT, and TTF, and validate using SEM and ANN.	309	Age, Gender, Nationality	Tech - TTF: 0.430 TC - TTF : 0.450 TTF - BI : 0.269	Cross-sectional survey design	Saudi Arabia	Questionnaire survey	Perceived ease of use, usefulness, task features, and mobility positively impact intention to adopt the app; privacy risk was not significant
7	Lianra Wijaya, Kah Choon Ng, Paudoman Robinson Sihombing, 2023	To explore factors affecting the continuance usage intention (CUI) of telemedicine apps in Jakarta using Task-Technology Fit (TTF) theory.	137	Gender, Age, Education Level	Tech - TTF: 0.396 TC - TTF : 0.104 TTF - BI : 0.191	Cross-sectional survey	Indonesia	Online questionnaire survey	Technology characteristics and self-efficacy significantly impact telemedicine continuance usage, while task characteristics and awareness have minimal effects.
8	Mazen El-Masri, Katim Al-Yafi, Muhammad Mustafa Kamal, 2023	To develop and validate a Task-Technology-Identity Fit (TTF) model that predicts smartwatch utilisation and satisfaction using both SEM and ANN analysis.	248	Gender, Age, Occupation, Nationality	Tech - TTF: 0.382 TC - TTF : 0.382 TTF - BI : 0.680	Cross-sectional survey design	Qatar	Questionnaire survey	Task-technology fit and technology-identity fit both significantly affect user satisfaction with smartwatches. Identity fit has a stronger effect on satisfaction, while actual task fit impacts smartwatch choice.
9	Tung-Cheng Lin, 2014	To investigate the effectiveness of mobile nursing information systems (MNIS) using an extended TTF model that includes organizational readiness and separates TTF into task-technology fit (TaTeF) and technology-individual fit (TeIF).	144	Gender, Education, Work Experience	Tech - TTF: 0.298 TC - TTF : 0.352 TTF - BI : 0.300	Cross-sectional survey	Taiwan	Questionnaire survey	Technology-individual fit and organizational readiness significantly influence MNIS usage; organizational support plays a crucial role in adoption.
10	Mohammad Ali Yousef Yamin, Bader A. Aljroubi, 2020	To investigate Saudi citizens' behavioral intention to adopt wireless sensor network (WSN)-based telemedicine apps during COVID-19, extending UTAUT with TTF.	348	Age, Gender	Tech - TTF: 0.209 TC - TTF : 0.489 TTF - BI : 0.508	Cross-sectional survey	Saudi Arabia	Online questionnaire survey	Task-technology fit and facilitating conditions are crucial for the adoption of telemedicine apps based on wireless sensor networks; R ² explains 79.5% of variance in adoption behavior.
11	Muhammad Shahbaz et al., 2021	To propose a big data analytics (BDA) system for environmental air pollution (EAP) management and examine user adoption behavior using TTF and UTAUT frameworks.	412	Gender, Age, Education Level	Tech - TTF: 0.267 TC - TTF : 0.284 TTF - BI : 0.140	Cross-sectional survey	Pakistan	Structured questionnaire survey	Task-technology fit, along with UTAUT factors, strongly influences the adoption of the BDA-EAP management system. Resistance to change negatively impacts adoption, while extrinsic motivation positively moderates adoption.

ID	Author & Year	Aims & Objectives	Sample Size	Demographic Variables	Path Coefficient (Beta)	Study Design	Country	Methodology	Key Findings
12	Ju-Ling Hsiao, Rai-Fu Chen, 2012	To investigate the relationship between nursing task characteristics, m-NIS characteristics, TTF, and nursing performance using a modified TTF model.	219	Gender, Age, Work Experience	Tech - TTF : 0.270 TC - TTF : 0.308 TTF - BI : 0.312	Cross-sectional survey design	Taiwan	Questionnaire survey	Task-technology fit (TTF) positively influences nursing performance through factors such as information identification, acquisition, integration, and interpretation.
13	Hyo-Jin Kang, Jieun Han, Gyu Hyun Kwon, 2022	To analyze the user acceptance behavior of smart home health care services (SHHSs) in South Korea using an integrated UTAUT and TTF model.	487	Gender, Age, Marital Status	Tech - TTF : 0.517 TC - TTF : -0.007 TTF - BI : -0.209	Cross-sectional survey	South Korea	Online survey	Performance expectancy, effort expectancy, social influence, and task-technology fit significantly influence the behavioral intention and adoption of SHHS. Task characteristics did not affect task-technology fit.
14	Al-Rahmi et al., 2022	To integrate UTAUT and TTF models to evaluate social media use in teaching and learning in higher education and assess their impact on academic performance.	383	Gender, Age (18<41+), Discipline (Engineering, Science, Management, Social Sciences)	Tech - TTF : 0.199 TTF - BI : 0.107	Survey-based SEM (SmartPLS 3.3.3)	Malaysia	Structural Equation Modeling integrating UTAUT and TTF	Task-Technology Fit (TTF) and Behavioral Intention (BI) significantly impact Academic Performance (AP), PEX, EEX, Social Characteristics, and Technology Characteristics influence TTF and BI. Strong interrelationships among UTAUT and TTF constructs improve student performance via social media use.
15	Hsieh & Lin, 2020	To examine the performance impact and user resistance behavior toward the epidemic prevention cloud (EPC) using an integrated Task-Technology Fit (TTF) and Status Quo Bias (SQB) model.	116	Gender, Age, Education, Work Position, Work Experience	Tech - TTF : 0.474 TC - TTF : 0.277 TTF - BI : 0.712	Cross-sectional field survey using SmartPLS	Taiwan	Integrated TTF and Status Quo Bias (SQB) theoretical model	Task and technology characteristics positively influence TTF. TTF enhances utilisation and performance while reducing resistance to use. Resistance is influenced by uncertainty costs (positive), perceived value (negative), but not sunk costs. Resistance negatively affects utilisation but not performance. Utilisation and TTF significantly affect performance.

Appendix B

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Analyzing the Determinants of Healthcare Technology Adoption Using the Task-Technology Fit (TTF) Model: A Systematic Review and Meta-Analysis.
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	The abstract follows PRISMA 2020 for Abstracts guidelines, providing a structured summary including background (TTF's limited application in healthcare), objectives (investigate determinants of healthcare technology adoption), methods (SLR and meta-analysis), results (TTF as a significant predictor), and conclusions (theoretical and practical implications).
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Section 1 describes the rationale, highlighting TTF's limited application in healthcare compared to other domains, inconsistent findings, and the need to understand task-technology alignment for adoption
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Section 1 explicitly states the objective: to investigate determinants of healthcare technology adoption using an extended TTF model via SLR and meta-analysis, focusing on healthcare-specific tasks and technology characteristics. The research question is specified in Section 3.1.
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Section 3.2 lists inclusion criteria (healthcare technologies, English, 2012–2025, quantitative data with β coefficients) and exclusion criteria (non-TTF, non-healthcare, non-quantitative, pre-2012, non-English). Studies were grouped by TTF relationships (TC, TechC, TTF, BI) for meta-analysis
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Section 3.2 specifies databases searched: Scopus, Web of Science, IEEE Xplore, ScienceDirect, PubMed, and Google Scholar. The paper does not specify the exact date of the last search, but studies span 2012–2025.
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	keywords ("Task-Technology Fit," "TTF," "healthcare technology adoption," "mHealth," "telehealth," "technology characteristics," "chronic disease management") and use of controlled vocabularies (e.g., MeSH terms in PubMed). Hand-searching of bibliographies was also conducted to identify additional relevant studies
Selection process	8	Specify the methods used to decide whether a	Section 3.2 describes the selection process: two reviewers independently screened titles and abstracts

Section and Topic	Item #	Checklist item	Location where item is reported
		study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	using Covidence software, with discrepancies resolved by discussion or a third reviewer. Full-text reviews were also conducted by two independent reviewers. No automation tools were used.
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Section 3.3 outlines data collection: two reviewers extracted publication details, study characteristics, and quantitative data (e.g., β coefficients) independently, using a template (Appendix A). No contact with study investigators is mentioned. Covidence and EndNote were used for data management
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Section 3.3 defines outcomes sought: β coefficients for relationships between TC, TechC, TTF, and BI. All compatible results (e.g., standardized path coefficients) were sought from each study
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Additional variables: publication details, study characteristics (e.g., technology type), sample size, reliability statistics (CR or Cronbach's α). Constructs with similar labels were harmonized. No assumptions about missing data are explicitly stated.
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if	Section 5.3 notes no formal methodological quality assessment was conducted due to the focus on quantitative data, a limitation. No specific tools or independent reviewer processes for bias assessment are described.

Section and Topic	Item #	Checklist item	Location where item is reported
		applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Section 4.4 specifies the effect measure: standardized path coefficients (β) for TC-TTF, TechC-TTF, and TTF-BI relationships, with significance levels ($p < 0.01$) and confidence intervals in Table 2.
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Section 3.2 describes eligibility for synthesis: studies providing β coefficients for TTF relationships were included, grouped by path (TC-TTF, TechC-TTF, TTF-BI).
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Section 3.3 notes data preparation: constructs with similar conceptualizations were harmonized to ensure consistency. No handling of missing statistics is described.
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Section 4 presents results visually in tables (Table 1 for β coefficients, Table 2 for meta-analysis results).
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Section 3.4 describes synthesis: meta-analysis used a random-effects model to estimate weighted mean β values, with heterogeneity assessed via I ² and Q statistics. The choice of meta-analysis is justified by the need to combine diverse study findings. meta-analysis was conducted using R.
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Section 5.1 explores heterogeneity causes, suggesting technology type (e.g., EHR vs. mHealth) and demographics as moderators, but no formal subgroup analysis or meta-regression is conducted.
	13f	Describe any sensitivity analyses conducted to assess robustness of the	No sensitivity analyses are described to assess the robustness of synthesized results.

Section and Topic	Item #	Checklist item	Location where item is reported
		synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	No methods are described to assess reporting bias (e.g., publication bias via funnel plots or Egger's test), a limitation not explicitly noted
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	No formal certainty assessment (e.g., GRADE) is described. Section 5.3 acknowledges the small sample size (15 studies) and lack of quality assessment as limiting result credibility.
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Section 4.1 describes the search process: 691 records identified, 171 duplicates removed, 520 screened, 361 full-text reviewed, and 15 studies included. Figure 2 (PRISMA Flow Diagram) visualizes this.
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Section 4.1 does not cite specific excluded studies or reasons beyond general criteria (e.g., non-TTF, nonquantitative)
Study characteristics	17	Cite each included study and present its characteristics.	Section 4.2 and Appendix A cite all 15 included studies and present characteristics (e.g., sample size, β coefficients), objectives, design, outcomes, and findings.
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	No risk of bias assessments are presented for included studies, consistent with the limitation in Section 5.3
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g., confidence/credible interval), ideally using structured tables or plots.	Section 4.3 and Appendix A present β coefficients, sample sizes, and significance for each study. Table 1 summarizes ranges, averages, and significance. Table 2 provides meta-analytic β estimates with 95% confidence intervals.
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Section 4.2 summarizes study characteristics (e.g., technologies, geographic distribution). No risk of bias assessment is included.
	20b	Present results of all	Section 4.4 and Table 2 present meta-analysis results:

Section and Topic	Item #	Checklist item	Location where item is reported
		statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	$\beta = 0.445$ (TechC-TTF), $\beta = 0.263$ (TCTTF), $\beta = 0.271$ (TTF-BI), with 95% CIs and I ² values (86.15%–94.87%) indicating high heterogeneity.
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Section 5.1 discusses heterogeneity, attributing it to technology type and demographics, but no statistical analysis (e.g., subgroup analysis) is presented.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	No sensitivity analyses are reported.
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	No assessment of reporting biases is presented for the syntheses.
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	No formal certainty assessment is provided. Section 5.3 notes limitations (small sample, no quality assessment) affecting confidence in results
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Section 5.1 interprets results in context, comparing findings to prior studies (e.g., Wang et al., 2020; Tao et al., 2023) and emphasizing TTF's role in healthcare.
	23b	Discuss any limitations of the evidence included in the review.	Section 5.3 discusses limitations of included evidence: small sample (15 studies), predominance of Asian studies, and focus on quantitative data
	23c	Discuss any limitations of the review processes used.	Section 5.3 discusses review process limitations: no protocol registration, no quality assessment, and limited geographic diversity.
	23d	Discuss implications of the results for practice, policy, and future research.	Sections 5.1 and 5.2 discuss implications for theory (extending TTF with BI), practice (user-centered design, training), and future research (qualitative studies, broader demographics).
OTHER INFORMATION			
Registration and	24a	Provide registration information for the	Section 5.3 states the review was not registered.

Section and Topic	Item #	Checklist item	Location where item is reported
protocol		review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	No protocol is mentioned or accessible, consistent with the limitation in Section 5.3
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	No amendments are mentioned, as no protocol was prepared.
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	No financial or non-financial support is mentioned in the paper.
Competing interests	26	Declare any competing interests of review authors.	No competing interests are declared in the paper.
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Appendix A provides extracted data (study IDs, authors, β values, sample sizes). No analytic code or additional materials are mentioned as publicly available.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71

EXPLORING THE MATURITY OF SMART PUBLIC GOVERNANCE: EVIDENCE FROM TWO POLICY SECTORS IN SLOVENIA

PETRA VUJKOVIĆ,^{1,2} TINA JUKIĆ¹

¹ University of Ljubljana, Faculty of Public Administration, Ljubljana, Slovenia
petra.vujkovic@fu-uni-lj.si, tina.jukic@fu-uni-lj.si

² University of Rijeka, Faculty of Economic and Business, Rijeka, Croatia
petra.vujkovic@fu-uni-lj.si

This paper applies a case study approach to test a multi-attribute decision-support model for assessing smart public governance maturity in real-life settings. Developed on prior research, verified through synthetic cases, and validated through expert focus group discussions, the model enables a structured, criteria-based assessment of smart public governance maturity across four dimensions: ICT-enabled innovations for public sector governance, institutional changes, empowered citizens, and outcomes: public value creation. Two public administration organisations in charge of policy making in Slovenia, each operating in a distinct policy sector, were used as case studies for testing the proposed assessment tool. Data were collected through semi-structured interviews with policy makers, and responses were aligned with predefined attribute value scales. The findings confirm that the model can be applied in practice as a useful tool for public sector organisations seeking to assess their smart public governance maturity level and identify areas for further improvement.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.16](https://doi.org/10.18690/um.fov.4.2025.16)

ISBN
978-961-286-998-4

Keywords:
smart public governance,
maturity assessment,
public administration,
policy making,
multi-attribute decision-
support model



University of Maribor Press

1 Introduction

As a research domain, smart public governance (SPG) is only emerging. Nevertheless, it has received growing scientific and policy attention, particularly due to its potential to address complex socio-economic challenges and public policy issues. A crucial first step for any government striving to advance SPG is to assess its current maturity level. Such an assessment enables public sector institutions, including those responsible for policy making, to understand where they stand and to identify actions that need to be taken to make progress. However, in searching for a suitable assessment tool, we encountered a problem. While a few frameworks, toolboxes, roadmaps, and models for assessing SPG have been proposed by scholars (notably Ruijter et al., 2023; Lin, 2018; Šiugždinienė et al., 2017; Bolívar & Meijer, 2016; Scholl & Scholl, 2014), these efforts have left room for further research, as they remain theoretical and lack empirical validation that would make them applicable in real-life settings. Considering these challenges, a multi-attribute decision-support model (hereinafter referred to as "the model") was developed based on the Decision Expert (DEX) – a qualitative, hierarchical, rule-based method within the Multi-Criteria Decision Analysis (MCDA) family (Berčič et al., 2024; Vrbek & Jukić, 2024; Bohanec, 2022: 2017: 2012; Jukić et al., 2022; Dobnik et al., 2018). Assessing SPG maturity requires the consideration of qualitative criteria, making DEX a suitable method for qualitative multi-attribute modelling. This approach is particularly well-suited for solving less formalised problems and has already been successfully applied in various real-life decision-making contexts (Kljajić Borštnar & Pucihar, 2021). The developed model is briefly outlined in Section 2.1 of this paper.

In this context, the study sets out to test the model in a real-life setting, establishing a "social experiment" aimed at demonstrating its practical usefulness. The main research question this paper seeks to answer is: *Can a previously developed multi-attribute decision-support model be effectively applied to assess and compare SPG maturity levels in public administration organisations in charge of policy making (PAO-PM)?*

This paper is structured as follows. Section 2 outlines the methodology, including the development of the model and the case study design. Section 3 presents the results of the SPG maturity assessment conducted in two PAO-PM, structured

around four aggregated attributes (criteria). Finally, Section 4 discusses the key findings and outlines the study's limitations.

2 Methodology

2.1 Development of a multi-attribute decision-support model

The model for assessing SPG maturity was developed following the Design Science Research (DSR) approach (Dresch et al., 2015). The development process included four main steps:

1. We conducted a structured literature review, i.e., a content analysis of Web of Science and Scopus papers. The analysis focused on the concept of SPG, including related terms such as smart city governance, smart urban governance, smart local governance, and smart public administration. Despite being different concepts, these terms have been used by scholars in developing frameworks, toolboxes, roadmaps, or models for measuring SPG – at both local and national levels.
2. We developed the model using the DEX methodology.
3. We verified the model using 20 synthetic cases, in which input values for the basic attributes (criteria) were randomly generated using Random.org – a true random number generator (Haahr, 2004; 2006; Bigger et al., 2008).
4. We validated the model by organizing two focus group discussions with experts, who participated either in face-to-face or online workshops to discuss the model's attributes (criteria), hierarchical structure, and domain values (attribute scales).

Fig. 1 shows the developed model, implemented in DEXi and drawn in a DEXiTree, a companion tool to the DEXi software (Bohanec, 2025: 2024: 2007; Jereb et al., 2003). The model root attribute, "assessment of SPG maturity", serves as the main output, having no parents and representing the highest level of aggregation. It is divided into four aggregated attributes (also known as subtrees): "ICT-enabled innovation for public sector governance," "institutional changes," "empowered citizens," and "outcomes: public value creation." Each of these subtrees has at least one descendant $\mathcal{S}(x) \neq \emptyset$, and is therefore considered a partial output of the model. The first subtree is further subdivided into three aggregated attributes ("use of

technology," "use of data," and "privacy and security"), which are then decomposed into basic attributes with no descendants $\mathcal{S}(x) = \emptyset$ and serve as model inputs. Overall, the model consists of 20 basic and 8 aggregated attributes, following the DEX method recommendation that complex decision-making occurs at higher hierarchical levels, while the number of attributes (criteria) decreases from lower-level to higher-level (Bohanec, 2012).

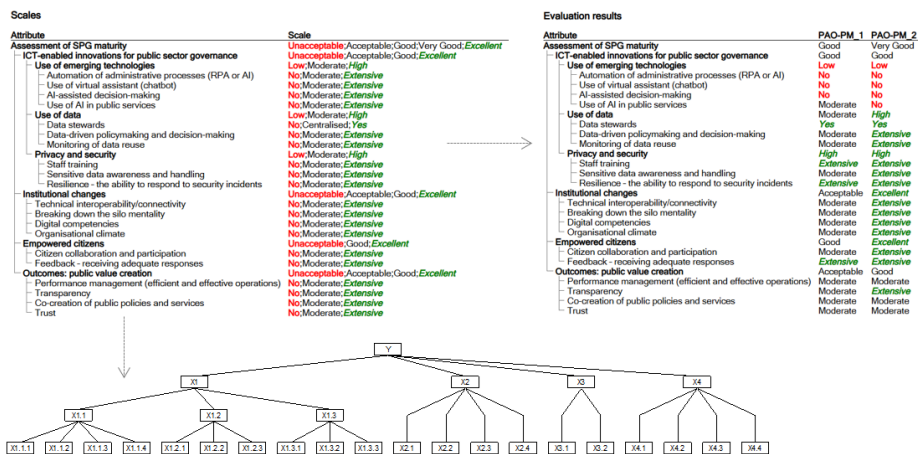


Figure 1: Developed model implemented in DEXi and drawn in a DEXiTree – align algorithm

2.2 Case study design

This study applied a qualitative case study approach to assess the maturity of SPG in two PAO-PM in Slovenia. These cases were selected based on their potential to provide rich informational content, as they operate in distinct policy sectors – one PAO-PM is responsible for science, technology and innovation policy, while the other oversees information society.

Interview instrument: the developed model consists of 20 basic attributes (criteria), each operationalised through a structured question with three predefined response options. These options represent ordinal attribute value scales (see Fig. 1) and were randomly ordered to minimise response bias. To encourage deeper reflection, each question also included an open explanation box, where policy makers justified their responses and provided practical examples. The responses from the policy makers

were used to determine the values of the model's basic attributes (criteria), which were then aggregated by the model to assess the overall SPG maturity.

Data collection and analysis: data were collected through in-person, semi-structured interviews with key policy makers from the two PAO-PM. One interview was conducted in October 2024 and the second in November 2024, each lasting approximately two hours. The interviews were conducted in Slovenian, audio-recorded, transcribed, and supplemented with field notes. The author coded the transcripts, aligning the responses with the attribute value scales defined in the model. Preliminary findings were shared with the policy makers in December 2024, followed by a short verification discussion in January 2025 to confirm whether the findings reflected the SPG maturity in their PAO-PM.

3 Results

Based on policy makers' responses, the model derived the final assessment of SPG maturity. As can be seen in Fig. 2, PAO-PM_1 was assessed as "good", while PAO-PM_2 was assessed as "very good" in terms of SPG maturity.

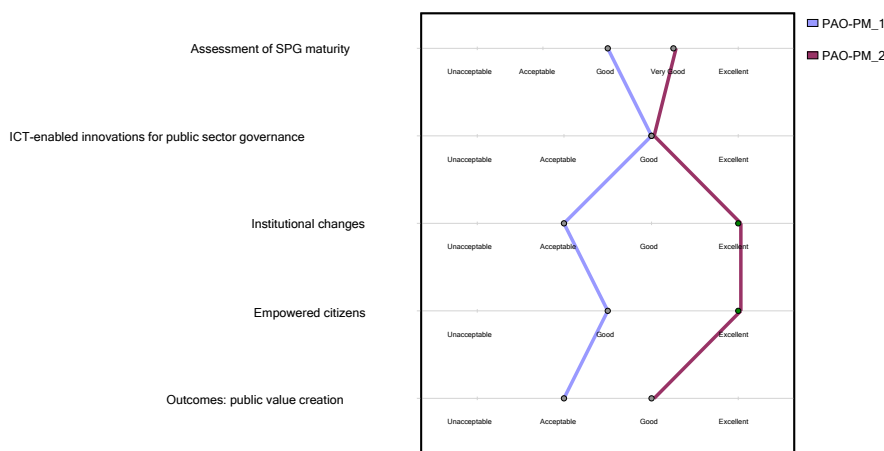


Figure 2: Assessment of SPG maturity, drawn in DEXiWin – linear chart

The remainder of this section presents the results for PAO-PM_1 and PAO-PM_2, structured around the four aggregated attributes (subtrees) of the developed model (see Section 2.1).

3.1 ICT – enabled innovations for public sector governance

The analysis showed that in both PAO-PM, most administrative (bureaucratic) processes, including human tasks, are not automated (see Fig. 3). Civil servants use various information systems (e.g., Krpan, Skrinja), which provide support for the performance of their tasks. However, these information systems, as functional tools, still require human intervention and do not operate in a fully automated manner. Furthermore, neither of the two PAO-PM provides a virtual assistant (chatbot) to citizens on their website (see Fig. 3). According to the policy makers we interviewed, the use of citizen-facing virtual assistants (chatbots) in the state administration is still in its embryonic phase. In addition, decisions taken rely solely on the experience and knowledge of the decision makers, without any support from AI algorithms (see Fig. 3). In PAO-PM_1, to some extent, integration of AI into a computer application improves public service delivery by optimising internal processes through back-end systems. On the other hand, PAO-PM_2 pointed out that AI currently has no impact, since functional tools are not supported by AI (see Fig. 3).

We found that both PAO-PM have in-house data stewards (see Fig. 3). While in PAO-PM_2 data is used in various stages of policy making, and decisions are always data-driven, PAO-PM_1 uses data in some stages of policy making due to the nature of the state administration, which functions as the administrative machinery of the state, reflecting its bureaucratic approaches to decision making (see Fig. 3). Similarly, PAO-PM_2 regularly (systematically) monitors its data re-use by users on the Open Data Portal – OPSI. Meanwhile, PAO-PM_1 monitors data re-use by users only when information on data re-use is requested (see Fig. 3).

Civil servants in both PAO-PM receive regular mandatory educational trainings on IS privacy and security (see Fig. 3), organised by the Administration Academy. In addition, civil servants may attend other non-mandatory workshops which are available according to their needs. In PAO-PM_2 all civil servants (regardless of their role) are aware of regulations governing the handling of sensitive (special categories of personal) data. Meanwhile, in PAO-PM_1 those civil servants who handle sensitive data know how data should be handled when transferred and stored (see Fig. 3). Resilience is a top priority for both PAO-PM, and civil servants are well informed about the security threats and practical guidelines on how to manage them (see Fig. 3).

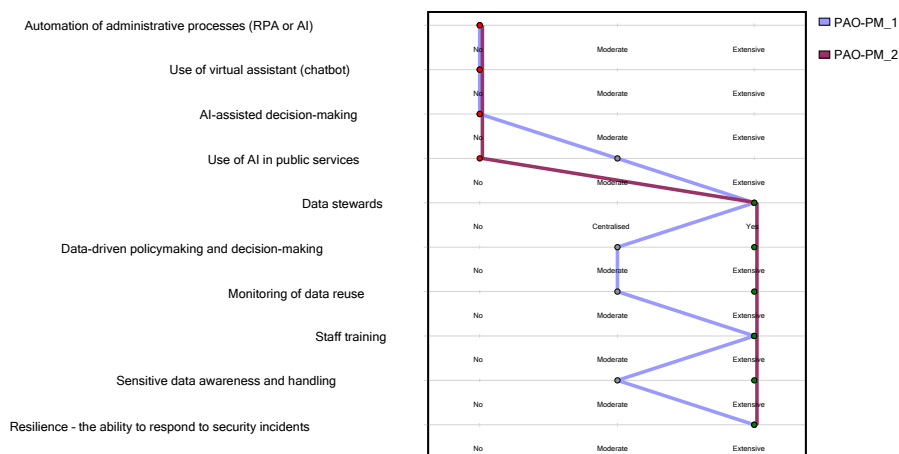


Figure 3: Assessment of the subtree: ICT – enabled innovations for public sector governance, drawn in DEXiWin – linear chart

3.2 Institutional changes

Information systems in both PAO-PM are interoperable, enable data semantics, and support the seamless distribution of information sources. However, PAO-PM_1 points out that the lack of interoperability among other information systems within the state administration hinders data and information distribution (see Fig. 4). According to the policy makers we interviewed in PAO-PM_2, efforts to break down silo mentality are ongoing and have resulted in a shift towards a different leadership style (switching to coaching). In addition, PAO-PM_1 noted that silo mentality is often perceived as beneficial by groups with a certain level of power and political control, leading to a lack of interest in breaking down silo mentality (see Fig. 4).

While PAO-PM_2 uses the 180-degree method to evaluate digital competence of civil servants, in PAO-PM_1, the digital competences are assessed solely based on employee task performance (see Fig. 4). In both PAO-PM, organisational climate is measured annually using a questionnaire. However, while in PAO-PM_1 the result does not lead to concrete changes, in PAO-PM_2 they are used to plan organisational interventions (see Fig. 4).

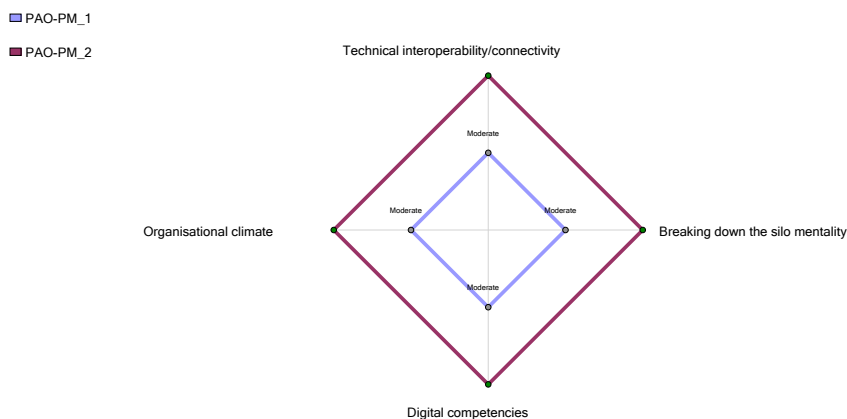


Figure 4: Assessment of the subtree: Institutional changes, drawn in DEXiWin – radar chart

3.3 Empowered citizens

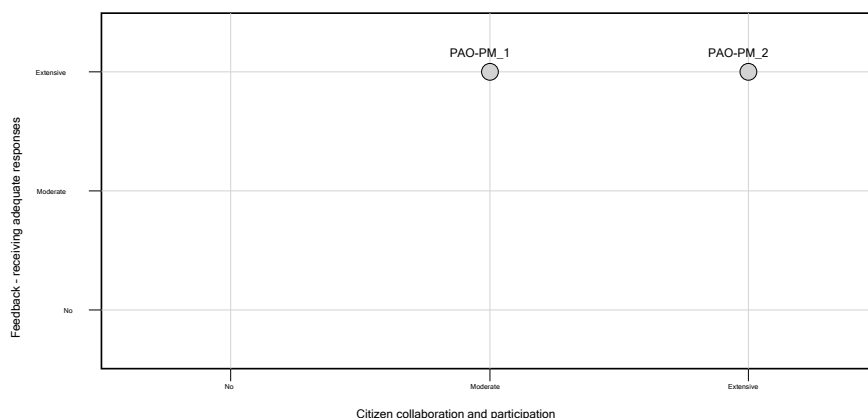


Figure 5: Assessment of the subtree: Empowered citizens, drawn in DEXiWin

The analysis showed that both PAO-PM engage with citizens through different channels, including digital platforms for citizen engagement (e.g., eZakonodaja, eDemokracija, Stop birokraciji, Predlagam vladi), as well as public meetings and debates held in physical or digital spaces (see Fig. 5). However, in PAO-PM_1, citizens do not participate in the drafting of the initial rulemaking due to a lack of active citizenship and potential conflict of interests (influencing regulatory processes and outcomes). In addition, in both PAO-PM citizens' voices matter – proposals or

opinions are considered (see Fig. 5), and acquired feedback is provided, even when initiatives cannot be considered.

3.4 Outcomes: public value creation

According to the policy makers we interviewed in PAO-PM, the goals are clearly defined and monitored at the operational level (see Fig. 6). However, setting strategic goals is more challenging due to their political nature. At the tactical level, goals can also be difficult to achieve, as public authorities may lose competencies at the end of their mandate. Both PAO-PM open proposed regulations and drafts policies to the public (see Fig. 6), yet in PAO-PM_1, the standard 30-day public comment period is not always observed. In practice, citizens contribute to the co-production of public policies and services (see Fig. 6) rather than full co-creation. Additionally, both PAO-PMs monitor perceived trustworthiness in government (see Fig. 6), but not citizen trust in their own organization.

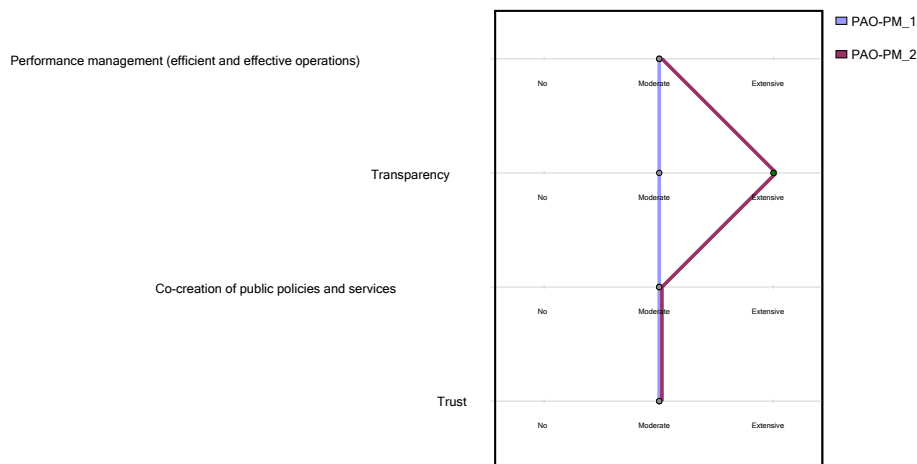


Figure 6: Assessment of the subtree: Outcomes: public value creation, drawn in DEXiWin – linear chart

4 Conclusion

The main objective of this study was to explore whether a previously developed model can be effectively applied to assess and compare SPG maturity levels in PAO-PM. The findings revealed several key weak points that PAO-PM need to address to advance their maturity in SPG.

The use of emerging technologies (such as RPA – robotic process automation, and AI-driven processing mechanisms) to automate human tasks, assist in the information provision to users, enhance decision-making processes, support evidence-based policymaking, and improve public service delivery remains largely absent in practice – primarily due to a complex legal and regulatory landscape. Attempts to implement RPA for rule-based routine process automation have failed in the past due to centralized ICT infrastructure, which prevents non-central authorities from intervening in the system. Adopting AI in state administration is a challenging process involving organisational, ethical, and legal barriers. In response, guidelines are being drafted to define the extent to which civil servants may use large language models (LLMs) and other AI-based text-generation tools, given the long-term risks associated with handling sensitive and confidential data. A shortage of data analysts further limits data-driven policymaking, as the volume and variety of collected data often remain underutilised.

Furthermore, the lack of architectural interoperability in state administration hinders information integration and cross-organisational sharing. This directly restrains the full digitization of procurement processes, preventing the sole use of electronic tendering systems. In addition, the fragmentation of digital public services across multiple platforms limits interoperability and user experience. The persistence of silo mentality is not rooted in bureaucratic structure per se but in the mindset of certain groups, which hinder collaboration across departments. Although digital upskilling initiatives exist, low digital self-awareness among civil servants reduces motivation in optional training programmes.

While formal mechanisms for citizen engagement are in place, opportunities for true co-creation of public policies and services with public and private actors remain limited. This is primarily due to strict regulatory frameworks designed to ensure transparency and prevent undue influence on decision-makers. In practice, due to

undergo legislative process, emergency situations, or international obligations and commitments, public authorities occasionally provide less than 30 days for public consultation on proposed regulations and draft policies. Citizen trust in public authorities is closely tied to general trust in the political system. The absence of a meritocratic system makes institutional trust more vulnerable to political fluctuations.

Finally, it is important to acknowledge the limitations that characterize this study, as they may represent valuable starting points for further research. First, the maturity of SPG has so far been assessed in only two PAO-PM. Therefore, additional applications are needed to confirm whether the developed model is capable of generating new knowledge and remains context-appropriate across a broader range of policy sectors. Second, as the model was validated by a group of experts in Slovenia, we acknowledge that it may reflect country-specific conditions when assessing SPG maturity. Consequently, this paper leaves the door open for further research, particularly to test the model in different countries, administrative traditions, and organisational settings to ensure its broader applicability.

References

- Berčić, T., Bohanec, M., & Ažman Momirski, L. (2024). Integrating Multi-Criteria Decision Models in Smart Urban Planning: A Case Study of Architectural and Urban Design Competitions. *Smart Cities*, 7, 786–805. <https://doi.org/10.3390/smartcities7020033>
- Biggar, P., Nash, N., Williams, K., & Gregg, D. (2008). An experimental study of sorting and branch prediction. *ACM Journal of Experimental Algorithmics*, 12 (1.8), 1-39. <https://doi.org/10.1145/1227161.1370599>
- Bohanec, M. (2025). DEXi: A Program for Multi-Attribute Decision Making (Version 5.05.) [Computer program]. DEXi: A Program for Multi-Attribute Decision Making
- Bohanec, M. (2024). *IJS Delovno Poročilo DEXiWin: DEX Decision Modeling Software User's Manual Version 1.2*. Jožef Štefan Institute
- Bohanec, M. (2022). DEX (Decision EXpert): A Qualitative Hierarchical Multi-criteria Method. In Kulkarni, A.J. (eds.), *Multiple Criteria Decision Making. Studies in Systems, Decision and Control* (pp. 39–78). Springer.
- Bohanec, M. (2017). Multi-Criteria Dex Models: An Overview and Analysis. In *Proceeding of the 14th International Symposium on Operational Research* (pp. 155-160).
- Bohanec, M. (2012). *Odločanje in modeli*. Ljubljana: DMFA založništvo.
- Bohanec, M. (2007). DEXiTree: a program for pretty drawing of trees. In *Proceedings of the Information Society IS 2007* (pp. 8–11). Ljubljana
- Bolivar, P.M.R. & Meijer, A. (2016). Smart Governance: Using a Literature Review and Empirical Analysis to Build a Research Model. *Social Science Computer Review*, 34(6), 1-20. <https://doi.org/10.1177/0894439315611088>
- Dresch, A., Lacerda, P.D., & Antunes, V.A.J. (2015). *Design Science Research*. London: Springer.

- Dobnik, D., Gruden, K., Zel, J., Bertheau, Y., Holst-Jensen, A., & Bohanec, M. (2018). Decision support for the comparative evaluation and selection of analytical methods: Detection of genetically modified organisms as an example. *Food Analytical Methods*, 11(8), 2105-2122. <https://doi.org/10.1007/s12161-018-1194-1>
- Haahr, M. (2024, December 17). *RANDOM.ORG: True Random Number Service*. Retrieved from <https://www.random.org>
- Haahr, M. (2006). Random Numbers. In N.J. Salkind (Eds.), *Encyclopedia of Measurement and Statistics*. Thousand Oaks: Sage.
- Jereb, E., Bohanec, M., & Rajkovič, V. (2003). *Računalniški program za večparametrsko odločanje: Uporabniški priročnik*. Kranj: Moderna organizacija.
- 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>
- Kljajić Borštnar, M. & Pucihar, A. (2021). Multi-Attribute Assessment of Digital Maturity of SMEs. *Electronics*, 10(8), 885. <https://doi.org/10.3390/electronics10080885>
- Lin, Y. (2018). A comparison of selected Western and Chinese smart governance: The application of ICT in governmental management, participation and collaboration. *Telecommunications Policy*, 42(10), 800-809. <https://doi.org/10.1016/j.telpol.2018.07.003>
- Ruijter, E., Twist, A., Haaker, T., Tartarin, T., Schuurman, N., Melenhorst, M., & Meijer, A. (2023). A Systematic Literature Review. *Smart Cities*, 6(0), 878-896. <https://doi.org/10.3390/smartcities6020042>
- Scholl, H.J. & Scholl, C.M. (2014). Smart Governance: A Roadmap for Research and Practice. In *iConference 2014 Proceeding* (pp. 163–176). <https://doi.org/10.9776/14060>
- Šiugždinienė, J., Gaule, E., & Rauleckas, R. (2017). In search of smart public governance: the case of Lithuania. *International Review of Administrative Sciences*, 85(0), 1-20. <https://doi.org/10.1177/0020852317707814>
- 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>

ADVANCED RAG-LLM PROTOTYPE AI ON PUBMED FOR CARDIAC HEALTH

LUUK P.A. SIMONS, PRADEEP K. MURUKANNAIAH,
BUDI S. HAN, MARK A. NEERINCX

Delft University of Technology, Faculty of EEMCS, Delft, the Netherlands
l.p.a.simons@tudelft.nl, P.K.Murukannaiah@tudelft.nl, b.s.han@student.tudelft.nl,
M.A.Neerincx@tudelft.nl

Healthy lifestyle behaviours are effective in preventing and treating cardiovascular disease. However, the growing body of scientific literature and the prevalence of conflicting studies make it challenging for healthcare practitioners and patients to stay informed. Large Language Models (LLMs), combined with Retrieval-Augmented Generation (RAG), enable automated claim verification and summarization. We enhanced RAG-LLM with extra modules and evaluated performance. Inclusion-Criteria-based filtering of PubMed papers improved verdict performance. Next, for health claims, PICO-based (Population, Intervention, Comparison, Outcome) paper mapping and summarization improves transparency of evidence used for verdict generation (like ‘Berries reduce blood pressure’). Still, the RAG-LLM models we tested have biases towards positivity (too many foods deemed heart healthy) and neutrality (no clear direction). We discuss mechanisms at play and challenges on the route forward.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.17](https://doi.org/10.18690/um.fov.4.2025.17)

ISBN

978-961-286-998-4

Keywords:

health self-management,
eHealth,
AI,
LLM,
RAG,
claim verification,
hypertension,
cardiovascular disease,
nutrition



University of Maribor Press

1 Introduction

Cardiovascular disease (CVD) is a leading cause of global mortality (Badimon, 2019; Gaidai, 2023). Hypertension, or high blood pressure, is a significant risk factor for CVD and the leading global cause of disease. Recent literature highlights the importance of evidence-based Health Self-Management (HSM) in improving cardiovascular health and reducing healthcare costs (Dineen-Griffin, 2019).

Literature on HSM and cardiac health is increasing rapidly (Qama, 2022). Information overload hinders timely access to insights useful in HSM support. As a special challenge in nutritional science, there is a general perception that studies are often contradictory (Nagler, 2014; Armitage, 2019). Moreover, fabricated science¹ and large food industry lobbies exist, resulting in fabricated guidelines². Studies have found that nutrition confusion is associated with nutrition backlash. For example, nutrition backlash decreased engagement in fruit and vegetable consumption (Lee, 2018). Given these challenges, we focus on nutrition (foods) and scientific evidence on how they help (or not) improve cardiac health and hypertension.

Our previous research on information needs and sources found that it is difficult for patients and practitioners to find actionable lifestyle advice which incorporates state-of-the-art scientific evidence (Simons, 2021, 2022a, 2023a): the Top 3 Dutch health institutes (for either hypertension or type 2 diabetes) provided watered down and inconsistent health advice, whereas as Google Scholar search heuristics analysis showed that the returned papers drown people in details and nonactionable research. In 2024, we surveyed 'expert users' (with LLM experience and who had just completed an intensive hypertension improvement challenge; Simons, 2022b, 2023b, 2024a, 2025) for their information needs and perceived added value of LLM's to help summarize health literature findings. In summary, they expressed concern about LLM's output and usefulness regarding (Simons, 2024b):

¹ Dr Neal Barnard (2018) eloquently explains how claims on cardiac health of eggs have (incorrectly) become more positive in the past decades, exactly because the previous decades had been so exhaustive on the negative cardiac health effects. In short, 'serious research' moved elsewhere, leaving a void filled by the egg industry to fabricate recent studies & reviews with designs to 'prove' healthiness.

² Even in the US Dietary Guidelines Advisory Committee, where objectivity should be key, 19 out of 20 members have clear industry affiliations and conflicting interests (Mialon, 2022).

- LLM information quality (and hallucinations, Sallam, 2023, Raina, 2024),
- dealing with conflicting health claims,
- explaining why updated advice is distinct from traditional/familiar advice,
- correct links & transparency regarding original studies used.

Still, advances in Artificial Intelligence (AI) like explicit claim verification and Retrieval-Augmented Generation (RAG) may mitigate these risks. So, we developed and tested an enhanced prototype, to help answer the **Research Question**:

To which extent can enhanced RAG-LLM models improve evidence inclusion and verdict transparency in mining nutrition science for cardiac health and hypertension claims?

2 Related work & Prototype design

The **claim verification** task is studied under the umbrella of automated fact-checking (Guo, 2022). The task involves automatically verifying the authenticity of claims based on the retrieval of evidence. A conventional framework of claim verification consist of three modules: the retrieval of relevant documents given a claim, selecting evidence from documents, and predicting a label (true, false, or not sure) based on the top-k evidence (Wadden, 2020; Pradeep, 2021; Soleimani, 2020). Research gap analyses (Gao, 2023; Wu, 2024; Liu, 2024) led us to focus on improving quality of evidence included and of transparency of verdict generation.

Liu et al. (2024) used the traditional three-step approach of claim verification but focused on an **Retrieval Augmented Generation (RAG)** module to specifically focus on RCT studies as evidence base for a given COVID claim. Instead of retrieving evidence from a prepared database, this augmented retrieval module presented a real life scientific use case. Below, we introduce Inclusion-Criteria-based filtering to enhance RAG by improving relevance of the inputs used. Retrieval-augmented methodologies harness the capabilities of multiple information retrieval techniques such as document vectorization, semantic similarity-based retrievers, and similarity ranking mechanisms.

We formulate concise claims using the ***Population, Intervention, Comparison, and Outcome (PICO) framework*** (Richardson, 1995). This framework is commonly used to formulate good clinical research questions, which can be utilized to formulate clinical claims by adapting the elements to suit the nature of the claim being made (Huang, 2006). For example Liu et al. (2024) used the PICO framework to construct a Covid Verification dataset of 15 PICO-encoded drug claims.

We introduce PICO use to translate ***nutrition*** science into health claims. PICO-encoded health claims can enhance document retrieval by guiding the search toward semantically relevant papers. Next, PICO-based paper summarization towards claim verdicts improves transparency. As an example PICO-based health claim:

- Population: Adults with high cholesterol levels
- Intervention: Consumption of flax seeds
- Comparison: Standard diet without flax seeds
- Outcome: Reduces LDL cholesterol levels
- Claim: *Consumption of flax seeds reduces LDL cholesterol in adults with high cholesterol.*

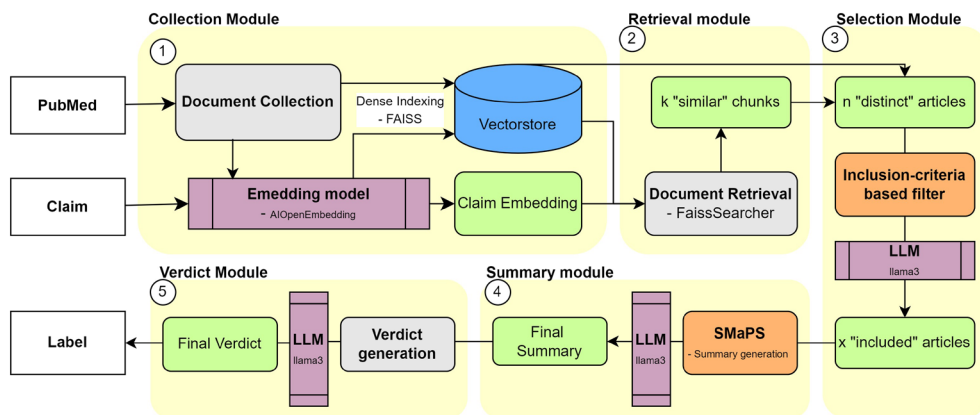


Figure 1: Framework with enhanced selection and summary modules

Figure 1 shows the overall ‘Advanced RAG-LLM’ framework we developed, including our extensions in orange. The numbers 1 to 5 highlight the key steps:

1. Document collection: based on semantic similarity, for each claim the most relevant articles are selected through the PubMed API. Each article (full text) is chunked into pieces of 1000 tokens to enable processing. FAISS³ (Facebook AI Similarity Search) is used for indexing, similarity search and clustering of dense vectors, to store these vectors.
2. Retrieval: In our case, a health claim serves as input query and FAISS searches the vectorstore for the documents that are semantically similar to the query vector. However, it is crucial to understand that semantic similarity does not necessarily equate to relevance or quality, see step 3.
3. Selection: An Inclusion-Criteria based filter was added to increase relevance of the selected papers. The PICO elements were used: Population = human adults (e.g. not animals). Intervention = Dietary Intervention (e.g. not a prospective study, or a medication intervention). Comparison = Control group or -condition stated. Outcome = blood pressure or cardiovascular health (e.g. not bone density etc).
4. Summary: Using SMaPS (see Figure 2), PICO- and summary texts per article are summarized into a final summary for a given claim.
5. Verdict: Based on the final summary for a claim, a verdict is created, a score from 1 (‘strongly refuted’) to 5 (‘strongly supported’).

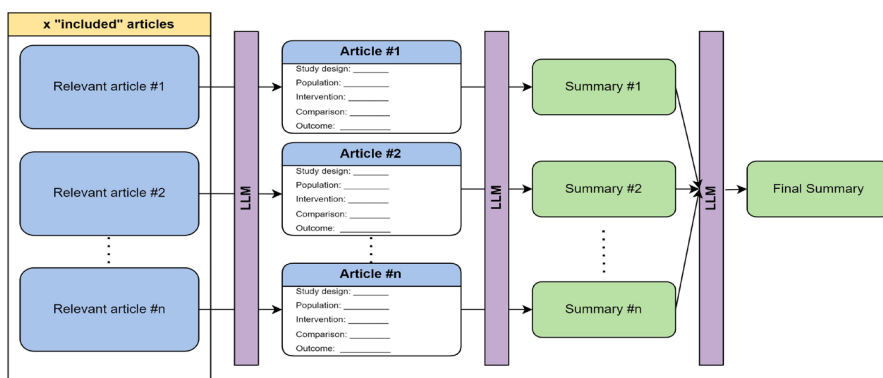


Figure 2: PICO-based Sequential Mapping (SMaPS)

³ <https://faiss.ai/>

Figure 2 shows how the top k articles for a claim (e.g. ‘Berries reduce blood pressure’) are synthesized using LLM’s towards PICO results and summaries, to create a final summary that can be used for verdict generation for that claim.

3 Method: Evaluation

Together with a cardiac health and nutrition expert, an initial set of 50 food and cardiac health claims was created (e.g. ‘Legumes lower blood pressure in human adults’). Half of them focused on blood pressure, half of them on cardiovascular health. Next, from the 200,000 PubMed papers specifically on nutrition and cardiac health, 10,000 papers were selected, based on highest semantic similarity with the 50 claims. Given the fact that we were processing full texts of the papers, 10,000 papers was the maximum which was feasible with the computing resources we had available to develop and test our prototype.

Next, we evaluated the effectiveness of the ‘Advanced RAG-LLM’ prototype. The *three main subquestions for evaluation*:

1. Selection Module: How accurate is the Inclusion-Criteria-based Selection Module?
2. Summary Module: How accurate and useful are PICO- and summary-synthesis?
3. Verdict Module: How accurate are the verdicts of the ‘Advanced RAG-LLM’ prototype model, in comparison to expert opinion?

For subquestion 1, on accuracy of the Inclusion-Criteria-based Selection Module, we conducted a manual check for 100 articles which were predicted by the model as ‘not’ fitting and 100 articles as ‘yes’ fitting the Inclusion-Criteria.

For subquestion 2, on accuracy of PICO synthesis and usefulness of summaries, we conducted a first, low level ‘expert user’ evaluation of PICO- and summary texts for specific scientific articles: which were read by these experts as part of their evaluation. Given the time constraints in this prototype engineering study, we followed the approach of early-stage, iterative small scale user testing. Generally, in this approach the first 80% of design flaws appear with the first 5 user tests (Faulkner, 2003). For the evaluation, we used 5 healthy lifestyle experts who were

familiar with both interpreting scientific studies and with coachee information needs. Each of them evaluated 2 articles; 5 articles were evaluated, and each article was evaluated by 2 experts. Overall, 10 article evaluations were done and we observed a high degree of consistency across expert evaluations.

For subquestion 3, on accuracy of the food health verdicts of the ‘Advanced RAG-LLM’ prototype model, for all 50 claims, the model verdict was compared to an expert verdict from 1 (‘strongly refuted’) to 5 (‘strongly supported’). The assignment for the expert was to generate verdicts following NutritionFacts⁴ state-of-the-art. We compared three different model versions, see Figure 3: BaseLLM, Standard RAG-LLM, and Advanced RAG-LLM to highlight module differences. To evaluate model accuracy for all three models we used confusion matrices (‘model predicted’ versus ‘expert’ scores) and Cohen’s weighted Kappa scores.

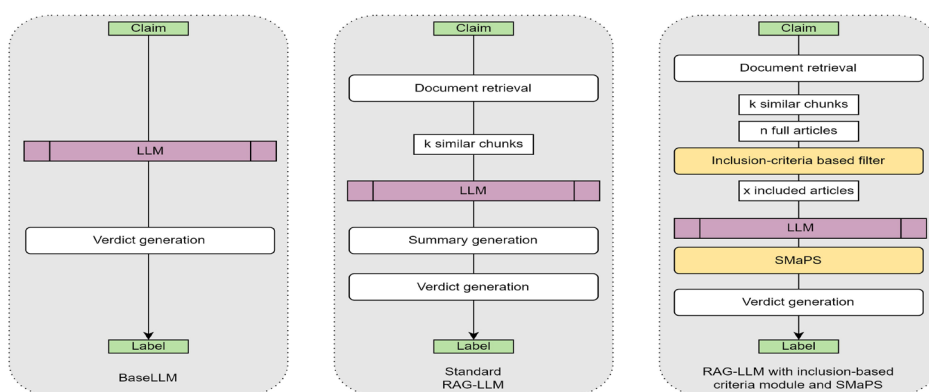


Figure 3: Comparing models: BaseLLM, RAG-LLM and Advanced RAG-LLM

4 Results

In response to *subquestion 1, on accuracy of the Inclusion-Criteria-based Selection Module*, Figure 4 shows that the number of False Negatives (FN) are modest in comparison to True Negatives: 3 out of 100 model predictions. Compared to the actual $n=83$ papers for inclusion the 3 FN papers are 3,6%. False Positives (FP) however form quite a group: 20 (17.1%) of the actual $n=117$ papers for

⁴ <https://nutritionfacts.org/>

exclusion are FP. Further analysis shows this is mostly due to incorrect flagging as intervention studies (of for example systemic reviews). Additionally, Population (animals) and Outcome (not really blood pressure) contributed to FP.

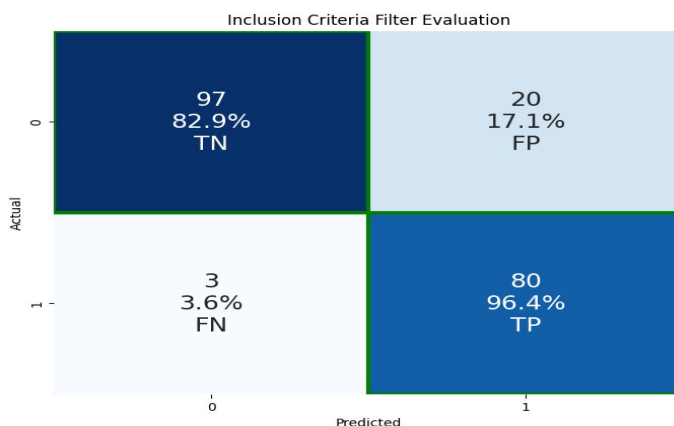


Figure 4: Confusion matrix for inclusion criteria: 0 = “excluded“ and 1 = “included“

In response to *subquestion 2, on accuracy of PICO synthesis and usefulness of summaries*, we found concerning *low performance*, see Figure 5. Whereas Intervention (4.2) and Comparison (4.1) score okay, accuracy scores for Population (2.0) and Outcome (1.9) descriptions were low. For Population accuracy in comparison to the original papers, some of the typical expert complaints were:

- “Missing info on sample size, selection method, cholesterol levels of study population.”
- “Sample size incorrect, for intervention and for control. Selection method info absent.”

For Outcome accuracy in comparison to the original papers, some of the typical expert complaints were:

- “Outcome incorrect; omission of the decrease in systolic blood pressure.”
- “Primary outcome measures from paper were missing.” → (10x out of 10)

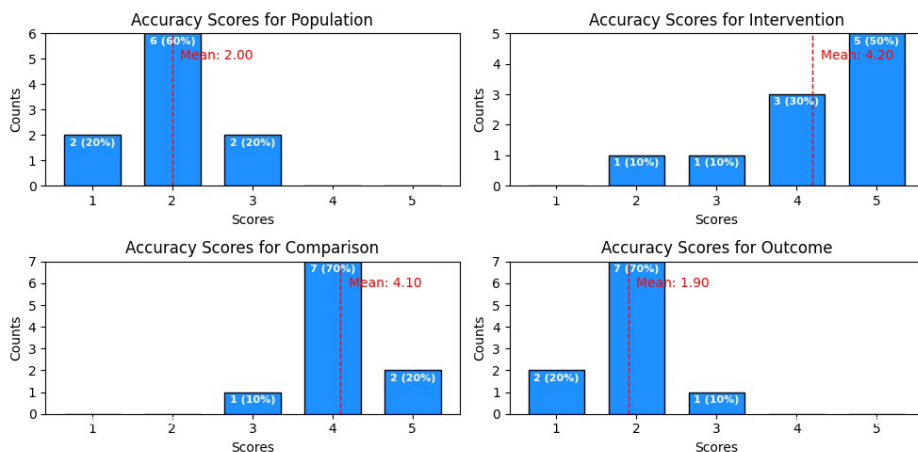


Figure 5: Expert Accuracy scores on PICO elements (n=10 paper evaluations)

Moreover, when *evaluating usefulness of a paper's summary text*, the average expert score is 2.1 (on range of 1 to 5). Illustrative expert remarks:

- “Blood pressure is left out of the results. Which is strange because it’s one of the main outcomes.” [multiple times]
- “Claim cannot be made on this study. Study focusses on kiwi’s only.”

Regarding *subquestion 3, on accuracy of the 50 food health verdicts of the three LLM models we compared*, we find that the Advanced RAG-LLM prototype generates improvements, but still LLM flaws persist, see Figures 6 and 7.

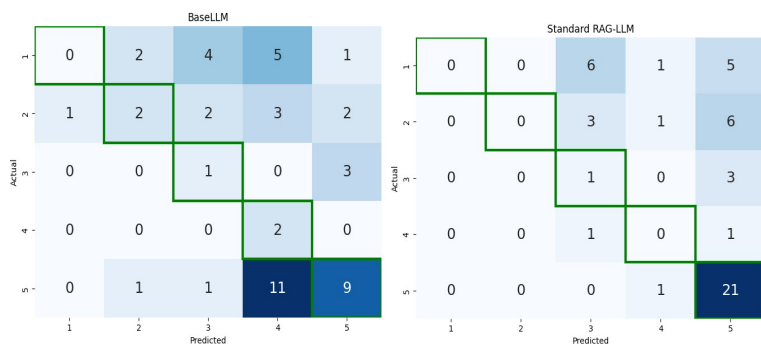


Figure 6: Confusion matrices BaseLLM & Standard RAG-LLM (n=50 claims predict/actual)

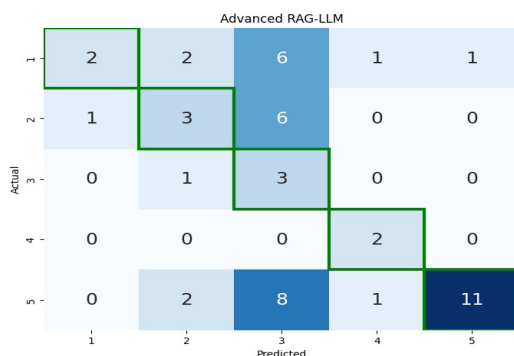


Figure 7: Confusion matrix Advanced RAG-LLM (n=50 claims: predict/actual)

The *BaseLLM model has a ‘positivity bias’*: most claims (foods) are predicted to be neutral (3) or healthy (4 or 5). The *Standard RAG-LLM has even more ‘positivity bias’*: Most claims (foods) are predicted to be clearly healthy (36 claims score 5; and 3 claims score 4). The rest (11 claims) are predicted as neutral: score 3. Interestingly, *the Standard RAG-LLM model labelled no foods at all as unhealthy for blood pressure or cardiac health* (not even sugar, alcohol, red meat, eggs, full fat dairy, cheese, salty foods, etc). As if there were not a single food detrimental to blood pressure or cardiovascular health in PubMed science...

Table 1: Model-Expert agreement: Cohen’s Weighted Kappa

Model	Cohen’s Weighted Kappa
BaseLLM	0.31
Standard RAG-LLM	0.27
Inclusion-based RAG-LLM	0.48

The *Advanced RAG-LLM prototype generates more balance* (positive and negative scores) and its verdicts are closer to expert opinion, see also its Kappa score of 0.48 in Table 1, meaning ‘weak’ agreement. Which is better than the ‘minimal’ agreements of 0.31 and 0.27 of the BaseLLM and Standard RAG-LLM models. Still, there is a bias towards neutral (3) scores and the Kappa scores indicate that there is significant room for improvement, which we discuss below.

5 Discussion & Conclusion

This prototyping study has several *limitations*. Most importantly, due to limits in computing resources, only 10,000 paper full texts were used (5% of the 200,000 cardiac health nutrition papers present in PubMed: the 5% most semantically similar to the 50 claims used). We expect that including more papers will improve results, see also our discussion of verdict flaws below. Besides, the three evaluations conducted each have their limitations. First, the evaluation of Inclusion-Criteria based filter (subquestion 1) was a simple face value check. This could have been made more rigorous by adding more evaluators. Still, for humans these are simple evaluations (e.g. this is not a nutrition intervention on adults, but a meta study, an in vitro study or a mouse model study), so we expect low error rates here.

The second evaluation, expert-based assessment of paper PICO- and summary texts, was relatively small-scale (subquestion 2): 10 paper evaluations were done, by a total of 5 domain experts. Still, from a design perspective, the fact that for example in 10 out of 10 paper evaluations the Outcome text was found lacking to some extent in relevant information, is a sign to first improve this sequential summarization module before proceeding to larger-scale evaluations.

The third evaluation, for subquestion 3, on accuracy of the 50 food health verdicts of the three LLM models we compared, used the inputs of only one expert in translating the evidence-base from NutritionFacts⁵ to verdicts. In a next phase, when the Advanced RAG-LLM is more robust, we can add more experts.

Still, the various user and expert evaluations do have face value: we can recognize the LLM model flaws that appear. Moreover, when digging deeper, we learn *lessons* why these specific flaws are the weaker points of model performance. For example, the *‘positivity bias’* we found, especially in the BaseLLM and Standard RAG-LLM models, is interesting. We see two possible reasons. Firstly, standard LLM’s (we used Meta’s Llama3)⁶ appear to be programmed to please the user (as various chess and other anecdotes in the history of LLM’s have illustrated). But the second reason may be more substantial: in hindsight, all our health claims were formulated as positive

⁵ <https://nutritionfacts.org/>

⁶ We used Llama3 model from <https://ollama.com/library/llama3> with a temperature setting of 0 to encourage the model to make prediction is purely based on the verdict with little creativity or variation.

statements (e.g. ‘Berries reduce blood pressure’). In follow up research it would be interesting to see if different top k articles are selected when all claims are formulated as negative statements: ‘... does not reduce blood pressure’).

Next, for some of the claims, not enough relevant intervention papers were included, which was exacerbated by including the wrong papers for a claim. This is illustrated by one of our explorative analyses: of the 50 claims, there were ***four claims (8%) with large differences (> 2 points) between expert- and Advanced RAG-LLM verdicts***. For these four claims (two claims on bananas and two on full fat dairy: claims for both blood pressure and cardiac health), we analysed the papers selected as evidence and a clear pattern appeared. After Inclusion-Criteria based filtering, only 3 to 5 papers were left as base material for these claims. Moreover, these were ***mostly ‘false positive’ papers***: included even though they were not specifically studying bananas or full-fat dairy. ***Thus, the LLM-based ‘final summary’ per claim was ‘forced into’ hallucination and overgeneralisation***. Though the prompting for study-referencing created transparent summaries (e.g. “The studies did not specifically investigate the effects of banana consumption on blood pressure.” And “foods like fruits [...] support overall health”) the evidence basis for a grounded verdict on the specific claim was absent (e.g. the n=5 studies included for the full fat and blood pressure claim were either focussing on low-fat dairy or dairy in general, but not on full-fat dairy).

Finally, ***one of the most pressing flaws we found, based on evaluating the PICO- and summary texts for papers, is the relatively poor representation of key elements/nuances from the scientific papers, as judged by the domain experts***. The outputs are lacking in nuance or context, rendering paper summaries less useful (score 2.1, from 1 to 5). Since these poor representations are used as inputs for creating final summaries and verdicts per claim, these verdicts become ‘averages of averages’. But in science, variance matters. For example, it makes a difference if we compare 0%-fat milk health effects to beans (0%-milk is comparatively ‘unhealthy’) or to twinkies (0%-milk is comparatively ‘healthy’). Or if we make equal-weight or equal-calorie food swaps. So averaging averages is often not helpful, e.g. “If you want to know how fast a cheetah can run, taking its average speed of the day generates a very wrong number.” This effect was aggravated by the fact that numbers (for study/control populations or outcomes) were often wrongly summarized or left

out. In short, the Llama 3 LLM and Facebook FAISS models we used appear to lack nuance for finding and summarizing the most relevant scientific facts.

Conclusion

The Advanced RAG-LLM prototype we tested does show improvements over base LLM and standard RAG-LLM performance. Firstly, we showed benefits from Inclusion-Criteria-based filtering. Secondly, we improved transparency with PICO-frame-based summarizing for verdict generation. Still, this could not disguise the fact that LLM's (= 'probable word generators') lack reasoning abilities and still disregard or misrepresent relevant facts from scientific papers. In short: The LLM-based models we tested are unable to distinguish relevant, solid science and findings from 'fabricated science' or low-relevance facts, thus they need extra reasoning tools.

Improved filtering and reasoning modules are needed to raise performance in:

- Paper selection
- Extracting key information (from: findings, control condition, population characteristics, study design, and expected causality of interventions)
- Context-based interpretation of numbers (e.g. 'Is this effect size large?')
- Highlighting findings which are most useful for patients and practitioners
- Recognizing and discarding 'fabricated science'

Acknowledgment

This research was (partly) funded by the <https://www.hybrid-intelligence-centre.nl/> a 10-year programme funded the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research, grant number 024.004.022 and by EU H2020 ICT48 project "Humane AI Net" under contract \$ \# \\$ 952026.

References

- Armitage, H. (2019). Any way you slice it, there's a lot to say about nutrition studies. Scope blog Stanford, accessed 28-2-2025: <https://scopeblog.stanford.edu/2019/01/28/any-way-you-slice-it-nutrition-studies-are-controversial/>.
- Badimon, L., Chagas, P., and Chiva-Blanch, G. (2019). Diet and Cardiovascular Disease: Effects of Foods and Nutrients in Classical and Emerging Cardiovascular Risk Factors. *Current Medicinal Chemistry*, 26(19):3639–3651.
- Barnard R. J. (2018). How the Egg Industry Skews Science. YouTube, accessed 26-2-2024: <https://www.youtube.com/watch?v=FyG8wr0gWIA>.

- Dineen-Griffin, S., Garcia-Cardenas, V., Williams, K., and Benrimoj, S. I. (2019). Helping patients help themselves: A systematic review of self-management support strategies in primary health care practice. *PLoS ONE*, 14(8), e0220116.
- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, 35, 379-383.
- Gaidai, O., Cao, Y., and Loginov, S. (2023). Global Cardiovascular Diseases Death Rate Prediction. *Current Problems in Cardiology*, 48(5):101622.
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*, 2.
- Guo, Z., Schlichtkrull, M., and Vlachos, A. (2022). A Survey on Automated Fact-Checking. *arXiv:2108.11896* [cs].
- Huang, X., Lin, J., and Demner-Fushman, D. (2006). Evaluation of PICO as a knowledge representation for clinical questions. *AMIA Annual Symposium proceedings. AMIA Symposium*, 2006, 359–363.
- Liu, H., Soroush, A., Nestor, J. G., Park, E., Idnay, B., Fang, Y., Pan, J., Liao, S., Bernard, M., Peng, Y., and Weng, C. (2024). Retrieval augmented scientific claim verification. *JAMLA Open*, 7(1), o0ae021.
- Lee, C.-j., Nagler, R. H., and Wang, N. (2018). Sourcespecific Exposure to Contradictory Nutrition Information: Documenting Prevalence and Effects on Adverse Cognitive and Behavioral Outcomes. *Health communication*, 33(4):453–461.
- Mialon, M., Serodio, P., Crosbie, E., Teicholz, N., Naik, A., & Carriedo, A. (2022). Conflicts of interest for members of the US 2020 Dietary Guidelines Advisory Committee. *Public Health Nutrition*, 1-28.
- Nagler, R. H. (2014). Adverse outcomes associated with media exposure to contradictory nutrition messages. *Journal of health communication*, 19(1):24–40.
- Pradeep, R., Ma, X., Nogueira, R., and Lin, J. (2021). Scientific Claim Verification with VerT5erini. In Holderness, E., Jimeno Yepes, A., Lavelli, A., Minard, A.-L., Pustejovsky, J., and Rinaldi, F., editors, *Proceedings of the 12th International Workshop on Health Text Mining and Information Analysis*, pages 94–103, online. Association for Computational Linguistics.
- Qama, E., Rubinelli, S., and Diviani, N. (2022). Factors influencing the integration of self-management in daily life routines in chronic conditions: a scoping review of qualitative evidence. *BMJ Open*, 12(12):e066647.
- Raina, A., Mishra, P., & Kumar, D. (2024). AI as a Medical Ally: Evaluating ChatGPT's Usage and Impact in Indian Healthcare. *arXiv preprint arXiv:2401.15605*.
- Richardson, W. S., Wilson, M. C., Nishikawa, J., and Hayward, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. *ACP Journal Club*, 123(3):A12. Publisher: American College of Physicians.
- Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. In *Healthcare MDPI*, Vol. 11, No. 6, p. 887.
- Simons, LPA, Neerincx MA, Jonker CM (2021). Health Literature Hybrid AI for Health Improvement; A Design Analysis for Diabetes & Hypertension, pp. 184-197, *34th Bled eConference*. June 27-30, Bled, Slovenia, Proceedings retrieval from www.bledconference.org. ISBN-13: 978-961-286-385-9, DOI: <https://doi.org/10.18690/978-961-286-385-9>.
- Simons, LPA, Neerincx MA, Jonker CM (2022a). Is Google Making us Smart? Health Self-Management for High Performance Employees & Organisations, *International Journal of Networking and Virtual Organisations*, Vol 27, No 3, pp.200-216. DOI: 10.1504/IJNVO.2022.10053605
- Simons, LPA, Gerritsen, B, Wielaard, B, Neerincx MA (2022b). 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, (2023a). Health 2050: Faster Cure via Bioinformatics & Quantified Self; A Design Analysis, *International Journal of Networking and Virtual Organisations*, Vol 28, No 1, pp.36-52. DOI:<https://doi.org/10.1504/IJNVO.2023.130957>.
- Simons, LPA, Gerritsen, B, Wielaard, B, Neerincx MA (2023b). Hypertension Self-Management Success in 2 weeks; 3 Pilot Studies, pp.19-34, *36th Bled eConference*. June 25-28, Bled, Slovenia, Proceedings retrieval www.bledconference.org. ISBN-13: 978-961-286-751-5, DOI: 10.18690/um.feri.6.2023
- Simons, LPA, Gerritsen, B, Wielaard, B, Neerincx MA (2024a). 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, Murukannaiah, PK, Neerincx, MA (2024b). Designing and Evaluating an LLM-based Health AI Research Assistant for Hypertension Self-Management; Using Health Claims Metadata Criteria, pp.283-298, *37th Bled eConference*. June 9-12, Bled, Slovenia, Proceedings. ISBN-13: 978-961-286-871-0, DOI: 10.18690/um.fov.4.2024.
- 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>
- Soleimani, A., Monz, C., and Worring, M. (2020). BERT for Evidence Retrieval and Claim Verification. In Jose, J. M., Yilmaz, E., Magalhaes, J., Castells, P., Ferro, N., Silva, M. J., and Martins, F., editors, *Advances in Information Retrieval*, pages 359–366, Cham. Springer International Publishing.
- Wadden, D., Lin, S., Lo, K., Wang, L. L., van Zuylen, M., Cohan, A., and Hajishirzi, H. (2020). Fact or Fiction: Verifying Scientific Claims. In Webber, B., Cohn, T., He, Y., and Liu, Y., editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7534–7550, Online. Association for Computational Linguistics.
- Wu, S., Xiong, Y., Cui, Y., Wu, H., Chen, C., Yuan, Y., ... & Xue, C. J. (2024). Retrieval-augmented generation for natural language processing: A survey. *arXiv preprint arXiv:2407.13193*.

SHAPING PARTICIPATION: HOW DIGITAL PLATFORM DESIGN INFLUENCES USER PARTICIPATION

GABRIEL MÄRZ

University of Turku, Turku School of Economics, Finland
gabriel.g.marz@utu.fi

This paper provides an integrative perspective on the factors influencing platform participation, a key driver of platform success. Based on a structured literature review of 99 scholarly articles from leading journals, this study synthesizes insights from the strategic, technical, and economic perspectives on digital platform design and governance from a participant viewpoint. It develops a participant-centric model that outlines the major determinants of platform adoption and homing decisions, considering both monetary and non-monetary costs, opportunity costs, and perceived benefits. The model also captures the dynamics of network effects and their implications for platform value creation, emphasizing the importance of strategic resource orchestration and adaptation over time. By unifying diverse concepts and frameworks, the paper addresses the complexity of platform ecosystems and provides actionable insights for platform providers seeking to enhance participant engagement and long-term platform viability. It contributes to the academic discourse by offering a comprehensive reference for future research and practical guidance for effective platform design and management.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.18](https://doi.org/10.18690/um.fov.4.2025.18)

ISBN
978-961-286-998-4

Keywords:

digital platforms,
platform design,
network effects,
platform scaling,
literature review



University of Maribor Press

1 Introduction

Digital platforms have garnered sustained academic interest due to their economic significance and their tendency to outcompete traditional pipeline businesses (Reuver et al., 2018). Various platform types have been studied with regard to technical and managerial design options (Evans & Schmalensee, 2005; McIntyre & Srinivasan, 2017), with particular focus on network effects, the primary differentiator from pipeline models, due to their distinctive dynamics (McIntyre & Srinivasan, 2017; Rietveld & Schilling, 2021). Initially seen as a self-reinforcing mechanism that could help early entrants dominate markets, network effects are now understood as strategic assets that must be actively managed to remain effective (Eisenmann et al., 2006; Knee, 2018; Srinivasan, 2023). Accordingly, the focus has shifted from pure scaling to resource orchestration for sustained advantage (Hagiu & Wright, 2024). Despite diverse literature streams offering guidance on how design choices shape network effects and platform success, an integrative view remains underdeveloped (Reuver et al., 2018; Rietveld & Schilling, 2021). Seminal works like McIntyre & Srinivasan (2017), Rietveld & Schilling (2021), and Hagiu & Wright (2024) attempt to unify related concepts, but typically adopt a provider-centric lens.

This paper contributes by offering a participant-centric, integrative model of platform participation, one of the key drivers of platform success, based on a structured literature review. The model maps how platform design choices and exogenous contextual factors influence adoption decisions and interact over time. In line with Keen & Williams (2013), who emphasize that user preferences shape market outcomes (Cennamo & Santalo, 2013; Panico & Cennamo, 2022), this paper reframes platform success from the participant's perspective. The resulting model integrates strategic, technical, and economic views (McIntyre & Srinivasan, 2017), identifying key determinants of homing decisions and their underlying drivers. Due to its neutral formulation, the model applies across platform types and contexts, offering a unified framework for scholars and actionable guidance for practitioners seeking to influence participant behavior through design interventions.

2 Theoretical Background

Digital platforms are virtual intermediaries that provide value to its participants by facilitating (intergroup and potentially intragroup) interactions and by providing valuable services (i.e. boundary resources) (Mantena & Saha, 2012). Overall, they aggregate internally and/or externally provided resources for its participants to

foster value creation and facilitate interaction (Hagiu, 2014; Leong et al., 2019; Srinivasan, 2023). Platforms typically exist in complex ecosystems, not solely having horizontal ties with players in the same or adjacent markets, but also having horizontal ties with firms operating in other layers (Chung et al., 2024; Cohen & Zhang, 2022). These ties or relationships may differ even within a platform, depending on the dependences with other entities on each level.

A key distinction from pipeline businesses is the presence of multiple participant groups whose interactions create intra-group and inter-group externalities, making platform governance and design features essential (Chen, Yi, et al., 2022; Constantinides et al., 2018; Srinivasan, 2023; van Alstyne et al., 2016). Platforms guide the interactions of their participants through the features available, as well as through rules and regulations that determine the openness of a platform along several dimensions (Broekhuizen et al., 2021; Kazan et al., 2018; Ondrus et al., 2015; Rietveld & Schilling, 2021; van Alstyne et al., 2016). While more lenient rules allow for flexible use of the platform and more complementary innovation, they also carry risks, such as platform forking or disintermediating the platform, increased complexity, or the loss of control over aspects such as quality (Broekhuizen et al., 2021; Kazan et al., 2018; Leong et al., 2019; McIntyre & Srinivasan, 2017; Rietveld & Schilling, 2021; Zhu & Iansiti, 2019).

Accordingly, digital platforms differ from pipeline businesses through their dynamic value proposition, shaped by inherent resources and the influence of participant characteristics and contributions on one another (Knee, 2018; Leong et al., 2019; Mantena & Saha, 2012; Srinivasan, 2023). This stems from network effects, positive or negative externalities among users, which depend on platform design, user preferences, homing decisions, and exogenous factors like market environment and competition (Carroni et al., 2024; Eisenmann et al., 2006; Farronato et al., 2024; Hagiu & Wright, 2024; Hinz et al., 2020; Mantena & Saha, 2012; Srinivasan, 2023; Weyl, 2010; K. Zhang & Sarvary, 2015).

A participant's adoption decision is thus driven by the platform's perceived value minus the cost of joining and using it. To optimize adoption, platforms often subsidize the more price-sensitive group or the group generating stronger externalities (Bar-Gill, 2019; Eisenmann et al., 2011; H. Li et al., 2021; H. Li & Zhu, 2021; McIntyre & Srinivasan, 2017; Parker et al., 2016).

Once a critical mass is achieved, network effects are considered to become self-reinforcing. However, reaching that threshold is time-sensitive and hinges on early adopters, creating a 'chicken and egg' dilemma that can limit a platform's market potential and lead to failure (Evans, 2009; Evans & Schmalensee, 2010; Hagiu & Rothman, 2016; McIntyre & Srinivasan, 2017; Ondrus et al., 2015; Parker et al., 2016). Consequently, diverse launch strategies have been developed to address this challenge (Caillaud & Jullien, 2003; Chen et al., 2019; Edelman, 2015; Evans, 2009; Knee, 2018; Schirmacher et al., 2017).

Beyond traditional network effects, data network effects have gained relevance with advances in AI and machine learning. These arise when platforms leverage growing data volumes to improve algorithms, refine services, and better understand user needs (Bergemann & Bonatti, 2024; Gregory et al., 2021; Varga et al., 2023). The ability to generate user insights enables personalized offerings, such as recommender systems or buyer analytics, and creates competitive barriers through accumulated data advantages (Bergemann & Bonatti, 2024; Gregory et al., 2021; Malgonde et al., 2022; Zhu & Iansiti, 2019).

3 Literature Review

3.1 Methods

To understand the drivers of platform success, a structured literature review was conducted following the guidelines of Webster & Watson (2002), Fink (2019), and the PRISMA-S extension for transparent and reproducible search strategies (Page et al., 2021; Rethlefsen et al., 2021). The systematic review was conducted in September 2024 using EBSCOhost (covering Business Source Ultimate, APA PsycInfo, and Communication & Mass Media Complete) and Web of Science. The review focused on identifying studies published in English-language journals listed in the FT50 or the AIS Senior Scholars' Basket of Journals, both of which represent high-quality, peer-reviewed outlets.

Search terms were developed by identifying common terminology in the field of digital platforms and refined iteratively to improve recall. The final search string was:

TI(("platform*"OR "two-sided*" OR "multi-sided*" OR "digital ecosystem*" OR "marketplace*") AND ("network" OR "winner" OR "competing platforms" OR "multi*hom*")) OR AB(("platform*"OR "two-sided*" OR "multi-sided*" OR "digital ecosystem*" OR "marketplace*") AND ("network" OR "winner" OR "competing platforms" OR "multi*hom*"))

No additional filters (e.g., date range) were used. The search strategy was peer-reviewed by a second researcher using the PRESS framework (McGowan et al., 2016). All references were exported to Citavi for citation management. Deduplication was performed using platforms' built-in tools and then manually by scanning title and author fields. This process resulted in 352 unique results.

Screening was conducted manually and independently by one researcher, starting with titles and abstracts, followed by full-text review where needed. Studies were included if they investigated digital platforms with a focus on factors influencing participation or design-related decisions. Overall, screening resulted in 69 relevant articles. Additionally, forward and backward citation tracking was applied to the included papers with a maximum of two iterations to identify further relevant literature, leading to a total of 99 relevant papers. A PRISMA 2020 flowchart (Figure 2, Appendix) summarizes the search and screening process (Page et al., 2021).

Relevant articles were thematically coded and synthesized using a manual, iterative process. Recurring themes and commonly used terms were grouped into a broader cost-benefit framework, offering a multidimensional view of factors contributing to platform success. Based on these findings, a conceptual model was systematically developed to illustrate the relationships influencing platform adoption and retention, which were identified as core drivers of platform success (van Alstyne et al., 2016; Varga et al., 2023). Using an inductive approach, key components and their interrelations were identified, clustered by similarity, and integrated into a value function simulating rational, value-based decision-making. The resulting model synthesizes diverse insights into a coherent framework, providing both conceptual clarity and a foundation for future empirical validation.

3.2 Findings

The literature yielded several insights regarding the competitive dynamics of platforms and the factors that contribute to a platform's success. Although there are numerous metrics for success, it was opted for a more nuanced approach. The dimensions and the factors influencing them that shape participants' homing and adaptation decisions were delineated. By identifying these dimensions, platform providers can gain insight into their participants' needs and adapt their strategies and design choices accordingly to achieve their goals and hence achieve success. The key points were clustered into different groups and will be outlined in the following sections, guided by Figure 1. It is important to acknowledge that the model's applicability is constrained by certain limitations. These include the assumption of some degree of rationality, despite the incorporation of bounded rationality and motivational factors (Hossain & Morgan, 2013; Sun & Gregor, 2023). This is in line with the insights from Simon (1966) and Petty et al. (1981). Hence, the model accounts for the absence of full information and the influence of motivational factors. However, it does not account for participants, fully relying on emotional or irrational decisions.

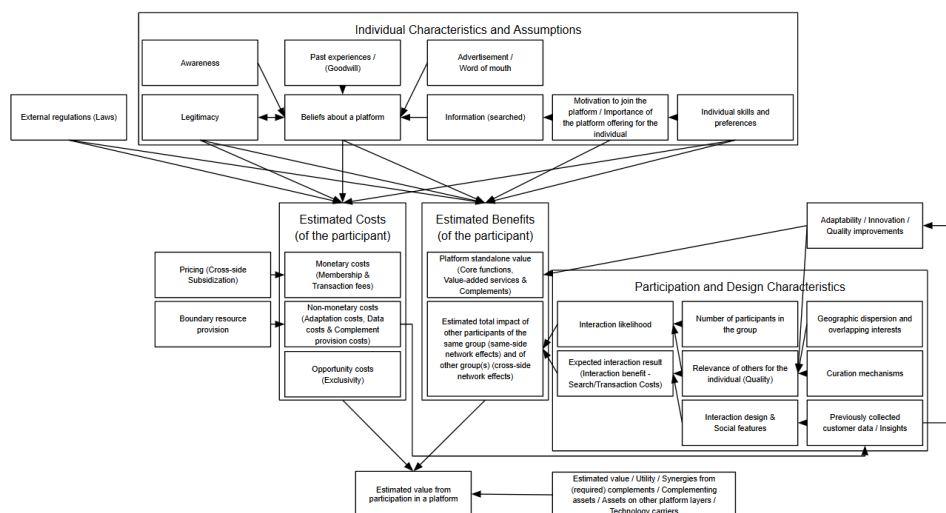


Figure 1: Influential Factors for a Participant's Platform Participation Decision

Source: Own

3.1.1 Fundamental Premise

The fundamental premise is organized as follows: Each potential participant of the different parties possesses distinctive characteristics, varying degrees of knowledge, capabilities, and preferences (Birge et al., 2021; Eisenmann et al., 2006; Farronato et al., 2024; Hagiü & Wright, 2024; Panico & Cennamo, 2022; Sun & Gregor, 2023; Veiga et al., 2017; K. Zhang & Sarvary, 2015). Therefore, the monetary and non-monetary costs and benefits associated with the platform may vary between the different parties and participants (Birge et al., 2021; Farronato et al., 2024). After estimating the value of the different options with the limited information available to them, rational potential participants will typically choose to participate in the platform that offers the highest expected total value for them, provided that the estimated overall value is positive (Edelman, 2015; Hagiü & Spulber, 2013; Hossain & Morgan, 2013; Keen & Williams, 2013; Mantena & Saha, 2012; McIntyre & Srinivasan, 2017). As participants are not limited to participation in a single platform, they subsequently evaluate the other options and estimate the additional value that would be gained from joining them (Bakos & Katsamakas, 2008; Barua & Mukherjee, 2021; Cennamo & Santalo, 2013; Hagiü & Wright, 2024; Tian et al., 2022). It is important to consider that the valuations are influenced by the previous choices made (Bakos & Halaburda, 2020). For instance, previous learning costs may be reflected to some extent in the valuations of other platforms (H. Li & Zhu, 2021; Polidoro & Yang, 2024). Conversely, solely new forms of interactions with previously engaged participants or various kinds of interaction with new participants translate into value (Bakos & Halaburda, 2020). However, the same kinds of interactions with the same participants that also multi-home do not provide any additional value (Bakos & Halaburda, 2020; Barua & Mukherjee, 2021). Therefore, in the case of overlapping participant groups, the value of each subsequent choice may be diminished to a certain extent in comparison to the initial valuation. This is contingent upon the presence of non-monetary costs that outweigh the benefits associated with the choice in question. Furthermore, positive interactions may help build goodwill towards a platform, increasing the likelihood of future participation (Calmon et al., 2021; Yang et al., 2022). In general, participants may utilize multiple platforms concurrently if they continue to derive value or utility from them (Bakos & Halaburda, 2020; Bakos & Katsamakas, 2008; Barua & Mukherjee, 2021). However, it is essential to consider that these decisions are subject to revision at specific intervals, and that preferences and, consequently, valuations may evolve

over time, especially as the platform and its ecosystem and industry mature (Broekhuizen et al., 2021; McIntyre & Srinivasan, 2017; Panico & Cennamo, 2022; Varga et al., 2023). This underscores the vital importance of platforms to reinvent themselves or to facilitate complementary innovation in order to maintain relevance (Cusumano & Gawer, 2002; Panico & Cennamo, 2022; van Alstyne et al., 2016; Varga et al., 2023; Zhu & Iansiti, 2019).

So, given this context, what drives the adaptation decision in detail? As indicated before, the decision process depends on the individual (Keen & Williams, 2013). First of all, if a participant is aware of the platform, past experiences, advertisements, and word-of-mouth shape a participant's beliefs about, and potentially the goodwill toward, a platform (Calmon et al., 2021; Edelman, 2015; Hagiú & Spulber, 2013; Keen & Williams, 2013; Koh & Fichman, 2014; Qiu & Rao, 2024; Sun & Gregor, 2023). In general, advertising and word-of-mouth increase a participant's awareness of a platform, which is influenced by the relevance of the platform's services to the participant and the participant's motivation to use the platform (Bar-Gill, 2019; Petty et al., 1981; van Alstyne et al., 2016; Varga et al., 2023; Yang et al., 2022; K. Zhang & Sarvary, 2015). Depending on a participant's motivation, the beliefs about a platform may be challenged by additional information gathered about a platform and its environment. This includes examining data about platform features, associated costs, legislation affecting the platform (Hagiú & Rothman, 2016; Ng & Pan, 2024), alternatives (Keen & Williams, 2013), the coherence of services offered (Taeuscher & Rothe, 2021), facilitated interactions, participating parties, beliefs about the platform's future development (Brouthers et al., 2016; Edelman, 2015; Zhu & Iansiti, 2012), risks involved with the platform usage (Gong et al., 2022; Hagiú & Rothman, 2016; Ng & Pan, 2024), etc. On this basis, the legitimacy or focality of a platform is assessed (Cennamo & Santalo, 2013; Ng & Pan, 2024; Sun & Gregor, 2023), as well as the potential costs and benefits associated with it, which ultimately influence a platform's valuation (Gong et al., 2022; Gregory et al., 2021; Hagiú & Rothman, 2016; Halaburda & Yehezkel, 2019; Khanagha et al., 2022).

3.1.2 Costs

A platform's valuation is heavily influenced by the costs of participation, which can be grouped into monetary, non-monetary, and opportunity costs (Carroni et al., 2024; Rietveld & Schilling, 2021; Srinivasan, 2023). These costs represent any barrier

to joining or staying on a platform, and while perceived subjectively, they generally correlate with actual costs (H. Li & Zhu, 2021).

Monetary costs include membership and transaction fees, which may be fixed or dynamic and vary across and within user groups (Birge et al., 2021; Du et al., 2014; Mantena & Saha, 2012; Rietveld & Schilling, 2021; Tavalaei et al., 2024; C. Zhang et al., 2022). They can regulate participation and influence platform quality (Anderson Jr. et al., 2023; Bakos & Katsamakas, 2008; Bernstein et al., 2021; Huang et al., 2022; Kazan et al., 2018; Tavalaei et al., 2024). Under the assumption of single-homing, cross-side subsidization, depending on the strength of network effects and the price sensitivity and demand elasticity of the parties, was often mentioned as a strategy to stimulate participation on one side and thus attract the other side to the platform (Barua & Mukherjee, 2021; H. Li et al., 2021; Srinivasan, 2023; Tavalaei et al., 2024; Weyl, 2010). Under the assumption of multi-homing, however, it is doubtful whether this strategy remains a viable option (Bakos & Halaburda, 2020; Cennamo & Santalo, 2013).

Non-monetary costs include switching and adaptation costs, such as learning or information costs associated with joining a platform, which act as barriers to entry but may also lock in users (Barua & Mukherjee, 2021; Cennamo & Santalo, 2013; Eisenmann et al., 2011; Gong et al., 2022; Hagiu & Rothman, 2016; Hagiu & Wright, 2024; Leong et al., 2019; Srinivasan, 2023). Data-related costs ("pricing by privacy") are a recurring form of non-monetary costs, as platforms may use personal data for service customization or monetization (Bergemann & Bonatti, 2024; Gregory et al., 2021; Varga et al., 2023). Changes in platform design (versioning) or functionality can increase adaptation costs, making backward compatibility a critical design consideration (Edelman, 2015; Panico & Cennamo, 2022; Xu et al., 2010).

Complement provision costs similarly affect adoption and homing choices. These include connectivity (ease of offering existing complements via interfaces) and the ability to create new complementary offerings using platform resources (Anderson Jr. et al., 2023; Hagiu & Wright, 2024; Polidoro & Yang, 2024; Rietveld & Schilling, 2021; Srinivasan, 2023; Tan et al., 2020). High connectivity thereby supports multi-homing, potentially reducing platform differentiation (Barua & Mukherjee, 2021; Kazan et al., 2018; Landsman & Stremersch, 2011; H. Li & Zhu, 2021; Mantena & Saha, 2012; Tian et al., 2022). While multi-homing is not inherently negative,

especially when a single platform cannot fully support complementors, platforms often aim to incentivize single-homing to strengthen network effects (Bar-Gill, 2019; Bernstein et al., 2021; K. Zhang & Sarvary, 2015). Thus, strategic investments in boundary resources must balance reducing participation frictions with maintaining platform distinctiveness (Anderson Jr. et al., 2023; Kazan et al., 2018; Tan et al., 2020).

Enabling complementary value creation differs from basic complement provisioning, as investments in boundary resources can increase the variety of new complements, attracting more users and enabling scaling over time (Chen, Tong, et al., 2022; Chen, Yi, et al., 2022; Leong et al., 2019; Tan et al., 2020). While such one-time investments can help establish a strong user base (Srinivasan, 2023), their impact may diminish if complementors multi-home using resources provided by the platform (Anderson Jr. et al., 2023). Granting greater autonomy to complementors can reduce provision costs but risks platform forking and added complexity (Broekhuizen et al., 2021; Hagiwara & Wright, 2015; Karhu et al., 2018). Therefore, platform providers must carefully calibrate investments in boundary resources and design, considering factors such as market position, competition, participant characteristics, and the behavioral effects of platform changes (Bar-Gill, 2019; Barua & Mukherjee, 2021; Eisenmann et al., 2006; Karhu et al., 2018; Sun & Gregor, 2023; Tan et al., 2020).

Opportunity costs, the benefits forgone or penalties avoided by choosing one option over another, also influence platform decisions (Carroni et al., 2024; Tian et al., 2022; Zhu & Iansiti, 2019).

In instances where platforms are predicated on specific complements, hardware, or technology carriers (e.g., mobile phones and cars are indispensable for drivers to operate on Uber), it is imperative to assess the overall value, utility, and potential synergies when making a decision about participating (Chung et al., 2024; Cohen & Zhang, 2022; Leong et al., 2019). Considering the relative costs associated with each potential course of action, the decision regarding platform participation may be made in conjunction with, or as a precursor to, the acquisition of the complementary product. The valuation of the complementary product may have implications for the homing behavior of participants (Chung et al., 2024; Cohen & Zhang, 2022; Leong et al., 2019).

3.1.3 Benefits

Platform benefits can be broadly divided into standalone value perceived independently by users and the value derived from others' participation (Hagiu & Rothman, 2016; Rietveld & Schilling, 2021). Standalone value includes platform functions, content, and resources that initially attract participants and provide a sustainable competitive edge, as they are not reliant on complementors (Anderson et al., 2014; Mantena & Saha, 2012; Zhu & Iansiti, 2012). Matchmaking algorithms, although reliant on user participation, are also part of standalone value due to their design-based independence (Meyer et al., 2024; Zhu & Iansiti, 2019). Participant data can further tailor and improve offerings, driving data-network effects, a self-reinforcing loop that boosts long-term platform value (Bergemann & Bonatti, 2024; Gregory et al., 2021; Varga et al., 2023).

Standalone value can be divided into core functions, such as matchmaking and transaction facilitation, and value-added services (Anderson et al., 2014; Anderson Jr. et al., 2023; Bhargava & Choudhary, 2004; Hagiu, 2014; Hagiu & Spulber, 2013). Core functions, shaped by platform architecture, are critical for initial adoption but may lead to adverse effects if over-optimized (Bakos & Katsamakas, 2008; Eisenmann et al., 2011; Meyer et al., 2024; Wu et al., 2018).

Value-added services, in contrast, are additional services (or boundary resources), including tools for customization or analytics options that enhance user and complementor experiences (Anderson Jr. et al., 2023; Du et al., 2014; Veiga et al., 2017; Wang et al., 2024), but their benefits can vary across participant groups and may impact overall welfare (Bergemann & Bonatti, 2024; Bhargava & Choudhary, 2004; Wang et al., 2024). These services may be offered freely or monetized, depending on strategic goals (Anderson Jr. et al., 2023; Bhargava & Choudhary, 2004; Du et al., 2014; Tan et al., 2020; Wang et al., 2024).

Platforms may also offer proprietary complements or content to drive adoption and achieve a critical mass (Hagiu & Spulber, 2013; Hagiu & Wright, 2015; Qiu & Rao, 2024; Tian et al., 2022, 2022; K. Zhang & Sarvary, 2015). Such offerings can also help to differentiate the platform and facilitate market envelopment, though they may disrupt existing complementors depending on the context (Bar-Gill, 2019; Hagiu & Wright, 2015, 2024; Haviv et al., 2020; Qiu & Rao, 2024; Raj, 2024; van

Alstyne et al., 2016). As platforms expand, service portfolios must be strategically managed, with attention to complementarity, user base overlap, and cost implications (Bar-Gill, 2019; Cusumano & Gawer, 2002; Eisenmann et al., 2011; Khanagha et al., 2022; Schrieck et al., 2024; Sun & Gregor, 2023).

Platform externalities significantly influence platform valuation (Gong et al., 2022; Varga et al., 2023). Depending on the availability of desirable participant traits, cross-side substitution, shaped by externalities, demand elasticity, and price sensitivity, can help regulate participation (Bakos & Halaburda, 2020; Bakos & Katsamakos, 2008; Cennamo & Santalo, 2013). Similarly, differential intragroup pricing may promote quality control or leverage spillover effects within participant groups (Bhargava et al., 2022; Haviv et al., 2020; Huang et al., 2022; Raj, 2024).

The desired characteristics of participants (also often referred to as quality), varying by platform type, individual interests, offering, and type of interaction, signal the likelihood of satisfactory interactions (Birge et al., 2021; Knee, 2018). This likelihood depends on the availability and relevance of others, interaction frequency, and associated costs (Anderson Jr. et al., 2023; Birge et al., 2021; Halaburda et al., 2018; Keen & Williams, 2013; Taeuscher & Rothe, 2021; Wu et al., 2018). Participant quality, often shaped by geography, shared interests, or other traits (Brouthers et al., 2016; Chen et al., 2019; Ghemawat, 2005; Varga et al., 2023; Veiga et al., 2017; K. Zhang & Sarvary, 2015; Zhu & Iansiti, 2019), is critical but not always observable upfront, making interactions risky (Hagiu & Wright, 2024). To mitigate this, platforms use curation, and promote participant signaling, filtering, or other risk mitigation mechanisms to improve interaction outcomes (Belleflamme & Peitz, 2018; Edelman, 2015; Hagiu & Rothman, 2016; Kozinets et al., 2021; H. Li & Zhu, 2021; Pu et al., 2022; Sun & Gregor, 2023; Tunc et al., 2021; Z. Zhou et al., 2024).

Matchmaking algorithms, including recommender systems, discovery services, and filtering tools, enhance interaction quality by simplifying and improving match accuracy, especially when data volume is high (Banerjee et al., 2016; Belleflamme & Peitz, 2018; Gregory et al., 2021; Hagiu & Wright, 2024; Meyer et al., 2024; B. Zhou & Zou, 2023; Zhu & Iansiti, 2019). In their absence, high search costs can reduce platform value (L. Li et al., 2020; Malgonde et al., 2022; Wu et al., 2018). The effectiveness of such systems depends on understanding users' value drivers (Hagiu & Wright, 2024; Zhu & Iansiti, 2019), with paid prioritization potentially distorting

platform welfare (Banerjee et al., 2016; Bergemann & Bonatti, 2024; Guo & Easley, 2016; Malgonde et al., 2022; B. Zhou & Zou, 2023). Platforms must balance transparency, trust, applicability, and usability when designing curation and recommendation mechanisms (Hagiu & Rothman, 2016; Halaburda et al., 2018; Kozinets et al., 2021; H. Li & Zhu, 2021; Pu et al., 2022; Tunc et al., 2021).

Lastly, it needs to be considered that the relevance of interactions depends on the platform usage intentions of a participant; and, therefore, the different types of interactions are weighed unevenly when considering their impact on a platform's valuation (Birge et al., 2021; Farronato et al., 2024; Meyer et al., 2024; Panico & Cennamo, 2022; Rietveld & Schilling, 2021; Tauscher & Rothe, 2021; Veiga et al., 2017; K. Zhang & Sarvary, 2015).

Intergroup interactions are typically most valuable, aligning with core use cases and revenue streams (Chen, Yi, et al., 2022). In these kinds of interactions, the match quality, depending on the level of satisfaction with a match, is therefore especially relevant (Birge et al., 2021; Hagiu & Rothman, 2016; Huang et al., 2022; Varga et al., 2023; Veiga et al., 2017). In contrast, intragroup interactions are more optional and may in some cases foster competition, risking participant fragmentation or market imbalance (Hagiu & Wright, 2024; Koh & Fichman, 2014; Malgonde et al., 2022; Meyer et al., 2024; Qiu & Rao, 2024; Tucker & Zhang, 2010). Thus, platforms should carefully manage both inter- and intragroup dynamics to optimize participation and value creation, which is dependent on the relative stance of the participants (Anderson Jr. et al., 2023; Chu & Manchanda, 2016; Hagiu & Wright, 2024; Kazan et al., 2018; Leong et al., 2019; Meyer et al., 2024).

Accordingly, a platform's valuation evolves with each new participant due to the cumulative effect of network externalities (Chen, Yi, et al., 2022; Knee, 2018; Srinivasan, 2023). Therefore, even in the absence of modifications to the platform's design, the value proposition of the platform is dynamic (Mantena & Saha, 2012; Rietveld & Schilling, 2021). It is crucial to ascertain whether the initially positive network effects persist in a self-reinforcing manner and to identify strategies for achieving and maintaining this dynamic, given that various factors, including participants' characteristics and interests, influence these effects (Knee, 2018; Leong et al., 2019; Panico & Cennamo, 2022; Rietveld & Schilling, 2021; Zhu & Iansiti, 2019). Also, different kinds of network effects, like same-side, different kinds of

cross-side, and data network effects may vary in their impact at different levels of participation and platform maturity (Anderson Jr. et al., 2023; Hinz et al., 2020; McIntyre & Srinivasan, 2017; Panico & Cennamo, 2022; Srinivasan, 2023; Varga et al., 2023; Zhu & Iansiti, 2019). Overall, the strategic and adaptive management of the diverse array of network effects is crucial for the success of a platform (Cennamo & Santalo, 2013; Rietveld & Schilling, 2021; Srinivasan, 2023).

In conclusion, the value of a platform can be assessed differently at any given time, contingent on the information available to the participants regarding the platform, their own characteristics, and their motivation to use it (Anderson Jr. et al., 2023; Keen & Williams, 2013; Z. Zhou et al., 2024; Zhu & Iansiti, 2019). Additionally, there are numerous cost factors or potential benefits that are not solely influenced by platform features and design; rather, they are also affected by the participation of other individuals (Gong et al., 2022). These factors may or may not be taken into account by a participant when making such a valuation estimate (Mantena & Saha, 2012; K. Zhang & Sarvary, 2015). Platform providers exert an influence on a platform's valuation, as they may modify their design decisions over time, which in turn affects potential participant costs and benefits (Anderson Jr. et al., 2023; Keen & Williams, 2013; Varga et al., 2023). Furthermore, this valuation may also undergo changes, as any other factors, such as participant clusters or individual interests or characteristics, like the perception of risk, undergo changes over time (Ng & Pan, 2024; Zhu & Iansiti, 2019).

A final consideration of a platform provider, not directly associated with a participant's valuation of a platform, is the order and manner of market entry and/or platform envelopment (Hagiu & Rothman, 2016; Haiyang Feng et al., 2020; Srinivasan, 2023). The manner and timing of a platform's market entry have implications for its positioning and long-term success. The decision of whether to be a first or second mover in a market is influenced by the dynamics of the market and the strength of network effects (Feng et al., 2020; Haiyang Feng et al., 2020; McIntyre & Srinivasan, 2017; Srinivasan, 2023; Zhu & Iansiti, 2012). Network effects can either create an entry barrier for second movers, locking participants in, or provide second movers with the opportunity to target specific segments (Feng et al., 2020; Hagiu & Rothman, 2016; Leong et al., 2019; McIntyre & Srinivasan, 2017). Consequently, there are various market entry strategies, which typically revolve

around creating a high level of relevance in a specific market niche or targeting a specific participant group (Ondrus et al., 2015).

4 Conclusion

Digital platforms are complex constructs that must constantly adapt to heterogeneous and evolving participant needs across and within user groups (Birge et al., 2021; Farronato et al., 2024; Meyer et al., 2024; Reuver et al., 2018; Rietveld & Schilling, 2021; Weyl, 2010). However, much of the existing literature either narrowly addresses technical or managerial design features in isolated contexts or offers generalized findings with limited consideration of interdependencies. In addition, the few studies that have attempted to identify the interdependencies of design options and their potential impact on platform success have mostly been written from the perspective of platform providers, disregarding influencing factors from the perspective of participants (Hagiu & Wright, 2024; McIntyre & Srinivasan, 2017; Reuver et al., 2018; Rietveld & Schilling, 2021). This paper addresses these gaps by modeling the platform participation process from a participant-centric perspective and linking managerial design actions to platform adoption dynamics. Through a structured literature review and theoretical integration, it identifies key design elements, contextual factors, and governance choices that shape user participation, an indicator for platform success (Keen & Williams, 2013; Panico & Cennamo, 2022).

The model captures a broad range of adoption drivers, including individual-level factors such as prior experiences, word-of-mouth, and user motivation. It also accounts for barriers to participation, including monetary, non-monetary, and opportunity costs. In contrast, platform benefits are categorized into core functionalities, value-added services, third-party contributions, and network effects, offering a nuanced understanding of perceived value. By incorporating bounded rationality, imperfect information, and temporally iterative decision-making, the model addresses limitations of static or overly rational economic frameworks such as the Technology Acceptance Model (Davis, 1989; Kamal et al., 2020). Moreover, it is specifically tailored to the platform adoption process and integrates a wide array of both soft and hard factors influencing decision-making. As a result, it enables the simulation of a more holistic and realistic platform participation decision.

The aggregation of findings from papers in different contexts furthermore enables generalization across empirical studies and reveals causal interdependencies that are often overlooked (Rietveld & Schilling, 2021). As participant valuations are shown to vary significantly, the model challenges winner-takes-all assumptions and highlights the possibility of multiple coexisting dominant platforms in certain markets (Cennamo & Santalo, 2013; Srinivasan, 2023). By tracing how participant perceptions evolve over time and influence platform engagement, this paper provides platform providers with actionable guidance on adaptive governance, value provision, and strategic positioning (Z. Zhou et al., 2024).

5 Limitations and Future Research

While this paper offers a holistic view of platform participation and design interdependencies, several limitations must be acknowledged, both in relation to the study's scope and its methodological approach.

First, the literature review was limited to English-language publications appearing in journals listed on the FT50 and the AIS Senior Scholars' List. Although this ensured scholarly quality, it potentially excluded relevant insights from other high-quality journals, practitioner publications, or non-English sources. Extending the journal and language scope could generate additional perspectives or challenge some of the conclusions drawn here.

Second, the thematic synthesis involved an interpretive coding process based on conceptual and theoretical patterns across studies. Despite efforts to ensure analytical rigor, the absence of multiple coders or an intercoder reliability assessment introduces a degree of subjectivity. Similarly, the paper did not include a formal quality assessment of the reviewed studies, which could influence the weight or reliability of the synthesized findings.

Third, the conceptual model developed in this paper is grounded in a literature-based synthesis rather than empirical validation. While the model identifies key drivers of platform participation and their interdependencies, its explanatory and predictive validity has not yet been tested in real-world settings. Future research could strengthen these insights by applying the model in specific platform contexts and evaluating its outcomes.

Fourth, although this paper generalizes findings across multiple studies, some of the relationships identified would benefit from empirical testing and contextual grounding. For example, further work is needed to quantify the qualitative findings outlined here, particularly regarding the impact of design options, the influence of individual preferences and motivations, and the trade-offs between different governance choices.

Fifth, while the paper explores various forms of network effects, there remains a lack of research on their interdependencies, their relative significance at different stages of platform development, and the contextual variables that shape them (McIntyre & Srinivasan, 2017; Panico & Cennamo, 2022; Srinivasan, 2023; Varga et al., 2023). The model assumes that such effects vary by stage and user group, but empirical clarification is still lacking.

Finally, although the paper discusses the potential of emerging technologies such as AI, recommender systems, and machine learning to mitigate negative externalities (e.g., search costs, complexity), these claims remain conceptual. Further empirical investigation is needed to determine how such technologies can influence user experience and participation, and whether they offset the increasing complexity associated with platform growth (Gregory et al., 2021; Malgonde et al., 2022).

References

- Anderson, E. G., Parker, G. G., & Tan, B. (2014). Platform Performance Investment in the Presence of Network Externalities. *Information Systems Research*, 25(1), 152–172. <https://doi.org/10.1287/isre.2013.0505>
- Anderson Jr., E. G., Parker, G. G., & Tan, B. (2023). Strategic Investments for Platform Launch and Ecosystem Growth: A Dynamic Analysis. *Journal of Management Information Systems*, 40(3), 807–839. <https://doi.org/10.1080/07421222.2023.2229125>
- Bakos, Y., & Halaburda, H. (2020). Platform Competition with Multihoming on Both Sides: Subsidize or Not? *Management Science*, 66(12), 5599–5607. <https://doi.org/10.1287/mnsc.2020.3636>
- Bakos, Y., & Katsamakas, E. (2008). Design and Ownership of Two-Sided Networks: Implications for Internet Platforms. *Journal of Management Information Systems*, 25(2), 171–202. <https://doi.org/10.2753/MIS0742-1222250208>
- Banerjee, S., Sanghavi, S., & Shakkottai, S. (2016). Online Collaborative Filtering on Graphs. *Operations Research*, 64(3), 756–769. <https://doi.org/10.1287/opre.2016.1508>
- Bar-Gill, S. (2019). Game of Platforms: Strategic Expansion into Rival (Online) Territory Sagit Bar-Gill. *Journal of the Association for Information Systems*, 20(10), 1475–1502. <https://doi.org/10.17705/1jais.00575>

- Barua, A., & Mukherjee, R. (2021). MULTI-HOMING REVISITED: LEVEL OF ADOPTION AND COMPETITIVE STRATEGIES. *MIS Quarterly*, 45(2), 897–924. <https://doi.org/10.25300/MISQ/2021/15416>
- Belleflamme, P., & Peitz, M. (2018). Inside the engine room of digital platforms: Reviews, ratings, and recommendations.
- Bergemann, D., & Bonatti, A. (2024). Data, Competition, and Digital Platforms. *American Economic Review*, 114(8), 2553–2595. <https://doi.org/10.1257/aer.20230478>
- Bernstein, F., DeCroix, G. A., & Keskin, N. B. (2021). Competition Between Two-Sided Platforms Under Demand and Supply Congestion Effects. *M&SOM: Manufacturing & Service Operations Management*, 23(5), 1043–1061. <https://doi.org/10.1287/msom.2020.0866>
- Bhargava, H. K., & Choudhary, V. (2004). Economics of an Information Intermediary with Aggregation Benefits. *Information Systems Research*, 15(1), 22–36. <https://doi.org/10.1287/isre.1040.0014>
- Bhargava, H. K., Wang, K., & Zhang, X. (2022). Fending Off Critics of Platform Power with Differential Revenue Sharing: Doing Well by Doing Good? *Management Science*, 68(11), 8249–8260. <https://doi.org/10.1287/mnsc.2022.4545>
- Birge, J., Candogan, O., Chen, H., & Saban, D. (2021). Optimal Commissions and Subscriptions in Networked Markets. *M&SOM: Manufacturing & Service Operations Management*, 23(3), 569–588. <https://doi.org/10.1287/msom.2019.0853>
- Broekhuizen, T., Emrich, O., Gijzenberg, M. J., Broekhuis, M., Donkers, B., & Sloot, L. M. (2021). Digital platform openness: Drivers, dimensions and outcomes. *Journal of Business Research*, 122, 902–914. <https://doi.org/10.1016/j.jbusres.2019.07.001>
- Brouthers, K. D., Geisser, K. D., & Rothlauf, F. (2016). Explaining the internationalization of ibusiness firms. *Journal of International Business Studies*, 47(5), 513–534. <https://doi.org/10.1057/jibs.2015.20>
- Caillaud, B., & Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. *RAND Journal of Economics*, 309–328.
- Calmon, A. P., Ciocan, F. D., & Romero, G. (2021). Revenue Management with Repeated Customer Interactions. *Management Science*, 67(4), 2944–2963. <https://doi.org/10.1287/mnsc.2020.3677>
- Carroni, E., Madio, L., & Shekhar, S. (2024). Superstar Exclusivity in Two-Sided Markets. *Management Science*, 70(2), 991–1011. <https://doi.org/10.1287/mnsc.2023.4720>
- Cennamo, C., & Santalo, J. (2013). Platform competition: Strategic trade-offs in platform markets. *Strategic Management Journal* (John Wiley & Sons, Inc.), 34(11), 1331–1350. <https://doi.org/10.1002/smj.2066>
- Chen, L., Shaheer, N., Yi, J., & Li, S. (2019). The international penetration of ibusiness firms: Network effects, liabilities of outsidership and country clout. *Journal of International Business Studies*, 50(2), 172–192. <https://doi.org/10.1057/s41267-018-0176-2>
- Chen, L., Tong, T. W., Tang, S., & Han, N. (2022). Governance and Design of Digital Platforms: A Review and Future Research Directions on a Meta-Organization. *Journal of Management*, 48(1), 147–184. <https://doi.org/10.1177/01492063211045023>
- Chen, L., Yi, J., Li, S., & Tong, T. W. (2022). Platform Governance Design in Platform Ecosystems: Implications for Complementors' Multihoming Decision. *Journal of Management*, 48(3), 630–656. <https://doi.org/10.1177/0149206320988337>
- Chu, J., & Manchanda, P. (2016). Quantifying Cross and Direct Network Effects in Online Consumer-to-Consumer Platforms. *Marketing Science*, 35(6), 870–893. <https://doi.org/10.1287/mksc.2016.0976>
- Chung, H. D., Zhou, Y. M., & Ethiraj, S. (2024). Platform Governance in the Presence of Within-Complementor Interdependencies: Evidence from the Rideshare Industry. *Management Science*, 70(2), 799–814. <https://doi.org/10.1287/mnsc.2023.4706>
- Cohen, M. C., & Zhang, R. (2022). Competition and coopetition for two-sided platforms. *Production & Operations Management*, 31(5), 1997–2014. <https://doi.org/10.1111/poms.13661>

- Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Introduction—platforms and infrastructures in the digital age, 29(2).
- Cusumano, M. A., & Gawer, A. (2002). The Elements of Platform Leadership. *MIT Sloan Management Review*, 43(3), 51–58.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=6553425&site=ehost-live&scope=site>
- Davis, F. D. (1989). Technology acceptance model: TAM. Al-Suqri, MN, Al-Aufi, as: *Information Seeking Behavior and Technology Adoption*, 205(219), 5.
- Du, A. Y., Das, S., Gopal, R. D., & Ramesh, R. (2014). Optimal Management of Digital Content on Tiered Infrastructure Platforms. *Information Systems Research*, 25(4), 730–746.
<https://doi.org/10.1287/isre.2014.0548>
- Edelman, B. (2015). How to Launch Your Digital Platform. *Harvard Business Review*, 93(4), 90–97.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=101711629&site=ehost-live&scope=site>
- Eisenmann, T., Parker, G., & van ALSTYNE, M. W. (2006). STRATEGIES FOR TWO-SIDED MARKETS. *Harvard Business Review*, 84(10), 92–101.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=22316862&site=ehost-live&scope=site>
- Eisenmann, T., Parker, G., & van Alstyne, M. (2011). Platform envelopment. *Strategic Management Journal* (John Wiley & Sons, Inc.), 32(12), 1270–1285. <https://doi.org/10.1002/smj.935>
- Evans, D. (2009). How catalysts ignite: the economics of platform-based start-ups. *Platforms, Markets and Innovation*, 416.
- Evans, D., & Schmalensee, R. (2005). The industrial organization of markets with two-sided platforms. National Bureau of Economic Research Cambridge, Mass., USA.
- Evans, D., & Schmalensee, R. (2010). Failure to launch: Critical mass in platform businesses. *Review of Network Economics*, 9(4).
- Farronato, C., Fong, J., & Fradkin, A. (2024). Dog Eat Dog: Balancing Network Effects and Differentiation in a Digital Platform Merger. *Management Science*, 70(1), 464–483.
<https://doi.org/10.1287/mnsc.2023.4675>
- Feng, H., Jiang, Z., Li, M., & Feng, N. (2020). First- or Second-Mover Advantage? The Case of IT-Enabled Platform Markets. *MIS Quarterly*, 44(3), 1107–1141.
<https://doi.org/10.25300/MISQ/2020/15273>
- Fink, A. (2019). *Conducting research literature reviews: From the internet to paper*. Sage publications.
- Ghemawat, P. (2005). Regional Strategies for Global Leadership. *Harvard Business Review*, 83(12), 98–108.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=18916536&site=ehost-live&scope=site>
- Gong, X., Cheung, C. M. K., Liu, S., Zhang, K. Z. K., & Lee, M. K. O. (2022). Battles of mobile payment networks: The impacts of network structures, technology complementarities and institutional mechanisms on consumer loyalty. *Information Systems Journal*, 32(4), 696–728.
<https://doi.org/10.1111/isj.12366>
- Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, S. H. (2021). THE ROLE OF ARTIFICIAL INTELLIGENCE AND DATA NETWORK EFFECTS FOR CREATING USER VALUE. *Academy of Management Review*, 46(3), 534–551.
<https://doi.org/10.5465/amr.2019.0178>
- Guo, H., & Easley, R. F. (2016). Network Neutrality Versus Paid Prioritization: Analyzing the Impact on Content Innovation. *Production & Operations Management*, 25(7), 1261–1273.
<https://doi.org/10.1111/poms.12560>
- Hagiu, A. (2014). Strategic Decisions for Multisided Platforms. *MIT Sloan Management Review*, 55(2), 71–80.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=94342384&site=ehost-live&scope=site>

- Hagiu, A., & Rothman, S. (2016). Network Effects Aren't Enough. *Harvard Business Review*, 94(4), 64–71.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=113934519&site=ehost-live&scope=site>
- Hagiu, A., & Spulber, D. (2013). First-Party Content and Coordination in Two-Sided Markets. *Management Science*, 59(4), 933–949. <https://doi.org/10.1287/mnsc.1120.1577>
- Hagiu, A., & Wright, J. (2015). Marketplace or Reseller? *Management Science*, 61(1), 184–203.
<https://doi.org/10.1287/mnsc.2014.2042>
- Hagiu, A., & Wright, J. (2024). WILL THAT MARKETPLACE SUCCEED? *Harvard Business Review*, 102(4), 94–103.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=177779174&site=ehost-live&scope=site>
- Haiyang Feng, Zhengrui Jiang, Minqiang Li, & Nan Feng (2020). FIRST- OR SECOND-MOVER ADVANTAGE? THE CASE OF IT-ENABLED PLATFORM MARKETS. *MIS Quarterly*, 44(3), 1107–1141. <https://doi.org/10.25300/misq/2020/15273>
- Halaburda, H., Jan Piskorski, M., & Yildirim, P. (2018). Competing by Restricting Choice: The Case of Matching Platforms. *Management Science*, 64(8), 3574–3594.
<https://doi.org/10.1287/mnsc.2017.2797>
- Halaburda, H., & Yehzekel, Y. (2019). Focality advantage in platform competition. *Journal of Economics & Management Strategy*, 28(1), 49–59.
- Haviv, A., Huang, Y., & Li, N. (2020). Intertemporal Demand Spillover Effects on Video Game Platforms. *Management Science*, 66(10), 4788–4807.
<https://doi.org/10.1287/mnsc.2019.3414>
- Hinz, O., Otter, T., & Skiera, B. (2020). Estimating Network Effects in Two-Sided Markets. *Journal of Management Information Systems*, 37(1), 12–38.
<https://doi.org/10.1080/07421222.2019.1705509>
- Hossain, T., & Morgan, J. (2013). When Do Markets Tip? A Cognitive Hierarchy Approach. *Marketing Science*, 32(3), 431–453. <https://doi.org/10.1287/mksc.1120.0770>
- Huang, P., Lyu, G., & Xu, Y. (2022). Quality Regulation on Two-Sided Platforms: Exclusion, Subsidization, and First-Party Applications. *Management Science*, 68(6), 4415–4434.
<https://doi.org/10.1287/mnsc.2021.4075>
- Kamal, S. A., Shafiq, M., & Kakria, P. (2020). Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). *Technology in Society*, 60, 101212.
- Karhu, K., Gustafsson, R., & Lyytinen, K. (2018). Exploiting and defending open digital platforms with boundary resources: Android's five platform forks. *Information Systems Research*, 29(2), 479–497.
- Kazan, E., Tan, C.-W., Lim, E. T., Sørensen, C., & Damsgaard, J. (2018). Disentangling Digital Platform Competition: The Case of UK Mobile Payment Platforms. *Journal of Management Information Systems*, 35(1), 180–219. <https://doi.org/10.1080/07421222.2018.1440772>
- Keen, P., & Williams, R. (2013). Value architectures for digital business: beyond the business model. *MIS Quarterly*, 37(2), 643–647.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=87371444&site=ehost-live&scope=site>
- Khanagha, S., Ansari, S., Paroutis, S., & Oviedo, L. (2022). Mutualism and the dynamics of new platform creation: A study of Cisco and fog computing. *Strategic Management Journal* (John Wiley & Sons, Inc.), 43(3), 476–506. <https://doi.org/10.1002/smj.3147>
- Knee, J. A. (2018). Why Some Platforms Are Better Than Others. *MIT Sloan Management Review*, 59(2), 18–20.
<https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=127196159&site=ehost-live&scope=site>
- Koh, T. K., & Fichman, M. (2014). Multihoming users' preferences for two-sided exchange networks. *MIS Quarterly*, 38(4), 977–996.

- Kozinets, R. V., Ferreira, D. A., & Chimenti, P. (2021). How Do Platforms Empower Consumers? Insights from the Affordances and Constraints of Reclame Aqui. *Journal of Consumer Research*, 48(3), 428–455. <https://doi.org/10.1093/jcr/ucab014>
- Landsman, V., & Stremersch, S. (2011). Multihoming in Two-Sided Markets: An Empirical Inquiry in the Video Game Console Industry. *Journal of Marketing*, 75(6), 39–54. <https://doi.org/10.1509/jm.09.0199>
- Leong, C., Pan, S. L., Leidner, D. E., & Huang, J.-S. (2019). Platform leadership: Managing boundaries for the network growth of digital platforms. *Journal of the Association for Information Systems*, 20(10), 1.
- Li, H., Shen, Q., & Bart, Y. (2021). Dynamic Resource Allocation on Multi-Category Two-Sided Platforms. *Management Science*, 67(2), 984–1003. <https://doi.org/10.1287/mnsc.2020.3586>
- Li, H., & Zhu, F. (2021). Information Transparency, Multihoming, and Platform Competition: A Natural Experiment in the Daily Deals Market. *Management Science*, 67(7), 4384–4407. <https://doi.org/10.1287/mnsc.2020.3718>
- Li, L., Chen, J., & Raghunathan, S. (2020). INFORMATIVE ROLE OF RECOMMENDER SYSTEMS IN ELECTRONIC MARKETPLACES: A BOON OR A BANE FOR COMPETING SELLERS. *MIS Quarterly*, 44(4), 1957–1985. <https://doi.org/10.25300/MISQ/2020/14614>
- Malgonde, O. S., Zhang, H., Padmanabhan, B., & Limayem, M. (2022). Managing Digital Platforms with Robust Multi-Sided Recommender Systems. *Journal of Management Information Systems*, 39(4), 938–968. <https://doi.org/10.1080/07421222.2022.2127440>
- Mantena, R., & Saha, R. L. (2012). Co-opetition Between Differentiated Platforms in Two-Sided Markets. *Journal of Management Information Systems*, 29(2), 109–140. <https://doi.org/10.2753/MIS0742-1222290205>
- McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS peer review of electronic search strategies: 2015 guideline statement. *Journal of Clinical Epidemiology*, 75, 40–46.
- McIntyre, D. P., & Srinivasan, A. (2017). Networks, platforms, and strategy: Emerging views and next steps. *Strategic Management Journal* (John Wiley & Sons, Inc.), 38(1), 141–160. <https://doi.org/10.1002/smj.2596>
- Meyer, T., Kerkhof, A., Cennamo, C., & Kretschmer, T. (2024). Competing for attention on digital platforms: The case of news outlets. *Strategic Management Journal* (John Wiley & Sons, Inc.), 45(9), 1731–1790. <https://doi.org/10.1002/smj.3600>
- Ng, E., & Pan, S. L. (2024). Competitive strategies for ensuring Fintech platform performance: Evidence from multiple case studies. *Information Systems Journal*, 34(3), 616–641. <https://doi.org/10.1111/isj.12406>
- Ondrus, J., Gannamaneni, A., & Lyytinen, K. (2015). The Impact of Openness on the Market Potential of Multi-Sided Platforms: A Case Study of Mobile Payment Platforms. *Journal of Information Technology*, 30(3), 260–275. <https://doi.org/10.1057/jit.2015.7>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Panico, C., & Cennamo, C. (2022). User preferences and strategic interactions in platform ecosystems. *Strategic Management Journal* (John Wiley & Sons, Inc.), 43(3), 507–529. <https://doi.org/10.1002/smj.3149>
- Parker, G. G., van ALSTYNE, M. W., & Choudary, S. P. (2016). Platform revolution: How networked markets are transforming the economy and how to make them work for you. WW Norton & Company.
- Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41(5), 847.

- Polidoro, F., & Yang, W. (2024). Porting learning from interdependencies back home: Performance implications of multihoming for complementors in platform ecosystems. *Strategic Management Journal* (John Wiley & Sons, Inc.), 45(9), 1791–1821. <https://doi.org/10.1002/smj.3601>
- Pu, J., Nian, T., Qiu, L., & Cheng, H. K. (2022). Platform Policies and Sellers' Competition in Agency Selling in the Presence of Online Quality Misrepresentation. *Journal of Management Information Systems*, 39(1), 159–186. <https://doi.org/10.1080/07421222.2021.2023410>
- Qiu, Y., & Rao, R. C. (2024). Can Merchants Benefit from Entry by (Amazon-Like) Platform if Multiagent Prices Signal Quality? *Marketing Science*, 43(4), 778–796. <https://doi.org/10.1287/mksc.2021.0227>
- Raj, M. (2024). More is (sometimes) merrier: Heterogeneity in demand spillovers and competition on a digital platform. *Strategic Management Journal* (John Wiley & Sons, Inc.), 1. <https://doi.org/10.1002/smj.3647>
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., & Koffel, J. B. (2021). PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, 10, 1–19.
- Reuver, M. de, Sørensen, C., & Basole, R. C. (2018). The Digital Platform: A Research Agenda. *Journal of Information Technology*, 33(2), 124–135. <https://doi.org/10.1057/s41265-016-0033-3>
- Rietveld, J., & Schilling, M. A. (2021). Platform Competition: A Systematic and Interdisciplinary Review of the Literature. *Journal of Management*, 47(6), 1528–1563. <https://doi.org/10.1177/0149206320969791>
- Schirmacher, N.-B., Ondrus, J., & Kude, T. (2017). Launch strategies of digital platforms: Platforms with switching and non-switching users.
- Schreieck, M., Ondrus, J., Wiesche, M., & Krcmar, H. (2024). A typology of multi-platform integration strategies. *Information Systems Journal*, 34(3), 828–853. <https://doi.org/10.1111/isj.12450>
- Simon, H. A. (1966). *Theories of decision-making in economics and behavioural science*. Springer.
- Srinivasan, R. (2023). *Platform Business Models for Executives*. Springer Nature Singapore. <https://doi.org/10.1007/978-981-99-4910-6>
- Sun, R., & Gregor, S. (2023). Reconceptualizing platforms in information systems research through the lens of service-dominant logic. *Journal of Strategic Information Systems*, 32(3), 1–31. <https://doi.org/10.1016/j.jsis.2023.101791>
- Tauescher, K., & Rothe, H. (2021). Optimal distinctiveness in platform markets: Leveraging complementors as legitimacy buffers. *Strategic Management Journal* (John Wiley & Sons, Inc.), 42(2), 435–461. <https://doi.org/10.1002/smj.3229>
- Tan, B., Anderson, E. G., & Parker, G. G. (2020). Platform Pricing and Investment to Drive Third-Party Value Creation in Two-Sided Networks. *Information Systems Research*, 31(1), 217–239. <https://doi.org/10.1287/isre.2019.0882>
- Tavalaci, M. M., Santaló, J., & Gawer, A. (2024). Balancing Variety and Quality: Examining the Impact of Benefit-Linked Cross-Subsidization on Multisided Platforms. *Journal of Management Studies* (John Wiley & Sons, Inc.), 1. <https://doi.org/10.1111/joms.13120>
- Tian, J., Zhao, X., & Xue, L. (2022). PLATFORM COMPATIBILITY AND DEVELOPER MULTIHOMING: A TRADE-OFF PERSPECTIVE. *MIS Quarterly*, 46(3), 1661–1690. <https://doi.org/10.25300/MISQ/2022/16369>
- Tucker, C., & Zhang, J. (2010). Growing Two-Sided Networks by Advertising the User Base: A Field Experiment. *Marketing Science*, 29(5), 805–814. <https://doi.org/10.1287/mksc.1100.0560>
- Tunc, M. M., Cavusoglu, H., & Raghunathan, S. (2021). Online Product Reviews: Is a Finer-Grained Rating Scheme Superior to a Coarser One? *MIS Quarterly*, 45(4), 2193–2234. <https://doi.org/10.25300/MISQ/2021/15586>
- van Alstyne, M., Parker, G. G., & Choudary, S. P. (2016). Pipelines, Platforms, and the New Rules of Strategy. *Harvard Business Review*, 94(4), 54–62.

- <https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=113934414&site=ehost-live&scope=site>
- Varga, S., Cholakova, M., Jansen, J. J., Mom, T. J., & Kok, G. J. (2023). From platform growth to platform scaling: The role of decision rules and network effects over time. *Journal of Business Venturing*, 38(6), N.PAG-N.PAG. <https://doi.org/10.1016/j.jbusvent.2023.106346>
- Veiga, A., Weyl, E. G., & White, A. (2017). Multidimensional Platform Design. *American Economic Review*, 107(5), 191–195. <https://doi.org/10.1257/aer.p20171044>
- Wang, Z., Guo, H., & Liu, D. (2024). Economics of Analytics Services on a Marketplace Platform. *MIS Quarterly*, 48(2), 775–824. <https://doi.org/10.25300/MISQ/2023/16452>
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, xiii–xxiii.
- Weyl, E. G. (2010). A Price Theory of Multi-Sided Platforms. *American Economic Review*, 100(4), 1642–1672. <https://doi.org/10.1257/aer.100.4.1642>
- Wu, Y., Zhang, K., & Padmanabhan, V. (2018). Matchmaker Competition and Technology Provision. *Journal of Marketing Research (JMR)*, 55(3), 396–413. <https://doi.org/10.1509/jmr.16.0423>
- Xu, X., Venkatesh, V., Tam, K. Y., & Hong, S.-J. (2010). Model of Migration and Use of Platforms: Role of Hierarchy, Current Generation, and Complementarities in Consumer Settings. *Management Science*, 56(8), 1304–1323. <https://doi.org/10.1287/mnsc.1090.1033>
- Yang, C., Guo, L., & Zhou, S. X. (2022). Customer Satisfaction, Advertising Competition, and Platform Performance. *Production & Operations Management*, 31(4), 1576–1594. <https://doi.org/10.1111/poms.13632>
- Zhang, C., Chen, J., & Raghunathan, S. (2022). Two-Sided Platform Competition in a Sharing Economy. *Management Science*, 68(12), 8909–8932. <https://doi.org/10.1287/mnsc.2022.4302>
- Zhang, K., & Sarvary, M. (2015). Differentiation with User-Generated Content. *Management Science*, 61(4), 898–914. <https://doi.org/10.1287/mnsc.2014.1907>
- Zhou, B., & Zou, T. (2023). Competing for Recommendations: The Strategic Impact of Personalized Product Recommendations in Online Marketplaces. *Marketing Science*, 42(2), 360–376. <https://doi.org/10.1287/mksc.2022.1388>
- Zhou, Z., Zhang, L., & van Alstyne, M. (2024). HOW USERS DRIVE VALUE IN TWO-SIDED MARKETS: PLATFORM DESIGNS THAT MATTER. *MIS Quarterly*, 48(1), 1–30. <https://doi.org/10.25300/MISQ/2023/17012>
- Zhu, F., & Iansiti, M. (2012). Entry into platform-based markets. *Strategic Management Journal* (John Wiley & Sons, Inc.), 33(1), 88–106. <https://doi.org/10.1002/smj.941>
- Zhu, F., & Iansiti, M. (2019). Why Some Platforms Thrive and Others Don't. *Harvard Business Review*, 97(1), 118–125.

Appendix

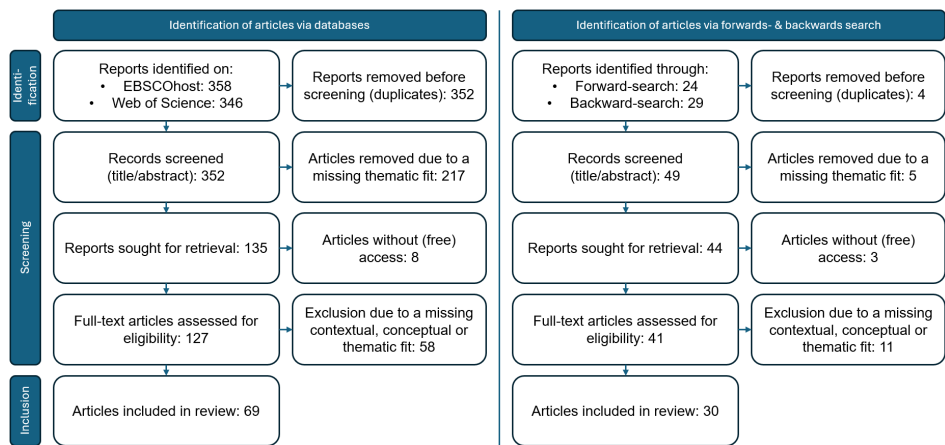


Figure 2: PRISMA 2020 flow diagram summarizing the article selection process

THE CONSTRUCTION OF SOFT SKILLS SCALE TARGETED GENERATION Z FOR IT SECTOR

ORKUN YILDIZ,^{1, 2, 3, 4} TUĞBA ELIF TOPRAK-YILDIZ^{5, 6}

¹ Izmir Democracy University, Faculty of Economics and Administrative Sciences, Izmir, Türkiye

yildiz.orkun@gmail.com

² METU, Faculty of Informatics, Ankara, Türkiye

yildiz.orkun@gmail.com

³ Saarland University, The European Institute for Advanced Behavioural Management (EIABM), Saarbrücken, Germany

yildiz.orkun@gmail.com

⁴ UNMSM, Department of Industrial Engineering, Lima, Peru

yildiz.orkun@gmail.com

⁵ Izmir Democracy University, Faculty of Arts and Sciences, Izmir, Türkiye

tugbaeliftoprak@gmail.com

⁶ Saarland University, The Institute of Language and Communication, Saarbrücken, Germany

tugbaeliftoprak@gmail.com

This research focuses on the increasing importance of soft skills in the Information Technology sector. The rise of new job roles, evolving work structures, and greater customer engagement necessitate soft skills in potential employees. Additionally, global competition, higher demands for workforce skills, and more diverse and virtually connected work environments have spurred interest in soft skills. These changes underscore the value of creativity, social, emotional, organizational, teamwork, and problem-solving abilities, collectively referred to as soft skills. Both industry and academia acknowledge that technical skills alone are insufficient in the workplace. Recognizing the importance of soft skills, this study aims to develop a scale to assess relevant soft skills within the IT context. The research involves several phases, including defining the construct, item development, scale construction, pilot testing, and performing exploratory and confirmatory factor analyses. The developed soft skills scale is anticipated to be beneficial for educators, researchers, and professionals in the IT field.

DOI

<https://doi.org/10.18690/um.fov.4.2025.19>

ISBN

978-961-286-998-4

Keywords:

soft skills,
information technology,
measurement,
scale development,
generation Z



University of Maribor Press

1 Introduction

Information Technology (IT) job advertisements usually categorise skill requirements into hard and soft skills (Ahmed et al., 2012). While hard skills require technical knowledge and skills to perform specific tasks, soft skills relate to personality traits and attitudes that impact an individual's behaviour, complementing and enhancing technical skills. Soft skills are not meant to substitute for technical skills (Cacciolatti et al., 2017). However, ideally, a candidate is expected to have a balance of both skills. There is also a growing awareness in industry and academia that more than hard skills are needed to lead to desired outcomes and employability (Ahmed et al., 2012; Dogara et al., 2020; Joseph et al., 2010). Particularly due to the emergence of new occupations, work organization methods and the growing involvement of customers, employers these days require soft skills when they consider prospective workers (Bailly & Léné, 2012). With increasing global competition and demand for higher workforce skills, businesses have grown more interested in soft skills (Cacciolatti et al., 2017). Moreover, within the last decade, workplaces have become increasingly multicultural and virtually dispersed (Castro et al., 2022). While these drastic changes take place, there has been an anticipation of shifting skill requirements, placing considerable emphasis on creativity, social, emotional, communication, critical thinking, and problem-solving skills of individuals, in other words, their soft skills (Alshare et al., 2011; Batista & Romani-Diaz, 2022).

Therefore, being a team member, operating independently, working with and without supervision is assumed to increase the demand for social and emotional skills (Green, 2004) while a growing number of studies report that employers' demands for these skills are on the rise and possessing soft skills is regarded to be crucial to employability and success in work life (Albandea & Giret, 2018; Cuesta & Budría, 2017; Deming, 2017; Ito & Kawazoe, 2015). Moreover, as the competition in the IT market increases, possessing these skills, leading to improved job performance, becomes more crucial (Bailey & Mitchell, 2006).

Despite their centrality to success in the workplace, research also reports that the skills that particularly recent graduates lack are not technical but soft skills, which feature attributes such as leadership, problem-solving and time management (Andreas, 2018). Furthermore, survey results obtained from businesses and college

graduates indicate a common belief that colleges are falling short in adequately preparing their students for the workforce (Andreas, 2018; Antón-Sancho et al., 2021; Farner & Brown, 2008; Gruzdev et al., 2018).

A review of existing research on soft skills reveals that a number of studies to date have concentrated on generating a measure targeting at soft skills (e.g., Aasheim et al., 2012; Bak & Boulocher-Passet, 2013; Buchynska & Schlichter, 2019; Dunaway, 2013; Hefley & Bottion, 2021; Jordan & Bak, 2016; Lutz & Birou, 2013; Murphy & Poist, 1991; Turner & Lowry, 2002; Van Yperen Hagedoorn et al., 2021; Yadin & Rashkovits, 2009; Zaffar & Winter, 2008). However, few of these studies have been carried out with a focus on the IT sector (e.g., Aasheim et al., 2012; Buchynska & Schlichter, 2019; Dunaway, 2013; Turner & Lowry, 2002; Van Yperen Hagedoorn et al., 2021; Zaffar & Winter, 2008). In addition, all these studies were conducted before the Covid-19 pandemic period, so they lacked a remote working perspective.

Considering the gap in the relevant literature and the significance of soft skills to employability and success in the IT sector, the present study aims to develop a scale which intends to measure soft skills in the IT context. The study recognises the significance of remote working in the IT sector which has been emphasized in the relevant literature (Aasheim et al., 2012; Bak & Boulocher-Passet, 2013; Buchynska & Schlichter, 2019; Dunaway, 2013; Hefley & Bottion, 2021; Jordan & Bak, 2016; Lutz & Birou, 2013; Murphy & Poist, 1991; Turner & Lowry, 2002; Van Yperen Hagedoorn et al., 2021; Yadin & Rashkovits, 2009; Zaffar & Winter, 2008) and incorporates this aspect in scale/item development process. All of these existing scales focus on the general workforce or various industries and are not specifically tailored to reflect the unique needs of the IT sector.

This study aims to address this gap in literature by developing a new scale specifically designed to measure soft skills in the IT sector. Bridging this theoretical gap will enable more accurate sector-specific assessments and contribute to the development of targeted training strategies. The scale developed in this study will not only have practical implications for the industry but will also provide a significant theoretical contribution. Identifying the shortcomings of existing scales will help clarify which soft skills are critical in specific scenarios within the IT sector.

Unlike previous studies that focus on technical and operational aspects, our research emphasizes the critical role of soft skills in navigating the challenges and opportunities of remote work environments. By synthesizing existing subdomains of soft skills and proposing a unified framework, our study not only clarifies these attributes but also provides a practical tool for organizations to assess and enhance their employees' soft skills. Additionally, our research contributes to the literature by bridging the conceptual gap with a domain-specific definition of soft skills, developing a validated scale to systematically evaluate these competencies, offering actionable insights for fostering soft skills in remote work settings, and focusing on Generation Z to provide unique insights into the soft skills needs of this emerging workforce.

2 Literature Background

2.1 Soft skills and IT

Traditionally, hard and soft skills have been defined by their nature: hard skills have been linked to technical and administrative abilities, while soft skills have been associated with human, conceptual, leadership, and interpersonal abilities (Bak et al., 2019). However, defining what soft skills exactly refer to would be challenging considering the diversity in the definitions, construct representations, and measurements of soft skills proposed in different areas, including psychology, education, business, and IT. Moreover, when conceptualising and operationalising the construct of soft skills, some researchers consider these skills as personality traits linked to the character of individuals (Antón-Sancho et al., 2021). On the other hand, some researchers refer to these skills as non-cognitive, emotional, social, people, interpersonal skills, applied skills, 21st-century skills, transversal skills or behavioural attitudes (Albandea & Giret, 2018; Andreas, 2018; Bailly & Léné, 2012; Bak et al., 2019; Batista & Romani-Díaz, 2022; Caeiro-Rodríguez et al., 2021; Caggiano et al., 2020; Carvalho et al., 2022). Some researchers even pay attention to specific physical properties, such as good appearance, when they consider soft skills (Bailly & Léné, 2012). Therefore, the conceptualising of soft skills may vary across disciplines and frameworks employed in studies conducted.

Despite the diversity in conceptualisations, one significant aspect of soft skills that is commonly recognised is that, unlike personality traits, they can evolve and can be learnt at schools, in the family or the workplace and are transversal-- meaning that they are cross-cutting and shared across different sectors of the labour market and academia (Andreas, 2018; Albandea & Giret, 2018; Antón- Sancho et al., 2021). Soft skills are considered important to a career that typically demands interacting with others and features a set of attributes such as communication skills, team building, problem-solving and time management skills (Andreas, 2018).

Soft skills in this study are contextualized for the IT sector, with emphasis on competencies such as virtual collaboration, self-management, and digital communication—skills increasingly vital in remote and hybrid work environments. Sources were identified through systematic searches in databases such as Scopus and Web of Science, using keywords including ‘soft skills,’ ‘Generation Z,’ and ‘IT workforce.’ Selection was based on relevance, recency, and peer-reviewed status.

2.2 Remote working

In recent years, the effects of remote working and digital transformation on business processes have become the focus of academic attention. This literature review has identified significant studies. Schuh et al. (2024) present a four-dimensional classification aimed at sustaining professional work environments shaped by post-pandemic digitalization. Hafermalz and Riemer (2016) examine how employees cope with social isolation during remote work. Chen and Hung (2022) analyze the costs perceived by different occupational groups in remote work environments. Johnson (2001) investigates the impact of remote work on individual characteristics. Zaza and Erskine (2024) examine the effects of mandatory technology-mediated remote work on employee performance during the pandemic. Mirbabaie et al. (2020) focus on the risks of digital overuse associated with remote work and explore "digital detox" approaches. Virtaneva et al. (2021) investigate the impact of remote work on knowledge workers during the COVID-19 pandemic. Debowski et al. (2022) examine remote design thinking processes. These studies address the various effects of remote work on organizations, individuals, and business processes, revealing that this topic constitutes a multidimensional and interdisciplinary research field. However, the role of soft skills in remote work environments has been largely overlooked. Communication, teamwork, problem-solving, and time management

skills have gained importance in the post-pandemic era as remote work reshapes employee interaction and collaboration dynamics. Despite their critical role in supporting productivity, adaptability, and well-being in remote work environments, these skills remain understudied, particularly in the IT industry context. Therefore, our study aims to fill this gap in the literature by conducting research to identify soft skills at the individual level, focusing on remote work.

2.3 Scales in remote working context

The increasing prevalence of remote work has necessitated the development of various measurement tools and conceptual models to assess critical competencies for employees working in virtual environments. While early research focused on productivity, organizational trust, and work-life balance (Grant et al., 2019; Charalampous et al., 2023), recent studies highlight the increasing importance of soft skills to enable effective collaboration, adaptability, and decision-making processes in remote work environments (Benligiray et al., 2024; Yildiz & Toprak-Yildiz, 2024). Soft skills, which include aspects such as communication, teamwork, emotional intelligence, and problem-solving, are considered essential for successful remote work integration, especially in knowledge-intensive fields such as Information Technology (IT). Despite this recognition, assessing soft skills in the context of remote work has been under-researched, with existing scales focusing either on general remote work competencies (Tramontano et al., 2021) or industry-specific technical skills and do not provide a comprehensive measurement framework for behavioral competencies. Additionally, there is a distinct lack of studies addressing soft skills required in remote work for Generation Z, given their experiences working remotely during the COVID-19 pandemic. This focus is critical because Generation Z represents a key demographic group expected to dominate the remote workforce in the near future. The COVID-19 pandemic has provided Generation Z with the opportunity to gain first-hand experience with remote work, especially during their university education, making them a key group for understanding the future of remote work dynamics. This gap in literature prevents a comprehensive understanding of how soft skills operate in remote work environments and how they can be effectively measured and developed.

In response to these gaps, the Soft Skills Remote Work Culture Acceptance Model (SSRWCA) proposed by Yildiz and Toprak-Yildiz (2024) provides a theoretical basis for the integration of soft skills into remote work adaptation assessments. Based on Causal Action Theory (RAT), this model highlights the role of decision-making skills, behavioral skills, management skills, and negotiation skills in shaping employees' abilities to effectively engage in remote work environments. By incorporating MacKenzie et al.'s (2011) best practices in scale development, this study aims to fill an important gap in the literature by developing a scale that will strengthen the methodological rigor of assessing soft skills in the IS field.

3 Methodology

3.1 The sample

The sample utilized in the current study featured 139 Management Information Systems department undergraduates attending a university in a metropolitan city in Türkiye. The age of the participants ranged between 18-23. Participation in the study was voluntary.

The item development process involved literature synthesis and expert input. Exploratory and confirmatory factor analyses were conducted to ensure structural validity, with detailed factor loadings and fit indices reported.

3.2 Data collection and analysis

Data were collected using an online platform to which the participants had access. The participants received a link to the online platform, which took them to the scale administered in digital format. It took around 40 minutes for participants to answer all items on the scale. The data were initially analysed using exploratory factor analysis (EFA) to investigate the appropriateness and structure of the data, as well as scale and item properties. In the second phase, the data were analysed using confirmatory factor analysis (CFA). Confirmatory factor analysis enables researchers propose and test theoretical measurement models which account for the links between the underlying latent variables and the empirical measures (Knoke, 2005).

3.3 Construction of the Scale

The construction of the scale featured several stages. The first stage was the operationalisation of the construct at hand, soft skills. As previously noted, the construct of soft skills has been conceptualised and measured in various ways in the relevant literature. Therefore, a systematic review of existing research in IS, IT and business was undertaken to operationalise the construct. Implications and findings provided by relevant research formed the basis of construct definition, which also guided all item writing efforts. Initially, an item pool which included around 250 items was created. The items were subject to a qualitative analysis by three experts. The final form of the scale included four subscales: decision-making skills subscale, behavioural skills subscale, management skills subscale, and negotiation skills subscale. The subscales which constitute the soft skills scale were based on these four constructs (i.e., decision making skills, behavioural skills, management skills, and negotiation skills) underlying the main construct of soft skills.

While several soft skills scales exist, this study addresses the unique behavioral and communication patterns of Generation Z, particularly within the IT sector, which are not adequately captured by existing instruments.

4 Results

4.1 Initial quantitative analyses and exploratory factor analysis

Before conducting the main confirmatory factor analysis, an exploratory factor analysis (EFA) was initially undertaken to ensure the appropriateness and structure of the data, scale and item properties. Analyses were performed separately for each subscale of interest in the same manner by using SPSS (IBM SPSS Statistics 25). These were the analyses performed: Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy test, Bartlett's test of Sphericity, and Cronbach alpha for internal consistency. The results of the KMO and Bartlett's test of Sphericity indicated the appropriateness of the data for further EFA. The internal reliability scores for the decision-making skills subscale (25 items), behavioural skills subscale (18 items), management skills subscale (23 items), and negotiation skills subscale (4 items) were .92, .90, .93, and .65, respectively. The internal reliability values were found to be

very strong for the decision-making skills subscale, behavioural skills subscale, management skills subscale and acceptable for the negotiation skills subscale.

4.2 Confirmatory factor analysis

CFA is a statistical method that helps researchers investigate the underlying factor structure of a particular set of observed variables. CFA makes it possible to evaluate if there are significant relationships between observed variables and the unobservable latent variables (i.e., constructs) that underlie them (Hair et al., 2018). CFA analyses were conducted using Smart PLS software using the maximum likelihood (ML) estimation method. Information about the underlying structure of the subscales, R^2 values, overall item properties and model fit has been presented in Table 1.

Table 1: Sample table

	Decision-making skills subscale	Behavioural skills subscale	Management skills subscale	Negotiation skills subscale
1 Underlying structure	Problem solving skills, organisational skills, planning skills, flexibility, cognitive skills, initiative skills	Communication skills, time management skills, motivation and enthusiasm, stress management skills	People managementskills, teamwork skills, leadership skills, collaborationskills, managing complexity and change	Negotiation skills
2 R^2	Problem solving skills (.65), organisational skills (.65), planning skills (.61), flexibility (.65), cognitive skills (.66), initiative skills (.62)	Communication skills (.65), time management skills (.49), motivation and enthusiasm (.47), stress management skills (.53)	People managementskills (.64), teamwork skills (.78), leadership skills (.80), collaboration skills (.79), managing complexity and change (.76)	Negotiation skills(.85)
3 Item properties	All item loadings ranging between .60-.84	All item loadingsranging between .64-.83	All item loadings ranging between .60-.80	All item loadings ranging between .63-.78
4 Model fit	SRMR 0.09	SRMR 0.08	SRMR 0.08	SRMR 0.10

Overall, the R^2 values, which represent the proportion of variance in the observed variable explained by the latent factor, were found to be indicating a strong relationship between the factors and the items in most of the cases. The item loadings, which indicate effectively each item measures the targeted latent factor, indicated good representation since all item loadings varied between .60 and .84. Finally, the Standardized Root Mean Square Residual (SRMR) values, which represent the discrepancy between the observed correlations based on the sample data and the predicted correlations based on the model, were computed to assess the model-to-data fit. Values ranging from 0.10-0.08 indicated an acceptable model-to-data fit for four subscales generated.

4 Discussions and Conclusions

The present study aimed to develop a soft skills scale that would measure intended skills and attributes in the context of IT. The scale consisted of four subscales targeting critical soft skills and attributes specified in the relevant literature. These subscales were decision-making skills subscale (Bak & Boulocher-Passet, 2013; Jordan & Bak, 2016; Lutz & Birou, 2013; Murphy & Poist, 1991), behavioural skills subscale (Derwik & Hellstrom, 2017; Dubey & Gunasekaran, 2015; Jordan & Bak, 2016; Zuo et al., 2018), management skills subscale (Jordan & Bak, 2016; Kovács et al., 2012, Lutz & Birou, 2013), and negotiation skills (Jordan & Bak, 2016; Kovács et al., 2012). The results of the EFA and CFA analyses indicated that the subskills functioned effectively, as evident in the item loadings ranging between .60 and .84 and acceptable R^2 values representing the proportion of variance in the observed variables explained by the underlying factors in our model. The final version of the scale featured 70 items. The item number per subscale and subdimension is as follows. Decision-making skills subscale featured problem-solving skills (4 items), organisational skills (4 items), planning skills (4 items), flexibility (4 items), cognitive skills (4 items), initiative skills (5 items). The behavioural skills subscale featured communication skills (5 items), time management skills (4 items), motivation and enthusiasm (4 items), and stress management skills (5 items). Management skills subscale featured people management skills (5 items), teamwork skills (5 items), leadership skills (5 items), collaboration skills (4 items), and managing complexity and change (4 items). Finally, the negotiation skills subscale featured negotiation skills (4 items).

Each subdimension of the scale effectively reflects the dynamics of the IT industry. Given the increasing prevalence of hybrid work models and the growing importance of virtual collaboration, this scale provides a valuable roadmap for developing soft skills in post-pandemic workforce conditions. In hybrid work environments, the ability of employees to collaborate, communicate effectively within teams, and demonstrate leadership is becoming increasingly important. This study develops a sector-specific scale for measuring remote work and virtual collaboration skills, thereby contributing to the IS literature (Tromontano et al., 2021).

Negotiation skills, in particular, are among the most critical skills for IT professionals. Including subdimensions like conflict resolution in the negotiation skills subscale is highly meaningful, considering the importance of these skills. This approach ensures that IT professionals are equipped not only with technical skills but also with the ability to manage complex interpersonal relationships and resolve conflicts within teams. The strong emphasis on negotiation skills on this scale provides an important resource for evaluating leadership, collaboration, and changing management skills within the sector.

In conclusion, this study fills an important gap in IS literature and introduces a new measurement tool for evaluating sector-specific soft skills. While the literature often focuses on digital competencies and technostress (Molino et al., 2020; Yıldız & Toprak-Yıldız, 2024), this study provides an in-depth examination of decision-making, behavioral skills, management, and negotiation skills, making a significant contribution to the development of professionals in the IT sector.

The findings are discussed in relation to prior literature, highlighting both convergences and divergences. The study's main limitation is the homogeneous student sample, which restricts generalizability. Future research should validate the scale with broader and professional populations.

Acknowledgements

This study was supported by the Research Fund of the Izmir Democracy University. Project Number: HIZDEP-İİBF/2203.

References

- Aasheim, C., Shropshire, J., Li, L., & Kadlec, C. (2012). Knowledge and skill requirements for entry-level IT workers: A longitudinal study. *Journal of Information Systems Education*, 23(2), 193–204. <https://aisel.aisnet.org/jise/vol23/iss2/8/>
- Ahmed, F., Capretz, L. F., & Campbell, P. (2012). Evaluating the demand for soft skills in software development. *IT Professional*, 14(1), 44–49.
- Albandea, I., & Giret, J. F. (2018). The effect of soft skills on French post-secondary graduates' earnings. *International Journal of Manpower*, 39(6), 782–799.
- Alshare, K. A., Lane, P. L., & Miller, D. (2011). Business communication skills in information systems (IS) curricula: Perspectives of IS educators and students. *Journal of Education for Business*, 86(3), 186–194.
- Andreas, S. (2018). Effects of the decline in social capital on college graduates' soft skills. *Industry and Higher Education*, 32(1), 47–56.
- Antón-Sancho, Á., Vergara, D., & Fernández-Arias, P. (2021). Self-assessment of soft skills of university teachers from countries with a low level of digital competence. *Electronics*, 10(20), 2532.
- Bailey, J., & Mitchell, R. B. (2006). Industry perceptions of the competencies needed by computer programmers: technical, business, and soft skills. *Journal of Computer Information Systems*, 47(2), 28–33.
- Bailly, F., & Léné, A. (2012). The personification of the service labour process and the rise of soft skills: A French case study. *Employee Relations*, 35(1), 79–97.
- Bak, O., & Boulocher-Passet, V. (2013). Connecting industry and supply chain management education: Exploring challenges faced in a SCM consultancy module. *Supply Chain Management: An International Journal*, 18(4), 468–479.
- Bak, O., Jordan, C., & Midgley, J. (2019). The adoption of soft skills in supply chain and understanding their current role in supply chain management skills agenda: A UK perspective. Benchmarking. *An International Journal*, 26(3), 1063–1079.
- Batista, M. P., & Romani-Dias, M. (2022). AACSB international accreditation as a catalyst for soft skills in business schools. *Journal of Education for Business*, 97(4), 213–219.
- Basol, O., & Çömlekçi, M. (2022). Validity and reliability study of remote work attitude scale. *Journal Of Mehmet Akif Ersoy University Economics and Administrative Sciences Faculty*, 9(1), 243–261. <https://doi.org/10.30798/makuiibf.862439>
- Benligiray, S., Güngör, A., & Akbas, I. (2024). Measuring remote working skills: Scale development and validation study. *PLOS One*, 19(4). <https://doi.org/10.1371/journal.pone.0299074>
- Buchynska, T., & Schlichter, B. R. (2019). *Soft Skills in Co-Sourcing of Information System Development Services*. MCIS 2019, 14. <https://aisel.aisnet.org/mcis2019/14/>
- Cacciolatti, L., Lee, S. H., & Molinero, C. M. (2017). Clashing institutional interests in skills between government and industry: An analysis of demand for technical and soft skills of graduates in the UK. *Technological Forecasting and Social Change*, 119, 139–153.
- Caeiro-Rodríguez, M., Manso-Vázquez, M., Mikic-Fonte, F. A., Llamas-Nistal, M., Fernández-Iglesias, M. J., Tsalapatas, H., ... & Sørensen, L. T. (2021). Teaching soft skills in engineering education: An European perspective. *IEEE Access*, 9, 29222–29242.
- Caggiano, V., Schleutker, K., Petrone, L., & Gonzalez-Bernal, J. (2020). Towards identifying the soft skills needed in curricula: Finnish and Italian students' self-evaluations indicate differences between groups. *Sustainability*, 12(10), 4031.
- Carvalho, C., & Almeida, A. C. (2022). The adequacy of accounting education in the development of transversal skills needed to meet market demands. *Sustainability*, 14(10), 5755.
- Castro, M., Barcaui, A., Bahli, B., & Figueiredo, R. (2022). Do the Project Manager's Soft Skills Matter? Impacts of the Project Manager's Emotional Intelligence, Trustworthiness, and Job Satisfaction on Project Success. *Administrative Sciences*, 12(4), 141.

- Charalampous, M., Grant, C., & Tramontano, C. (2023). Getting the measure of remote e-working: A revision and further validation of the E-work life scale. *Employee Relations*. <https://doi.org/10.1108/ER-11-2021-0483>
- Chen, X., & Hung, W.-H. (2022, October 17). Understanding work behaviors in remote work environments during the COVID-19 pandemic: Transaction cost theory perspective. *ICEB 2022 Proceedings (Bangkok, Thailand)*. <https://aisel.aisnet.org/iceb2022/53>
- Cuesta, M. B., & Budría, S. (2017). Unemployment persistence: How important are non cognitive skills? *Journal of Behavioral and Experimental Economics*, 69, 29-37.
- Debowski, N., Tavanapour, N., & Bittner, E. A. (2022). Conversational agents in creative work—a systematic literature review and research agenda for remote design thinking. *Proceedings of the Pacific Asia Conference on Information Systems*.
- Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4), 1593-1640.
- Derwik, P., & Hellström, D. (2017). Competence in supply chain management: a systematic review. *Supply Chain Management: An International Journal*, 22(2), 200-218.
- Dogara, G., Saud, M. S. B., & Kamin, Y. B. (2020). Work-based learning conceptual framework for effective incorporation of soft skills among students of vocational and technical institutions. *IEEE Access*, 8, 211642-211652.
- Dubey, R., & Gunasekaran, A. (2015). Shortage of sustainable supply chain talent: an industrial training framework. *Industrial and Commercial Training*, 47(2), 86-94.
- Dunaway, M. M. (2013). IS learning: The impact of gender and team emotional intelligence. *Journal of Information Systems Education*, 24(3), 189–202. <https://aisel.aisnet.org/jise/vol24/iss3/4/>
- Farner, S. M., & Brown, E. E. (2008). College students and the work world. *Journal of Employment Counseling*, 45(3), 108-114.
- Grant, C., Wallace, L., Spurgeon, P., Tramontano, C., & Charalampous, M. (2019). Construction and initial validation of the E-Work Life Scale to measure remote e-working. *Employee Relations*, 41(1), 16–33. <https://doi.org/10.1108/ER-09-2017-0229>
- Green, F. (2004). Why has work effort become more intense? *Industrial Relations: A Journal of Economy and Society*, 43(4), 709-741.
- Gruzdev, M. V., Kuznetsova, I. V., Tarkhanova, I. Y., & Kazakova, E. I. (2018). University graduates' soft skills: The employers' opinion. *European Journal of Contemporary Education*, 7(4), 690- 698.
- Hafermalz, E., & Riemer, K. (2016, June 15). The Work of Belonging Through Technology in Remote Work: A Case Study in Tele-Nursing. *Research Papers*. https://aisel.aisnet.org/ecis2016_rp/106
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate Data Analysis* (8 ed.). Cengage Learning.
- Hefley, W. E., & Botton, M. (2021). Skills of junior project management professionals and project success achieved by them. *International Journal of Information Systems and Project Management*, 9(1), 56–75. <https://aisel.aisnet.org/ijispm/vol9/iss1/4/>
- Ito, H., & Kawazoe, N. (2015). Active learning for creating innovators: Employability skills beyond industrial needs. *International Journal of Higher Education*, 4(2), 81-91.
- Johnson, R. (2001, December 31). Toward a Deeper Understanding of Alternative Work Arrangements: The Impact of Core Self-Evaluations on Remote Work. *AMCIS 2001 Proceedings*. <https://aisel.aisnet.org/amcis2001/364>
- Jordan, C., & Bak, O. (2016). The growing scale and scope of the supply chain: A reflection on supply chain graduate skills. *Supply Chain Management: An International Journal*, 21(5), 610–626. <https://www.emerald.com/insight/content/doi/10.1108/SCM-02-2016-0059/full/html>
- Joseph, D., Ang, S., Chang, R. H., & Slaughter, S. A. (2010). Practical intelligence in IT: Assessing soft skills of IT professionals. *Communications of the ACM*, 53(2), 149-154.
- Knoke, D. (2005). Structural equation models. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (pp. 689–695). New York: Elsevier

- Kovács, G., Tatham, P., & Larson, P. D. (2012). What skills are needed to be a humanitarian logistician?. *Journal of Business Logistics*, 33(3), 245–258.
- Lutz, H., & Birou, L. (2013). Logistics education: A look at the current state of the art and science. *Supply Chain Management: An International Journal*, 18(4), 455–467. <https://www.emerald.com/insight/content/doi/10.1108/SCM-08-2012-0269/full/html>
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 293–334.
- Mirbabaie, M., Marx, J., Braun, L.-M., & Stieglitz, S. (2020, January 1). Digital Detox – Mitigating Digital Overuse in Times of Remote Work and Social Isolation. *ACIS 2020 Proceedings*. <https://aisel.aisnet.org/acis2020/87>
- Molino, M., Ingusci, E., Signore, F., Manuti, A., Giancaspro, M., Russo, V., Zito, M., & Cortese, C. (2020). Wellbeing Costs of Technology Use during Covid-19 Remote Working: An Investigation Using the Italian Translation of the Technostress Creators Scale. *Sustainability*, 12(15). <https://doi.org/10.3390/su12155911>
- Murphy, P. R., & Poist, R. F. (1991). Skill requirements of senior-level logisticians: practitioner perspectives. *International Journal of Physical Distribution & Logistics Management*, 21(3), 3–14.
- Sarkar, P., Posen, A., & Etemad, A. (2023). *AVCAffe: A Large Scale Audio-Visual Dataset of Cognitive Load and Affect for Remote Work* (B. Williams, Y. Chen, & J. Neville, Eds.; WOS:001243759700009; pp. 76–85).
- Schuh, G., Herkenrath, C., Boos, W., Hoeborn, G., & Boenig, J. (2024, January 3). How to establish a lasting remote work concept in organizations: A classification for the operational design of remote work. *Hawaii International Conference on System Sciences 2024 (HICSS-57)*. https://aisel.aisnet.org/hicss-57/cl/distributed_collaboration/4
- Tramontano, C., Grant, C., & Clarke, C. (2021). Development and validation of the e-Work Self-Efficacy Scale to assess digital competencies in remote working. *Computers in Human Behavior Reports*, 4. <https://doi.org/10.1016/j.chbr.2021.100129>
- Turner, R., & Lowry, G. (2002). *Towards a profession of information systems and technology: The relative importance of “Hard” and “Soft” skills for IT practitioners*. In: 2002 Information Resources Management Association International Conference, 19 May 2002–22 May 2002, Seattle, Washington, USA.
- Van Yperen Hagedoorn, J. M., Smit, R., Versteeg, P., & Ravesteyn, P. (2021). *Soft skills of the chief information security officer*. BLED 2021 Proceedings, 31.
- Virtaneva, M., Feshchenko, P., Hossain, A., Kariluoto, A., Himmanen, J., Kaitila, P., Kultanen, J., Kemell, K.-K., & Abrahamsson, P. (2021, June 17). COVID-19 Remote Work: Body Stress, Self-Efficacy, Teamwork, and Perceived Productivity of Knowledge Workers. *12th Scandinavian Conference on Information Systems*. <https://aisel.aisnet.org/scis2021/8>
- Yadin, A., & Rashkovits, R. (2009). *New Perspective On Developing Technical And Soft Skills For IS Graduates—The Case Of System Analysis And Design Workshop*. MCIS 2009 Proceedings, 5. <https://aisel.aisnet.org/mcis2009/5/>
- Yildiz, O., & Toprak-Yildiz, T. E. (2024, October 3). Acceptance of Remote Working Culture: Developing A Conceptual Model. *MCIS 2024 Proceedings*. <https://aisel.aisnet.org/mcis2024/24>
- Zaffar, M. A., & Winter, S. J. (2008). *Minding the IS soft skills gap: Evidence of discourse convergence and organizational field structure*. ICIS 2008 Proceedings, 69. <https://aisel.aisnet.org/icis2008/69/>
- Zaza, S., & Erskine, M. (2024). Thriving with Remote Work During a Crisis: Effects of a Mandatory Technology-mediated Work Environment on Job Outcomes. *AIS Transactions on Human-Computer Interaction*, 16(1), 110–138. <https://doi.org/10.17705/1thci.00202>
- Zuo, J., Zhao, X., Nguyen, Q. B. M., Ma, T., & Gao, S. (2018). Soft skills of construction project management professionals and project success factors: A structural equation model. *Engineering, Construction and Architectural Management*, 25(3), 425–442.

BARRIERS TO THE ADOPTION OF DIGITAL HEALTH SOLUTIONS AMONG TYPE 2 DIABETIC PATIENTS DURING THE COVID-19 PANDEMIC

AYESHA THANTHRIGE,¹ THU HA DANG,²

NILMINI WICKRAMASINGHE¹

¹ La Trobe University, School of Computing, Engineering and Mathematical Sciences,
Victoria, Australia

thanthrige87@gmail.com, n.wickramasinghe@latrobe.edu.au

² Swinburne University of Technology, School of Health Sciences, Victoria, Australia
thuha.dang@latrobe.edu.au

The world faces unexpected disruptions, such as the COVID-19 pandemic. These situations expose gaps, specifically in healthcare, and highlight the need for digital health solutions (DHSs). However, managing chronic conditions like Type 2 Diabetes Mellitus (T2DM) becomes challenging during such crises, especially in the U.S. where T2DM is highly prevalent. These challenges offer lessons for better preparedness. Prior studies have overlooked the specific barriers to DHS adoption among T2DM patients. This study addresses this gap by investigating adoption barriers for U.S. T2DM patients. Using the Unified Theory of Acceptance and Use of Technology (UTAUT) and a PRISMA-guided systematic review, we identified key barriers such as infrastructure limitations, usability issues, socioeconomic disparities, privacy concerns, and perceived usefulness limitations. Findings recommend enhancing digital infrastructure, simplifying interfaces, and strengthening security measures to improve adoption and long-term use. These insights guide DHS implementation for T2DM, with broader implications for chronic disease management globally, addressing healthcare disruptions beyond COVID-19.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.20](https://doi.org/10.18690/um.fov.4.2025.20)

ISBN

978-961-286-998-4

Keywords:

digital health solutions,
type 2 diabetes mellitus,
COVID-19 pandemic,
UTAUT,
healthcare



University of Maribor Press

1 Introduction

The COVID-19 pandemic disrupted global healthcare systems, especially revealing weaknesses in chronic disease management, particularly for Type 2 Diabetes Mellitus (T2DM). In the United States (U.S.), T2DM affects 38.4 million people (11.6% of the population) as of 2023, with an additional 97.6 million showing prediabetic symptoms (CDC, 2024). T2DM's high prevalence also led to high cost and demand for continuous monitoring. It places a significant burden on healthcare systems and the situation further worsened due to restricted access during the pandemic (Filip et al., 2022). Digital health solutions (DHSs), such as telemedicine and mobile apps, were needed to ensure continuity care, especially for T2DM, enabling remote consultations and glucose monitoring (Cimini et al., 2022). However, several unique adoption challenges emerged, particularly in the U.S. context, due to healthcare policy gaps, racial disparities, and infrastructure constraints (Tewari et al., 2023; Waizinger et al., 2022).

On the other hand, T2DM patients faced heightened risks during the pandemic, with increased complications due to disrupted care (Upsher et al., 2022). DHSs mitigated these issues by facilitating remote monitoring, but adoption varied across U.S. populations due to digital literacy, socio-economic status, and trust in technology. Given the U.S.'s uniquely fragmented healthcare system and the disparities it creates this study explores how T2DM impacts diverse communities' access to DHSs (Stange, 2009; Steinhardt et al., 2021). Examining these barriers during COVID-19 provides insights into future disruptions, ensuring DHSs remain relevant for chronic disease management post-pandemic.

Prior research on DHS adoption for T2DM in the U.S. often examines barriers in isolation (Petersen et al., 2020), such as technological constraints, without a unified framework. U.S. specific factors, like insurance barriers, racial inequities and policy fragmentation, also important and remain underexplored (Stange, 2009; Villagra et al., 2019). This study addresses these gaps by systematically analyzing barriers using the Unified Theory of Acceptance and Use of Technology (UTAUT) framework (Venkatesh et al., 2003) (Refer appendix A). UTAUT's application in this context is novel, as its use in digital health for T2DM during crises is limited in the U.S. context (Petersen et al., 2020).

The objective is to identify barriers to DHS adoption among T2DM patients during COVID-19 in the U.S., informing policymakers on equitable solutions. The study was narrowed to pandemic and U.S. context as it accelerated digital health use, highlighting unique challenges like access and usability and U.S. has a high T2DM prevalence. UTAUT's constructs guide the analysis of barriers systematically such as infrastructure and privacy concerns, bridging theoretical and practical insights.

Practically, it provides more important scalable strategies such as enhancing infrastructure and addressing inequities to improve DHS access for T2DM management, benefiting policymakers, providers, and developers, contributing new insights to chronic disease management literature with implications beyond the pandemic.

2 Methodology

2.1 Study Design

This systematic literature review examines barriers to DHSs adoption among U.S. T2DM patients during the COVID-19 pandemic (2020–2023), using the UTAUT framework (Venkatesh et al., 2003). The review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure transparency and rigor (Page et al., 2021). Refer figure 1.

2.2 Research Question

What are the key barriers to adopting DHSs among T2DM patients in the U.S., and how were these barriers affected by the COVID-19 pandemic?

2.3 Search Strategy

The research question was broken into four concepts focusing on four concepts namely T2DM management, DHSs, COVID-19, and adoption barriers. Keywords included (1) T2DM (“Type 2 diabetes mellitus”, “diabetes management”), (2) DHSs (“digital health”, “telemedicine”, “remote monitoring”, “mobile health”), (3) COVID-19 (“COVID-19”, “pandemic”), and (4) barriers (“adoption”, “barriers”, “challenges”). Search strings combined these using Boolean operators (e.g., “Type 2

diabetes AND telemedicine AND COVID-19 AND barriers”) across MEDLINE, EMBASE, and Scopus, incorporating MeSH terms and keywords.

2.4 Study Selection

Studies were selected based on predefined criteria (Table 1). Two reviewers independently screened titles and abstracts, resolving discrepancies via discussion or a third reviewer. A full-text review was conducted for the screened articles by assessing in detail by two independent reviewers and the discrepancies between the reviewers were resolved through a third reviewer.

Table 1: Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Studies on T2DM diagnosis/management during COVID-19 using DHSs	Non-T2DM or non-COVID-19 studies
U.S. based community T2DM populations	Non-U.S. populations
Published between 2020 to 2023	Studies published before 2020
Published in English language	Non-English publications

2.5 Data Extraction

A standardized form captured study design, setting, population, interventions, outcomes, and barriers. Two researchers extracted data independently, resolving discrepancies through discussion.

2.6 Data Synthesis

Thematic analysis followed Braun and Clarke’s (2006) six-step process (1) data familiarization, (2) coding barriers using UTAUT constructs (3) theme generation, (4) theme review, (5) theme definition, and (6) reporting. UTAUT constructs (Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Facilitating Conditions (FC), Behavioral Intention (BI), and Usage Behavior (UB) informed initial coding, with themes refined using NVivo software (Braun & Clarke, 2006).

3 Results

3.1 Study Selection and Characteristics

The review identified 426 publications and 30 duplicates were removed, leaving 396 for screening. After title/abstract screening, 380 were excluded, and 16 articles underwent full-text review resulting in 15 studies for final analysis. Refer figure 1. The small sample reflects the specific U.S. and COVID-19 focus but ensures relevance (Rethlefsen et al., 2021). The analysis shows 50% of the studies were conducted in mixed (both urban and rural) settings, indicating a broad applicability of DHSs across diverse environments. Urban settings accounted for 42% of the studies, highlighting a significant focus on densely populated areas where healthcare infrastructure is generally more accessible.

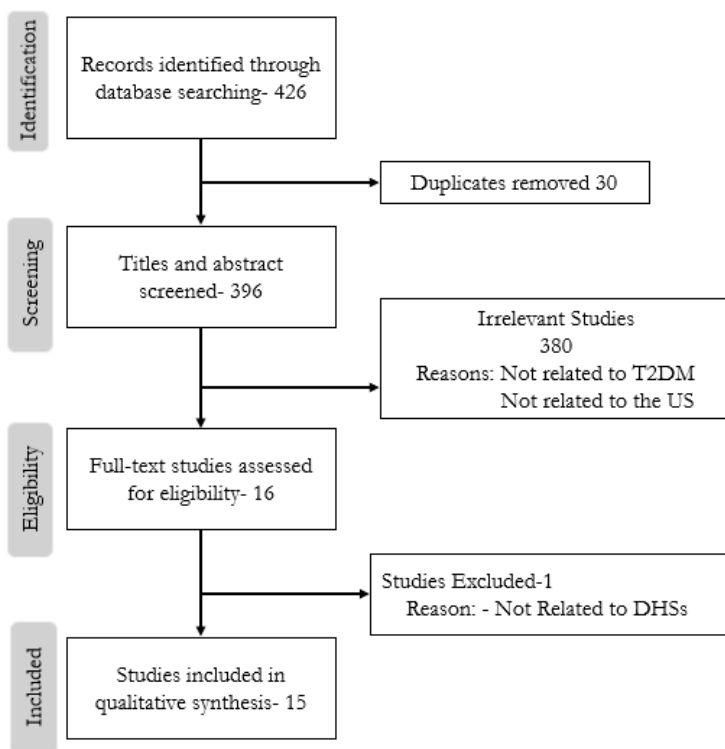


Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram of the study selection process

3.2 DHSs for T2DM Management

DHSs such as telemedicine, mobile apps, and remote monitoring, enabled remote consultations, glucose monitoring, and lifestyle interventions have been utilized during COVID-19. These tools ensured care continuity despite restricted in-person access, supporting T2DM management in urban and rural settings. Telemedicine dominated (50%) due to its scalability, while mobile apps facilitated patient engagement.

3.3 Thematic Analysis of Barriers

Five themes emerged namely infrastructure constraints, usability challenges, socio-economic disparities, privacy and trust issues, and perceived usefulness limitations, supported by direct evidence (Table 2). BI and UB were not standalone themes but impacted across barriers, as FC, EE, SI, and PE reduced adoption intentions and usage.

These barriers are visually summarized in a Fishbone Diagram (Figure 2), illustrating their contributions to the limited adoption of DHSs. UTAUT constructs shaped thematic synthesis, mapping barriers to PE, EE, SI, FC, BI, UB (Venkatesh et al., 2003). For example, “limited internet” was coded as an FC barrier which led to reducing BI. Infrastructure constraints (FC) like unstable internet lowered BI and UB, especially for rural patients (He et al., 2023). Usability challenges (EE) increased perceived effort, reducing BI among seniors. Privacy and trust issues reduced engagement in virtual diabetes programs due to 47% of African American participants distrusting health information. Socio-economic disparities, with 60% of low-income patients facing device access issues, and 27.1% reporting financial stress, amplified infrastructure barriers (Chunara et al., 2021; Steinhardt et al., 2021).

Table 2: Detailed Analysis of Barriers

Barrier Theme	UTAUT Construct	Description and Examples	Source(s)
Infrastructure Constraints	FC	Limited internet access and unstable connections restrict video teleconsultations for diabetes management, particularly in rural areas. For example, patients struggle with consistent connectivity for virtual appointments. "Not all families are comfortable using telehealth technology and others lack access to a stable internet connection."	(Monaghan & Marks, 2020; Petersen et al., 2020; Steinhardt et al., 2021; Vaughan et al., 2022)
Infrastructure Constraints	FC	Lack of access to devices (e.g., computers, smartphones) exacerbates the digital divide for diabetes patients. "During the COVID-19 pandemic, populations with health disparities are the same as those that are having difficulty accessing virtual health care. Access barriers include... not having the necessary hardware and software..."	(Kerr & Sabharwal, 2021; Steinhardt et al., 2021)
Infrastructure Constraints	FC	Location-based logistical barriers (e.g., lack of transportation, distance to facilities) hinder telehealth access in rural areas. For example, patients may need in-person visits for device setup.	(Merrill et al., 2022)
Infrastructure Constraints	FC	Lack of interoperability in health devices and data systems complicates virtual diabetes care. "An often-touted reason for reluctance to embrace interoperability is the perception that interoperability will hamper innovation..."	(Kerr & Sabharwal, 2021)
Infrastructure Constraints	FC	Logistical challenges for providers, such as inadequate clinic resources and training, limit telehealth implementation. For example, nurse practitioners face challenges adapting to virtual care.	(Vaughan et al., 2022; Waizinger et al., 2022)
Infrastructure Constraints	FC	Travel requirements and distance to facilities pose barriers, particularly in underserved areas. For example, patients may need in-person visits for initial telehealth setup.	(Casas et al., 2023; Lam et al., 2020; Steinhardt et al., 2021)
Usability Challenges	EE	Older adults with diabetes prefer traditional methods (e.g., landlines) and face physical impairments (e.g., hearing, vision), increasing effort for telehealth use. For example, 36.7% of U.S. seniors used landlines in 2021.	(Casas et al., 2023; He et al., 2023; Lam et al., 2020)

Barrier Theme	UTAUT Construct	Description and Examples	Source(s)
Usability Challenges	EE	Young adults with diabetes faced disruptions (e.g., college closures, unemployment, low digital literacy), increasing effort for virtual diabetes management. "Young adults with diabetes are particularly vulnerable during the COVID-19 pandemic..."	(Monaghan & Marks, 2020; Upsher et al., 2022)
Usability Challenges	EE	Telehealth platforms with technical limitations (e.g., complex interfaces) increase effort, particularly for low-digital-literacy patients. A layered approach is needed for self-efficacy.	(Beks et al., 2022; Petersen et al., 2020; Upsher et al., 2022)
Usability Challenges	EE	Lack of smartphone-focused telehealth platforms increases effort for diverse diabetes populations. "Smartphone-based internet access has substantially reduced the digital divide..."	(Kerr & Sabharwal, 2021)
Usability Challenges	EE	Telehealth's inability to conduct thorough physical examinations (e.g., foot exams) increases effort for accurate diabetes management.	(Casas et al., 2023)
Usability Challenges	EE	Providers require new expertise for virtual diabetes care, increasing effort for tasks like patient education.	(Merrill et al., 2022)
Usability Challenges	EE	Complex telehealth interfaces hinder self-management for type 2 diabetes patients, particularly those with limited digital skills. For example, patients struggle with remote glucose monitoring tools.	(Upsher et al., 2022)
Socio-Economic Disparities	FC	Financial instability and stress (e.g., job loss, insurance concerns) hindered diabetes management and telehealth access. "Financial instability during the pandemic increased worries about contracting COVID-19..."	(Monaghan & Marks, 2020)
Socio-Economic Disparities	FC	Community clinics face resource constraints (e.g., time, participant recruitment) in underserved areas with high diabetes prevalence.	(Beks et al., 2022)
Socio-Economic Disparities	FC	Healthcare disparities limit telehealth access for low-income and minority diabetes patients. For example, underserved areas face economic and systemic inequities.	(Chunara et al., 2021)

Barrier Theme	UTAUT Construct	Description and Examples	Source(s)
Socio-Economic Disparities	FC	Low health insurance literacy among minority and low-income diabetes patients creates barriers to telehealth access. For example, patients may not understand virtual visit coverage.	(Villagra et al., 2019)
Privacy and Trust Issues	SI	Privacy and cybersecurity concerns (e.g., inconsistent privacy notices) undermine trust, particularly among older adults. "A major concern for maintaining and developing trust is the serious issue of privacy..."	(Kerr & Sabharwal, 2021)
Privacy and Trust Issues	SI	Telehealth visits for mental health aspects of diabetes care require private spaces, which many patients lack. For example, crowded households limit privacy.	(He et al., 2023)
Privacy and Trust Issues	SI	Historical healthcare mistreatment contributes to distrust among minority diabetes populations. "Absence of trust in the healthcare system... is based on historical precedents."	(Kerr & Sabharwal, 2021)
Perceived Usefulness Limitations	PE	Telehealth's inability to support physical examinations reduces perceived usefulness. "Telehealth services currently face limitations in conducting thorough physical examinations..."	(Casas et al., 2023; He et al., 2023)
Perceived Usefulness Limitations	PE	Discordant health beliefs, particularly among Black Americans, reduce telehealth's perceived usefulness. "Among Americans with diabetes who are Black, discordant health beliefs are common..."	(Kerr & Sabharwal, 2021; Tewari et al., 2023)

PE was impacted when DHSs couldn't support thorough exams, lowering UB (Casas et al., 2023). UTAUT assumes adequate infrastructure, but U.S.-specific policy gaps (e.g., telehealth reimbursement limits) worsened barriers, requiring framework adaptation (Patel et al., 2021). This U.S.-focused analysis highlights racial disparities (e.g., African Americans facing higher costs), offering new insights into policy-driven barriers (Chunara et al., 2021).

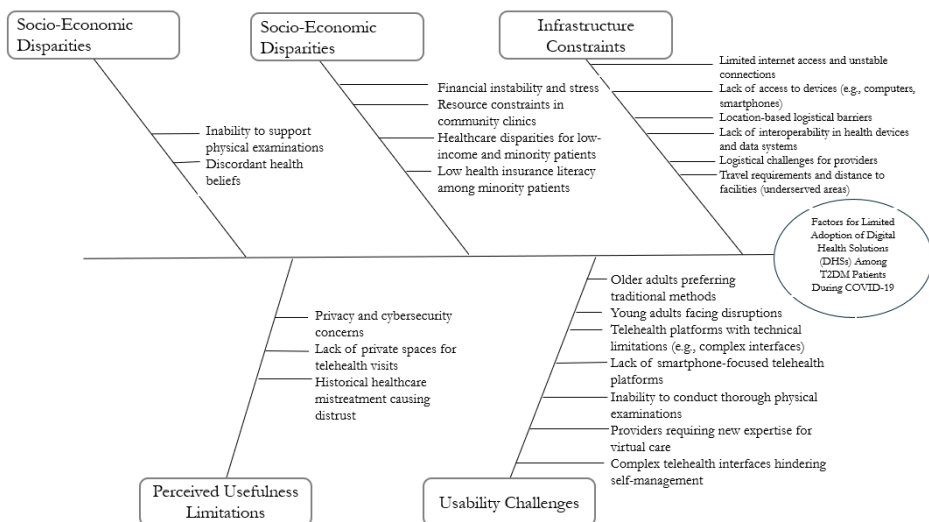


Figure 2: Fishbone Diagram of Barriers to DHS Adoption

4 Discussion

This review identifies infrastructure constraints, usability challenges, socio-economic disparities, privacy concerns, and perceived usefulness limitations as barriers to DHS adoption for T2DM management in the U.S. during COVID-19, analyzed through UTAUT. These U.S.-specific insights highlight policy impacts, addressing stakeholder adoption challenges.

Unreliable rural internet reduced FC, lowering BI (He et al., 2023). UTAUT assumes adequate infrastructure, but U.S. policy gaps, like limited broadband subsidies, exacerbated barriers, necessitating expanded broadband access. Complex interfaces increased EE, reducing BI among seniors, with 36.7% preferring landlines (Casas et al., 2023). Simplified, voice-navigated interfaces can enhance UB. Privacy fears diminished SI, as patients distrusted telehealth platforms. Robust encryption can rebuild trust (Ali et al., 2022). Telehealth's inability to support physical examinations reduced PE, as patients perceived DHSs as less effective, lowering BI (Casas et al., 2023). Demonstrating improved self-management outcomes can address this. Racial disparities, with African Americans facing higher costs due to insurance gaps, further limited FC, affecting 60% of low-income patients (Steinhardt et al., 2021; Tewari et al., 2023). Policies like the Affordable Care Act could ensure equity. Incompatible

systems hindered FC, reducing UB interoperable platforms can streamline care. Recommendations include investing in broadband, simplifying interfaces, enhancing security, promoting interoperability, subsidizing access, and highlighting benefits to boost adoption, aligning with UTAUT to improve BI and UB for T2DM management.

Future studies should test interventions like simplified interfaces or digital literacy programs to validate their impact on adoption. Longitudinal research can assess post-pandemic DHS effectiveness, while cross-regional studies may reveal global disparities in adoption patterns, building on UTAUT's application in diverse contexts.

This study has several limitations. The focus on literature from 2020 to 2023 limits the scope, potentially overlooking relevant studies conducted prior to the COVID-19 pandemic that might provide valuable context. Geographically, the study concentrates on the U.S., which may restrict the applicability of results to countries with different healthcare infrastructures and socio-economic dynamics. The study's reliance on secondary data from published literature, rather than primary data collection, also limits the ability to capture the real-time, nuanced experiences of patients and providers during the pandemic. Inclusion of potential non-empirical studies and lack of formal quality/bias assessment may limit rigor. Future research that incorporates a wider range of datasets, frameworks, and demographic variables could offer a more nuanced, globally relevant understanding of DHS adoption barriers for T2DM care.

5 Conclusion

This systematic review examined the barriers influencing the adoption of digital health interventions among individuals with T2DM during the COVID-19 pandemic, using the Unified Theory of Acceptance and Use of Technology (UTAUT) as a guiding framework. The findings highlight that PE, EE, FC, and SI significantly shaped users' engagement with digital tools, with pandemic-related disruptions amplifying these effects. This study contributes to the digital health literature by synthesizing recent U.S.-based evidence through a theory-driven lens, providing a timely understanding of how contextual and technological factors intersect during public health emergencies. The integration of UTAUT enhances

explanatory power and offers a structured basis for evaluating adoption barriers across different contexts. Moreover, rapid changes in digital health technology may affect the generalizability of these findings over time. Future research should explore the evolving role of digital literacy, personalization, and trust in post-pandemic health technology use, particularly among vulnerable or underserved populations. Expanding the theoretical scope beyond UTAUT may also help uncover deeper sociocultural and behavioral determinants that shape digital health engagement.

References

- Ali, A., Almaiah, M. A., Hajjej, F., Pasha, M. F., Fang, O. H., Khan, R., Teo, J., & Zakarya, M. (2022). An Industrial IoT-Based Blockchain-Enabled Secure Searchable Encryption Approach for Healthcare Systems Using Neural Network. *Sensors (Basel, Switzerland)*, 22(2), 572. <https://doi.org/10.3390/s22020572>
- Beks, H., King, O., Clapham, R., Alston, L., Glenister, K., McKinstry, C., Quilliam, C., Wellwood, I., Williams, C., & Wong Shee, A. (2022). Community Health Programs Delivered Through Information and Communications Technology in High-Income Countries: Scoping Review. *Journal of medical Internet research*, 24(3), e26515-e26515. <https://doi.org/10.2196/26515>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://url.au.m.mimecastprotect.com/s/VmMnCXLK3oUvE1JwtksxHWVmHz?domain=doi.org>
- Casas, L. A., Alarcón, J., Urbano, A., Peña-Zárate, E. E., Sangiovanni, S., Libreros-Peña, L., & Escobar, M. F. (2023). Telemedicine for the management of diabetic patients in a high-complexity Latin American hospital. *BMC health services research*, 23(1), 314-314. <https://doi.org/10.1186/s12913-023-09267-0>
- Chunara, R., Zhao, Y., Chen, J., Lawrence, K., Testa, P. A., Nov, O., & Mann, D. M. (2021). Telemedicine and healthcare disparities: a cohort study in a large healthcare system in New York City during COVID-19. *Journal of the American Medical Informatics Association : JAMIA*, 28(1), 33-41. <https://doi.org/10.1093/jamia/ocaa217>
- Cimini, C. C. R., Maia, J. X., Pires, M. C., Ribeiro, L. B., de Oliveira e Almeida Pinto, V. S., Batchelor, J., Ribeiro, A. L. P., & Marcolino, M. S. (2022). Pandemic-Related Impairment in the Monitoring of Patients With Hypertension and Diabetes and the Development of a Digital Solution for the Community Health Worker: Quasiexperimental and Implementation Study [Article]. *JMIR medical informatics*, 10(3), Article e35216. <https://doi.org/10.2196/35216>
- Filip, R., Gheorghita Puscaselu, R., Anchidin-Norocel, L., Dimian, M., & Savage, W. K. (2022). Global Challenges to Public Health Care Systems during the COVID-19 Pandemic: A Review of Pandemic Measures and Problems. *Journal of personalized medicine*, 12(8), 1295. <https://doi.org/10.3390/jpm12081295>
- He, Q., Keith, D., Eckhoff, D. O., Park, C., & Ng, B. P. (2023). Accessibility and Utilization of Telehealth Services During the COVID-19 Pandemic Among Medicare Beneficiaries by Diabetes Status. *Research in gerontological nursing*, 16(3), 134-146. <https://doi.org/10.3928/19404921-20230301-01>
- Kerr, D., & Sabharwal, A. (2021). Principles for virtual health care to deliver real equity in diabetes. *Lancet Diabetes Endocrinol*, 9(8), 480-482. [https://doi.org/10.1016/s2213-8587\(21\)00176-5](https://doi.org/10.1016/s2213-8587(21)00176-5)
- Lam, K., Lu, A. D., Shi, Y., & Covinsky, K. E. (2020). Assessing Telemedicine Unreadiness Among Older Adults in the United States During the COVID-19 Pandemic. *JAMA internal medicine*, 180(10), 1389-1391. <https://doi.org/10.1001/jamainternmed.2020.2671>

- Merrill, C. B., Roe, J. M., Seely, K. D., & Brooks, B. (2022). Advanced Telemedicine Training and Clinical Outcomes in Type II Diabetes: A Pilot Study. *Telemedicine reports*, 3(1), 15-23. <https://doi.org/10.1089/tmr.2021.0039>
- Monaghan, M., & Marks, B. (2020). Personal Experiences With COVID-19 and Diabetes Technology: All for Technology Yet Not Technology for All. *Journal of diabetes science and technology*, 14(4), 762-763. <https://doi.org/10.1177/1932296820930005>
- Patel, S. Y., Mehrotra, A., Huskamp, H. A., Uscher-Pines, L., Ganguli, I., & Barnett, M. L. (2021). Trends in Outpatient Care Delivery and Telemedicine During the COVID-19 Pandemic in the US. *JAMA internal medicine*, 181(3), 388-391. <https://doi.org/10.1001/jamainternmed.2020.5928>
- Petersen, F., Jacobs, M., & Pather, S. (2020, 2020//). Barriers for User Acceptance of Mobile Health Applications for Diabetic Patients: Applying the UTAUT Model. *Responsible Design, Implementation and Use of Information and Communication Technology*, Cham.
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & Group, P.-S. (2021). PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Journal of the Medical Library Association*, 109(2), 174-200. <https://doi.org/10.5195/jmla.2021.962>
- Stange, K. C. M. D. P. (2009). The Problem of Fragmentation and the Need for Integrative Solutions. *Annals of family medicine*, 7(2), 100-103. <https://doi.org/10.1370/afm.971>
- Steinhardt, M. A., Brown, S. A., Lehrer, H. M., Dubois, S. K., Wright, J. I., Whyne, E. Z., Sumlin, L. L., Harrison, L., Jr., & Woo, J. (2021). Diabetes Self-Management Education and Support Culturally Tailored for African Americans: COVID-19-Related Factors Influencing Restart of the TX STRIDE Study. *Sci Diabetes Self Manag Care*, 47(4), 290-301. <https://doi.org/10.1177/26350106211027956>
- Tewari, S., Coyne, K. D., Weinerman, R. S., Findley, J., Kim, S. T., & Flyckt, R. L. (2023). Racial Disparities in Telehealth Use During the COVID-19 Pandemic. *Fertility and sterility*. <https://doi.org/10.1016/j.fertnstert.2023.05.159>
- Upsher, R., Noori, Y., Kuriakose, L., Vassiliadou, I., Winkley, K., & Ismail, K. (2022). Needs, concerns and self-management experiences of people with type 2 diabetes during the COVID-19 pandemic: A qualitative study. *Diabetic medicine*, 39(8), e14883-n/a. <https://doi.org/10.1111/dme.14883>
- Vaughan, E., Foreyt, J., Virani, S., Balasubramanyam, A., Ballantyne, C., Amspoker, A., & Landrum, J. (2022, 2022). Mentored implementation to initiate a diabetes program in an underserved community: A pilot study.
- 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. <https://doi.org/10.2307/30036540>
- Villagra, V. G., Bhuva, B., Coman, E., Smith, D. O., & Fifield, J. (2019). Health insurance literacy: disparities by race, ethnicity, and language preference. *The American journal of managed care*, 25(3), e71-e75.
- Waizinger, O., Shpigelman, M., Shental, R., Yunis, B., Shimoni, P., Od Cohen, Y., & Kagan, I. (2022). Diabetes nurse practitioners in the shadow of the COVID-19 pandemic: Challenges, insights, and suggestions for improvement. *Journal of nursing scholarship : an official publication of Sigma Theta Tau International Honor Society of Nursing*, 54(4), 453-461. <https://doi.org/https://dx.doi.org/10.1111/jnu.12754>

Appendix A

Unified Theory of Acceptance and Use of Technology (UTAUT)

Overview: The Unified Theory of Acceptance and Use of Technology (UTAUT) explains user intentions to use a particular technology and subsequent usage behavior. It integrates elements from multiple technology acceptance models into one unified framework.

Origin: Proposed by Venkatesh et al. (2003).

Key Constructs:

- Performance Expectancy (PE): The degree to which an individual believes using the system will help achieve gains in job performance.
- Effort Expectancy (EE): The degree of ease associated with the use of the system.
- Social Influence (SI): The extent to which an individual perceives that others believe they should use the new system.
- Facilitating Conditions (FC): The degree to which an individual believes that an organizational and technical infrastructure exists to support system use.
- Behavioral Intention (BI): The intention to use the system, which influences actual usage (Use Behavior).
- Moderators: Age, Gender, Experience, and Voluntariness of Use can moderate the relationships between these constructs.

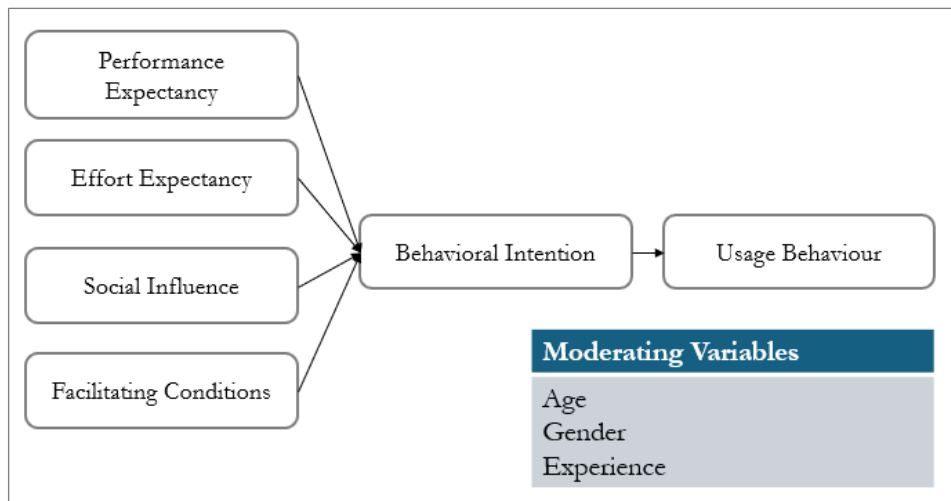


Figure 3: UTAUT Framework

AI'S ROLE IN COLLABORATIVE LEARNING SETTINGS IN HIGHER EDUCATION – A SYSTEMATIC REVIEW

HENNA PERNU, SANNA PAANANEN, SATU AKSOVAARA,
MINNA SILVENNOINEN

Jamk University of Applied Sciences, School of Professional Teacher Education,
Jyväskylä, Finland
henna.pernu@jamk.fi, sanna.paananen@jamk.fi, satu.aksovaara@jamk.fi,
minna.silvennoinen@jamk.fi

There is increasing demand for teamwork and collaboration proficiencies, which require recalibration of workforce capabilities. As artificial intelligence reshapes societal and professional demands, higher education must adapt by equipping students with essential skills for evolving workplace needs. This systematic review aims to map the perspectives explored and the observed effects of using AI in collaborative learning among higher education students. The reviewed 34 studies primarily explored how AI enhances collaborative learning, targeting group performance, collaboration, knowledge building, and social interactions. Findings showed that integrating AI into collaborative learning situations can promote both collaborative and individual learning in various higher education contexts. This review provides an overview of AI's role in collaborative learning settings, highlights current development, and identifies gaps for further research. It also serves as a source for educators, policymakers, and researchers interested in leveraging AI to foster enriching educational experiences.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.21](https://doi.org/10.18690/um.fov.4.2025.21)

ISBN
978-961-286-998-4

Keywords:
artificial intelligence,
collaborative learning,
higher education,
student,
knowledge building



University of Maribor Press

1 Introduction

Artificial intelligence (AI) has emerged as a powerful and rapidly advancing technology impacting our society (Vinothkumar & Karunamurthy, 2023), and in higher education (HE), it enhances human capabilities and opens new opportunities for teaching and learning (Popenici & Kerr, 2017). AI is commonly described as the capacity of computing systems to learn, adapt, and perform tasks that typically require human intelligence (Duan et al., 2019). In educational contexts, AI is further understood as the simulation of human-like cognitive processes such as data synthesis, adaptation, and error correction (Popenici & Kerr, 2017). As AI reshapes industries, professionals must acquire new competencies. There is a growing demand for higher-order cognitive skills as routine tasks are automated, necessitating a recalibration of the workforce's capabilities. The increasing complexity of tasks requires a diverse set of skills, such as problem-solving skills, and multidisciplinary knowledge (Katsamakas et al., 2024). This shift underlines the growing importance of 'human skills and competencies'—such as creativity, critical thinking, and emotional intelligence, and the ability to work and communicate effectively in teams is more critical than ever (Huang & Rust, 2018; Bae & Bozkurt, 2024).

While digital transformation presents both challenges and opportunities for HE (Rasul et al., 2024; Ritter et al., 2024), it is crucial that educational systems not only respond to these changes but also proactively equip students with the skills demanded by modern workplaces (Rêgo et al., 2024; Yang, 2023). Therefore, learning should shift from individual mastery to collaborative knowledge building that fosters essential skills like problem-solving and communication (Atchley et al., 2024; D'Mello et al., 2024). While AI reshapes the modern workplace, HE institutions (HEIs) must ensure students develop AI-complementary skills, those that support effective interaction with AI, informed decision-making, and navigation of complex social dynamics in professional settings (Katsamakas et al., 2024; Ritter et al., 2024; Poláková et al., 2023; Korteling et al., 2021). For HEIs to thrive in today's world, it is highly relevant to engage students in collaborative learning activities (Atchley et al., 2024).

Collaborative learning is grounded in social constructivism, which asserts that knowledge is constructed through collaborative interaction and shared understanding among individuals (Barkley et al., 2014; Dillenbourg, 1999), and it

may include sharing knowledge, assisting peers, and resolving viewpoints (Webb & Mastergeorge, 2003). Collaborative learning has been an integral part of HE for many decades. It was initiated by educators in the 1960s and 1970s with the aim of transforming the structure of authority within education (Yang, 2023). In HE, collaborative learning enhances student motivation, engagement, and academic success (Joseph et al., 2024; La Rocca, 2014; Loes, 2022; Webb & Mastergeorge, 2003) and it is also known to foster critical 21st-century workplace skills (Atchley et al., 2024).

The increased opportunities of AI in learning have led to its adoption in monitoring, assisting, and supporting collaborative situations (Järvelä et al., 2023; Atchley et al., 2024). Innovative digital practices are needed as HE increasingly faces situations where instructors cannot be present and there is a lack of in-person interactions (e.g., Otto et al. 2024). Recent research indicates that in the HE context, AI-based tools may effectively support collaboration by assisting both educators and students (Atchley et al., 2024). For example, generative AI tools can also support collaborative learning experiences through inquiry-based learning, critical thinking, and group communication, providing personalized experiences while fostering essential social and cognitive skills (Bae & Bozkurt, 2024).

Integrating AI into collaborative learning offers great potential but also raises concerns about bias, accuracy, ethics, and long-term educational impact (Monteith et al., 2024; Adeshola & Adepoju, 2024). To mitigate these challenges, AI should function as an assistive rather than a directive tool, fostering collaborative problem-solving and supporting rather than replacing human intellectual engagement (Atchley et al., 2024; Rêgo et al., 2024). In the systematic review, Chu et al. (2022) recognized that AI research in HE primarily focuses on student learning behaviour, as well as the accuracy and sensitivity of AI tools, with limited emphasis on AI's role in fostering higher-order thinking and collaboration. Due to the swift emergence of AI in HE, a literature review is needed to synthesize existing research and provide an evidence-based foundation for understanding its potential in supporting collaboration and collaborative learning. A systematic assessment of the current state of research aims to highlight impacts, identify gaps, and raise relevant research questions. It also aims to offer knowledge to support, for example, educators and education designers in making informed decisions about integrating AI tools in

collaborative learning settings in HE. This systematic review will address the following research questions:

- RQ1: What thematic categories can be identified in research on the use of artificial intelligence in the context of collaborative learning among higher education students?
- RQ2: What effects of students' use of artificial intelligence on collaborative learning are identified in research conducted in higher education settings?

2 Method

2.1 Search strategy and identification of relevant publications

This study adopted the PRISMA framework for tracking and reporting the selection of studies (Moher et al., 2009). We specifically included research articles that discussed collaborative learning and the use of AI applications by students in HE. We conducted two searches. The initial search was conducted by both authors on February 3, 2024, and the filtering process resulted in a reduced set of records for further screening. Following this, the backward snowballing method (Wohlin, 2014) was applied in alphabetical order of the included records by the second author using the same inclusion and exclusion criteria. The second search was conducted on October 30, 2024, to update the review with the latest articles. The overall search and screening process is summarized in Figure 1.

The selection process for relevant articles was conducted similarly after both searches as follows. To ensure a sufficient level of interrater reliability during the screening process, we implemented a multi-step procedure. Before the abstract screening began, we conducted a calibration phase in which both authors independently screened a shared sample of abstracts ($n = 26$) to align the inclusion and exclusion criteria into practice. Discrepancies were discussed jointly to reach a common understanding and refine the screening guidelines. After this calibration, the remaining abstracts were divided evenly between the two primary authors, with each screening being approximately half. During this phase, any potentially unclear or borderline cases were flagged and discussed collaboratively. If consensus could not be reached between the two, a fourth author acted as an adjudicator and made

the final inclusion decision. This process helped to ensure consistency in decision-making and minimize the risk of bias or misinterpretation.

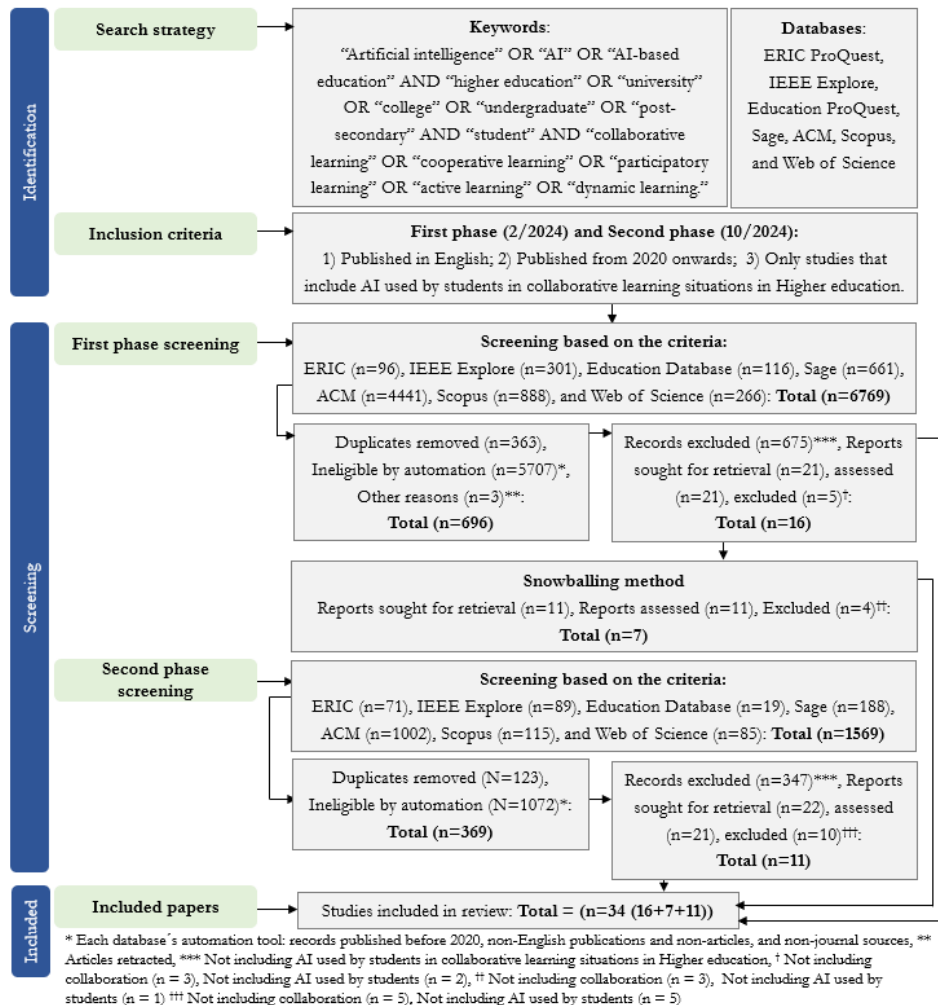


Figure 1: Flowchart of the selection process

2.3 Analysis

The analysis of the included articles followed a systematic and collaborative methodology, in line with PRISMA guidelines. Initially, the 34 records that met the inclusion criteria were independently reviewed by the first and second authors.

Relevant information from each article, including article details such as title, authors, publication year, country of origin, study design, journal ranking level, publication venue, and impact factor, was recorded in a structured Excel file to facilitate analysis and ensure consistency. A data-driven inductive content analysis method was employed to analyze the data extracted from the results of the selected articles. This approach followed the principles of qualitative content analysis as described by Elo and Kyngäs (2008), emphasizing systematic data reduction, category development, and transparent documentation of the analytical process. To systematically organize and interpret the data, key elements such as the original expression, condensed expression, keyword or phrase, subcategory, and main category were recorded in a structured Excel file. During the summarization process, the data collected was organized into initial subcategories, which were formed based on the observed recurring elements and similarities. These subcategories served as the basis for the creation of broader categories, which were built by combining thematically similar subcategories. The formation of categories was a critical step in the analysis, as it allowed for the structuring of the data and the presentation of key findings systematically and coherently. To facilitate this process, the main categories and subcategories were visually mapped on the MIRO board, enabling a clearer overview of the relationships between the key concepts and supporting the structured synthesis of the findings. The entire content analysis process was carried out in close collaboration among the research team, and each step of the process was carefully documented, ensuring transparency of the methods and the reproducibility of the research. The final categories are presented in the results section 3.1 of the article.

2.4 Quality appraisal

The quality of the articles included in the review was assessed using the Mixed Method Appraisal Tool (MMAT), which is a validated tool designed for the appraisal of empirical studies employing qualitative, quantitative, and mixed methods approaches. We used screening questions tailored to these research designs (Hong et al. 2018). The first and second authors conducted independent evaluations of the studies utilizing the MMAT scoring system. The final evaluations revealed that the studies included achieved scores ranging from 40% to 100% on the MMAT. Furthermore, the impact factors of the journals were screened. (See Table 1.)

Table 1: Quality assessment (MMAT score%, journal, IF)

MMAT Quality Score (%)	Studies	Total	Impact Factor (IF) [0,10]	Studies	Total
100	8	34	6-10	6	34
80	14		3-6	18	
60	10		1-3	8	
40	2		Unknown	2	

2.5 Included publications

A diverse range of research methodologies was employed across the identified studies. Specifically, four studies employed qualitative approaches. Among the quantitative studies, seven were randomized controlled trials, four were nonrandomized studies, and one utilized a descriptive quantitative approach. Additionally, 18 studies utilized mixed methods. Of these, eight combined qualitative and quantitative descriptive methods, while ten integrated qualitative approaches with non-randomized quantitative methods. Most of the studies (23) included in the review were published within the years 2023-2024. During the years 2020-2022, there were 10 studies published, and one article was published in 2025. The studies included in this research were collected from a range of 18 different countries. A significant portion of these studies originated from institutions in China (15), followed by the USA, Serbia, and Saudi Arabia, each with two studies. The remaining countries each contributed a single study.

3 Results

3.1 Thematic categories on the use of artificial intelligence in collaborative learning among higher education students (RQ1)

The analysis of the selected literature of 34 studies revealed five main thematic categories concerning the use of AI in collaborative learning among HE students. From these main categories, altogether 20 subcategories were identified (See Table 2).

Table 2: Thematic categories of the use of AI in collaborative learning situations

Main categories and frequency	Subcategories and frequency	References
Collaborative learning (28)	Group performance (8)	Hayashi (2020); Kumar (2021); Long, et al. (2024); Zheng, et al. (2022); (2023a); (2023b); (2023c); (2024)
	Collaboration (7)	Hayashi (2020); Ilic et al. (2021); Järvelä et al. (2023); Kuleto et al. (2021); Ouyang et al. (2023); Tegos et al. (2021); Van Horn (2024)
	Collaborative knowledge building (6)	Chen et al. (2023); Saadati et al. (2023); Zheng et al. (2023a); (2023b); (2023c); (2024)
	Social interaction (5)	Li X. et al. (2024); Ouyang et al. (2024); Zheng et al. (2021); (2022); (2023a)
	Socially shared regulation (2)	Zheng et al. (2023a); (2023c)
Learner characteristics (26)	Cognition (9)	Li T. et al. (2024); Li X. et al. (2024); Kong et al. (2025); Kumar (2021); Peng et al. (2022); Zheng et al. (2021); (2023a); (2023c); (2024)
	Engagement (6)	Kumar (2021); Ouyang et al., (2024); Peng et al. (2022), Ramos & Condotta (2024); Van Horn (2024); Xie et al. (2021)
	Metacognition (4)	Li T. et al. (2024); Malik (2024); Ouyang et al. (2023); Van Horn (2024)
	Reflection (3)	Chen et al. (2023); Lin et al. (2024); Ouyang et al. (2024)
	Self-efficacy (3)	Kong et al. (2025); Kumar (2021); Van Horn (2024)
	Demographics (1)	Joseph et al. (2024)
Learning process (19)	Learning performance (8)	Ilic et al. (2021); Kong et al. (2025); Kuleto et al. (2021); Kumar (2021); Peng et al. (2022); Ramos & Condotta (2024); Tegos et al. (2021); Li T. et al. (2024)
	Feedback (5)	Kumar (2021); Olga et al. (2023); Ouyang et al. (2023); Saadati et al. (2023); Van Horn (2024)
	Pedagogical support (3)	Chen et al. (2023); Ouyang et al. (2024); Ramos et al. (2024)
	Self-directed learning (3)	Järvelä et al. (2023); Richter et al. (2024); Van Horn (2024)
Use of AI tools in collaboration (12)	Skills and knowledge (5)	Almulla (2024); Ilieva et al. (2023); Richter et al. (2024); Tzirides et al. (2024); Ilic et al. (2021)
	Emotional response (4)	Ilieva et al. (2023); Li X. et al. (2024); Rosenberg-Kima et al. (2020); Peng et al. (2022)

Main categories and frequency	Subcategories and frequency	References
	Perception and attitude (3)	Joseph et al. (2024); Long et al. (2024); Tegos et al. (2021)
Learning outcomes (8)	Competency (5)	Li X. et al. (2024); Lin et al. (2024); Long et al. (2024); Qureshi (2023); Ouyang et al. (2024)
	Knowledge integration (3)	Hayashi (2020); Zheng et al. (2021); (2022)

3.2 Effects of students' use of AI on collaborative learning identified in research conducted in HE settings

To address the second research question, a more detailed overview of the studies previously summarised in the table is presented within each thematic category.

Collaborative learning: A synthesis of the reviewed studies highlights that the use of AI in collaborative learning situations can affect *group performance* (Hayashi, 2020; Zheng et al., 2021; 2023a; 2024) by enhancing teamwork skills (Kumar, 2021; Long et al., 2024; Zheng et al., 2023b) and teamwork competencies (Long et al., 2024). Furthermore, AI integration fosters *collaboration* among students (Järvelä et al., 2023; Tegos et al., 2021; Van Horn, 2024; Zheng et al., 2022), for example, by positively influencing mutual learning and knowledge exchange (Van Horn, 2024), as well as in fostering students' regulation-oriented feedback (Ouyang et al., 2024) and improving coordination (Hyashi, 2020). The synthesis also indicates that AI can facilitate collaborative learning environments for HE institutions (Ilic et al., 2021; Kuleto et al., 2021). Moreover, results showed that AI-based systems can enhance *collaborative knowledge building* (Zheng et al., 2023a; 2023b; 2024) through AI-generated feedback (Zheng et al., 2023c), support for constructing arguments (Chen et al., 2023), and by encouraging co-construction of knowledge and problem-solving (Saadati et al., 2023). According to the studies reviewed, AI tools support *social interaction* through enhanced social engagement (Zheng et al., 2023a), foster interactive relationships (Zheng et al., 2021; 2022), facilitate information sharing (Li X. et al., 2024), and improve peer interactions (Ouyang et al., 2024). Additionally, the use of AI-based tools in collaborative learning situations enhances *social shared regulation* among learners (Zheng et al., 2023a; 2023c).

Learner characteristics: The synthesis of reviewed studies indicates that AI can enhance cognition by promoting engagement (Zheng et al., 2024), deeper cognitive thinking (Peng et al., 2022), and cognitive presence (Kong et al., 2025). While AI-supported collaboration has been associated with reduced cognitive load in some cases (Li et al., 2024), other studies have found no significant impact (Kumar, 2021; Zheng et al., 2021, 2023a, 2023c). Furthermore, AI-supported learning was found to enhance student *engagement* in some studies (Ouyang et al., 2024; Peng et al., 2022; Ramos & Condotta, 2024; Van Horn, 2024; Xie et al., 2021). However, Kumar et al. (2021) found no significant influence on students' perception of learning. Moreover, studies also identified impacts on *metacognition*, as the use of AI tools in learning environments enhanced students' understanding by supporting reflective and critical thinking (Li T. et al., 2024). AI-supported learning was also found to enhance metacognitive skills and the development of metacognition (Malik, 2024; Van Horn, 2024), as well as increase metacognitive engagement (Ouyang et al., 2023). Studies also showed that AI tools use positively influenced students' *reflection*, as AI-assisted learning facilitated idea refinement, reflective thinking, and self-awareness through collaborative argumentation, structured observation, real-time feedback, and behaviour adjustment (Chen et al., 2023; Lin et al., 2024; Ouyang et al., 2024). While many studies reported positive outcomes, the effects on students' self-efficacy were mixed. Increased confidence was observed primarily in the specific skills targeted by the learning activities (Van Horn, 2024), but no significant effects were found on motivational beliefs or creative self-efficacy (Kumar, 2021), nor was there evidence of increased motivation (Kong et al., 2025). Moreover, Joseph et al. (2024) found that *demographic* factors such as gender, field of study, specialization, and place had no effect on AI tool use, digital literacy competence, or participation in AI-assisted learning.

Learning process: AI in collaborative learning demonstrated a positive impact on *learning performance* by enhancing learning retention and academic performance, as well as supporting dynamic problem-based learning activities (Kong et al., 2025; Kumar, 2021; Li, T. et al., 2024; Peng et al., 2022; Ramos et al., 2024). In addition, AI was found to support the development of customized learning skills and enhance personalized learning (Ilic et al., 2021; Kuleto et al., 2021). The importance of real-time *feedback* provided by AI was detected in several studies (e.g. Chen et al., 2023; Li X. et al., 2024; Ouyang et al., 2024; Van Horn, 2024; Zheng et al., 2022), for example, in fostering problem-solving approaches (Ouyang et al., 2024). Despite

occasional confusion caused by AI-generated feedback (Ouyang et al., 2023), students generally viewed it positively (Kumar, 2021; Ouyang et al., 2024). A combination of peer and AI feedback was considered the most effective (Olga et al., 2023). AI tools have also been successfully utilized as *pedagogical support* for monitoring and managing the learning process (Chen et al., 2023), and they have proven effective in enabling more active teaching by supporting instructional planning (Ramos et al., 2024). Furthermore, AI-driven learning was found to reduce the need for extensive instructor feedback and support (Ouyang et al., 2024). Finally, AI-enhanced learning was found to support *self-directed learning*, contributing to more personalized education and fostering responsible decision-making through experimental and ethical practices (Van Horn, 2024; Richter et al., 2024). It also improved self-regulation by providing real-time adaptive feedback in human-AI collaboration (Järvelä et al., 2023).

Use of AI tools in collaboration: According to the synthesis of the reviewed studies, the use of AI in interactive and collaborative learning settings supports students' AI-related *skills and knowledge* by enhancing their perception of its ease of use and benefits, while also fostering skill development (Almulla, 2024; Ilic et al., 2021). Combining AI with human intelligence was found to improve AI literacy and boost students' confidence in using AI technologies (Joseph et al., 2024; Tzirides et al., 2024). Additionally, Richter et al. (2024) emphasize the importance of collaborative and experimental learning methods for ethical and practical use of AI. Students' *emotional responses* to AI in learning were generally positive (Ilieva et al., 2023; Li X. et al., 2024; Peng et al., 2022) as were their *perceptions and attitudes* towards AI integration (Long et al., 2024; Rosenberg-Kima et al., 2020; Tegos et al., 2021).

Learning outcomes: AI integration in collaborative learning has been shown to foster understanding and *competency* development, e.g., by providing essential support for structuring inquiry-based tasks, enhancing collaborative academic writing by improving writing quality, fostering reflective thinking and active participation in discussions, and enhancing problem-solving behaviours and learning outcomes, although final performance improvements may vary (Long et al., 2024; Li et al., 2024; Lin et al., 2024; Qureshi, 2023; Ouyang et al., 2023). Furthermore, AI was found to promote deeper knowledge processing by fostering both knowledge convergence and elaboration through real-time feedback and automated interaction classification, thereby contributing to *knowledge integration* (Zheng et al., 2022; Hayashi, 2020).

4 Discussion and conclusions

The quality and credibility of the analysed studies appeared relatively high, with the majority demonstrating strong methodological quality. As a response to the **first research question**, the analysis of the selected studies revealed five main categories and 20 subcategories. In addressing **the second research question**, the findings suggest that AI integration in collaborative learning holds potential to support both group and individual learning. Most reported effects were positive, including enhanced teamwork skills, improved group performance, and strengthened collaboration and knowledge building (e.g., Hayashi, 2020; Kumar, 2021; Tegos et al., 2021; Saadati et al., 2023). In summary, AI was found in supporting learning by providing adaptive and real-time feedback, enhancing engagement, assisting in the development of metacognitive skills, and facilitating the management of cognitive load (e.g., Chen et al., 2023; Li et al., 2024; Malik, 2024; Van Horn, 2024; Zheng et al., 2023a, 2023c). **The limitations of this review** include its focus on peer-reviewed journal articles and English-language publications, which may have restricted the diversity of perspectives and limited the international scope of the findings.

Future research should examine AI's long-term impact on teaching and learning, particularly as a collaborative and self-directed learning companion, while expanding to diverse educational settings across academic levels, disciplines, and cultural contexts. Although most reviewed studies focused on the benefits of AI, critical perspectives were limited. As AI tools use expands, concerns about bias, inaccuracy, and ethics emerge (Adeshola & Adepoju, 2024; Monteith et al., 2024), highlighting the need to examine the ethical and responsible use of AI in collaborative learning, in order to evaluate potential drawbacks.

In conclusion, as AI, particularly Generative AI, continues to reshape work and education, the importance of human skills, such as communication, collaboration, and problem-solving, is growing. In this context, AI also offers new opportunities to support collaborative learning. These developments call for HEIs to rethink their role and pedagogy (see also Popenici & Kerr, 2017). AI should be integrated in various ways into learning environments and competency development in diverse ways, while ensuring that students' AI literacy is strengthened and that they can apply it both ethically and effectively (Richter et al., 2024; Tzirides et al., 2024). Maximizing

AI's educational benefits requires collaboration among educators, policymakers, researchers, and developers, along with a deeper understanding of its role across learning environments to refine best practices for its integration, ensuring that it meaningfully supports both individual and collaborative learning.

Acknowledgments

The first and second authors contributed equally to this work.

References

(*) Indicates review articles included in the synthesis

- Adeshola, I., & Adepoju, A. P. (2024). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, 32(10), 6159-6172.
- Almulla, M. A. (2024). Investigating influencing factors of learning satisfaction in AI ChatGPT for research: University students' perspective. *Heliyon*, 10(11). Scopus.
<https://doi.org/10.1016/j.heliyon.2024.e32220> *
- Atchley, P., Pannell, H., Wofford, K. et al. Human and AI collaboration in the higher education environment: opportunities and concerns. *Cogn. Research* 9, 20 (2024).
<https://doi.org/10.1186/s41235-024-00547-9>
- Bae, H., & Bozkurt, A. (2024). The Untold Story of Training Students with Generative AI: Are We Preparing Students for True Learning or Just Personalization? *Online Learning*, 28(3).
- Barkley, Elizabeth F., and K. Patricia Cross. (2014) *Collaborative Learning Techniques: A Handbook for College Faculty*, John Wiley & Sons, Incorporated.
- Chen, W., Tan, J. S., Zhang, S., Pi, Z., & Lyu, Q. (2023). AppleTree system for effective computer-supported collaborative argumentation: An exploratory study. *Educational technology research and development*, 71(5), 2103-2136.*
- Chu, H. C., Hwang, G. H., Tu, Y. F., & Yang, K. H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, 38(3), 22-42.
- Dillenbourg, P. (1999). What do you mean by collaborative learning?. *Collaborative-learning: Cognitive and computational approaches.*, 1-19.
- D'Mello, S. K., Biddy, Q., Breideband, T., Bush, J., Chang, M., Cortez, A., ... & Whitehill, J. (2024). From learning optimization to learner flourishing: Reimagining AI in Education at the Institute for Student-AI Teaming (iSAT). *AI Magazine*, 45(1), 61-68.
- Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data—evolution, challenges and research agenda. *International journal of information management*, 48, 63-71.
- Elo S. & Kyngäs H. (2008) The qualitative content analysis process. *J Adv Nurs*.Apr;62(1):107-15.
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative ai. *Business & Information Systems Engineering*, 66(1), 111-126.
- Hayashi, Y. (2020). Gaze awareness and metacognitive suggestions by a pedagogical conversational agent: an experimental investigation on interventions to support collaborative learning process and performance. *International Journal of Computer-Supported Collaborative Learning*, 15(4), 469-498. *

- Hingle, A. (2025, January). How Technology Students Develop AI Literacy through Collaboration, Reflection, and Situated Learning. In Companion Proceedings of the 2025 ACM International Conference on Supporting Group Work (pp. 118-120).
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., ... & Vedel, I. (2018). Mixed methods appraisal tool (MMAT), version 2018. *Registration of copyright*, 1148552(10), 1-7.
- Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of service research*, 21(2), 155-172.
- Ilic, M. P., Păun, D., Popović Šević, N., Hadžić, A., & Jianu, A. (2021). Needs and performance analysis for changes in higher education and implementation of artificial intelligence, machine learning, and extended reality. *Education Sciences*, 11(10), 568. *
- Ilieva, G., Yankova, T., Klisarova-Belcheva, S., Dimitrov, A., Bratkov, M., & Angelov, D. (2023). Effects of generative chatbots in higher education. *Information*, 14(9), 492. *
- Joseph, G. V., Athira, P., Thomas, M. A., Jose, D., Roy, T. V., & Prasad, M. (2024). Impact of Digital Literacy, Use of AI Tools and Peer Collaboration on AI Assisted Learning: Perceptions of the University Students. *Digital Education Review*, 45, 43-49. *
- Järvelä, S., Nguyen, A., & Hadwin, A. (2023). Human and artificial intelligence collaboration for socially shared regulation in learning. *British Journal of Educational Technology*, 54(5), 1057–1076. *Education Database*. *
- Katsamakas, E., Pavlov, O. V., & Saklad, R. (2024). Artificial Intelligence and the Transformation of Higher Education Institutions: A Systems Approach. *Sustainability*, 16(14), 6118.
- Kong, X., Liu, Z., Chen, C., Liu, S., Xu, Z., & Tang, Q. (2025). Exploratory study of an AI-supported discussion representational tool for online collaborative learning in a Chinese university. *Internet and Higher Education*, 64. Scopus. Available online 10 October 2024 *
- Korteling, J. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human-versus artificial intelligence. *Frontiers in artificial intelligence*, 4, 622364.
- Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, 13(18), 10424. *
- Kumar, J. A. (2021). Educational Chatbots for Project-Based Learning: Investigating Learning Outcomes for a Team-Based Design Course. *International Journal of Educational Technology in Higher Education*, 18. ERIC. *
- La Rocca, C. (2014). Collaborative learning in higher education. *Open Journal of Social Sciences*, 2(02), 61.
- Li, T., Ji, Y., & Zhan, Z. (2024). Expert or machine? Comparing the effect of pairing student teacher with in-service teacher and ChatGPT on their critical thinking, learning performance, and cognitive load in an integrated-STEM course. *Asia Pacific Journal of Education*, 44(1), 45-60. *
- Li, X., Jiang, S., Hu, Y., Feng, X., Chen, W., & Ouyang, F. (2024). Investigating the impact of structured knowledge feedback on collaborative academic writing. *Education and Information Technologies*, 1-29. *
- Lin, C. J., Lee, H. Y., Wang, W. S., Huang, Y. M., & Wu, T. T. (2024). Enhancing reflective thinking in STEM education through experiential learning: The role of generative AI as a learning aid. *Education and Information Technologies*, 1-23. *
- Loes, C. N. (2022). The Effect of Collaborative Learning on Academic Motivation. *Teaching & Learning Inquiry*, 10. <https://doi.org/10.20343/teachlearningqu.10.4>
- Long, T., Zheng, Z., Shi, Y., Tong, M., & Liu, Z. (2024). Understanding how pre-service science teachers design inquiry-based activities in a knowledge integration (KI) based collaborative learning environment: a network analytic approach. *Educational technology research and development*, 1-32. *
- Malik, U. (2024). Developing Student Metacognition Through Collaborative Work on the Quip Platform. *International Journal of Educational Reform*, 10567879241247557. *

- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097.
- Monteith, S., Glenn, T., Geddes, J. R., Whybrow, P. C., Achtyes, E., & Bauer, M. (2024). Artificial intelligence and increasing misinformation. *The British Journal of Psychiatry*, 224(2), 33-35.
- Olga, A., Saini, A., Zapata, G., Searsmith, D., Cope, B., Kalantzis, M., ... & Kastania, N. P. (2023). Generative AI: Implications and applications for education. *arXiv preprint arXiv:2305.07605*. *
- Otto, S., Bertel, L. B., Lyngdorf, N. E. R., Markman, A. O., Andersen, T., & Ryberg, T. (2024). Emerging digital practices supporting student-centered learning environments in higher education: A review of literature and lessons learned from the COVID-19 pandemic. *Education and Information Technologies*, 29(2), 1673-1696.
- Ouyang, F., Guo, M., Zhang, N., Bai, X., & Jiao, P. (2024). Comparing the effects of instructor manual feedback and ChatGPT intelligent feedback on collaborative programming in China's higher education. *IEEE Transactions on Learning Technologies*. *
- Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education*, 20(1), 4. *
- Peng, Y., Li, Y., Su, Y., Chen, K., & Jiang, S. (2022). Effects of group awareness tools on students' engagement, performance, and perceptions in online collaborative writing: Intergroup information matters. *The Internet and Higher Education*, 53, 100845. *
- Poláková, M., Sulcimanová, J. H., Madžik, P., Copuš, L., Molnárová, I., & Polednová, J. (2023). Soft skills and their importance in the labour market under the conditions of Industry 5.0. *Heliyon*, 9(8).
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and practice in technology enhanced learning*, 12(1), 22.
- Qureshi, B. (2023). Exploring the use of chatgpt as a tool for learning and assessment in undergraduate computer science curriculum: Opportunities and challenges. *arXiv preprint arXiv:2304.11214*. *
- 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. *
- Rasul, T., Nair, S., Kalendra, D., Balaji, M. S., de Oliveira Santini, F., Ladeira, W. J., ... & Hossain, M. U. (2024). Enhancing academic integrity among students in GenAI Era: A holistic framework. *The International Journal of Management Education*, 22(3), 101041.
- Rêgo, B. S., Lourenço, D., Moreira, F., & Pereira, C. S. (2024). Digital transformation, skills and education: A systematic literature review. *Industry and higher education*, 38(4), 336-349.
- Richter, S., Giroux, M., Piven, I., Sima, H., & Dodd, P. (n.d.). A Constructivist Approach to Integrating AI in Marketing Education: Bridging Theory and Practice. *Journal of Marketing Education*, 0(0), 0273475324128876. *
- Ritter, B. A., Gopalkrishnan, S. S., & Sockbeson, C. E. S. (2024). The Next Generation of Pedagogy: Reflections from EdTech Leaders. *Journal of Management Education*, 48(4), 610-633.
- Rosenberg-Kima, R. B., Koren, Y., & Gordon, G. (2020). Robot-supported collaborative learning (RSCL): Social robots as teaching assistants for higher education small group facilitation. *Frontiers in Robotics and AI*, 6, 148.*
- Saadati, Z., Zeki, C. P., & Vatankhah Barenji, R. (2023). On the development of blockchain-based learning management system as a metacognitive tool to support self-regulation learning in online higher education. *Interactive Learning Environments*, 31(5), 3148-3171.*
- Tegos, S., Mavridis, A., & Demetriadis, S. (2021). Agent-supported peer collaboration in MOOCs. *Frontiers in Artificial Intelligence*, 4, 710856. *
- Tzirides, A. O. O., Zapata, G., Kastania, N. P., Saini, A. K., Castro, V., Ismael, S. A., ... &

- Kalantzis, M. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. *Computers and Education Open*, 6, 100184. *
- Van Horn, K. R. (2024). ChatGPT in English Language Learning: Exploring Perceptions and Promoting Autonomy in a University EFL Context. *TESL-EJ*, 28(1). Scopus. *
- Vinothkumar, J., & Karunamurthy, A. (2023). Recent Advancements in Artificial Intelligence Technology: Trends and Implications. *Quing: International Journal of Multidisciplinary Scientific Research and Development*, 2(1), 1–11.
- Webb, N. M., & Mastergeorge, A. (2003). Promoting effective helping behavior in peerdirected groups. *International Journal of Educational Research*, 39(1–2), 73–97.
- 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 (EASE '14)*, 1–10.
- Xie, T., Liu, R., Chen, Y., & Liu, G. (2021). MOCA: A motivational online conversational agent for improving student engagement in collaborative learning. *IEEE Transactions on Learning Technologies*, 14(5), 653–664.*
- Yang, X. (2023). A Historical Review of Collaborative Learning and Cooperative Learning. *TechTrends* 67, 718–728.
- Zheng, L., Zhong, L., Niu, J., Long, M., & Zhao, J. (2021). Effects of Personalized Intervention on Collaborative Knowledge Building, Group Performance, Socially Shared Metacognitive Regulation, and Cognitive Load in Computer-Supported Collaborative Learning. *Educational Technology & Society*, 24(3), 174–193. *
- Zheng, L., Niu, J., & Zhong, L. (2022). Effects of a learning analytics-based real-time feedback approach on knowledge elaboration, knowledge convergence, interactive relationships and group performance in CSCL. *British Journal of Educational Technology*, 53(1), 130–149. *
- Zheng, L., Niu, J., Long, M., & Fan, Y. (2023a). An automatic knowledge graph construction approach to promoting collaborative knowledge building, group performance, social interaction and socially shared regulation in CSCL. *British Journal of Educational Technology*, 54(3), 686–711. *
- Zheng, L., Fan, Y., Chen, B., Huang, Z., & Long, M. (2023b). An AI-enabled feedback-feedforward approach to promoting online collaborative learning. *Education and Information Technologies*. Scopus. *
- Zheng, L., Long, M., Niu, J., & Zhong, L. (2023c). An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL. *International Journal of Computer-Supported Collaborative Learning*, 18(1), 101–133. *
- Zheng, L., Fan, Y., Gao, L., Huang, Z., Chen, B., & Long, M. (2024). Using AI-empowered assessments and personalized recommendations to promote online collaborative learning performance. *Journal of Research on Technology in Education*. Scopus. *

“RAGE AGAINST THE MACHINE?”: THE IMPACT OF CLINICAL DECISION SUPPORT SYSTEMS ON HOSPITAL NURSING DECISION- MAKING, WORKFLOW EFFICIENCY, AND PATIENT OUTCOMES: A RAPID REVIEW

MATTHIJS BERKHOUT,¹ ASTRID VAN BARNEVELD,²
KOEN SMIT,¹ THIJS VAN HOUWELINGEN³

¹ HU University of Applied Sciences Utrecht, Research Group Digital Ethics, Utrecht, the Netherlands

Matthijs.berkhout@hu.nl, Koen.smit@hu.nl

² Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands

a.vanbarneveld@prinsesmaximacentrum.nl

³ HU University of Applied Sciences Utrecht, Research Group Technology for Healthcare Innovations, Utrecht, the Netherlands

thijs.vanhouwelingen@hu.nl

Clinical Decision Support Systems (CDSS) are increasingly developed for hospital nursing practice, yet their impact on decision-making, workflow efficiency, and patient outcomes remains complex. This rapid review synthesizes findings from 21 studies, highlighting both the benefits and challenges of CDSS implementation focused on three key areas. CDSS can enhance nursing decision-making by reducing variability and improving standardization, but there are concerns about system usability and the tendency to override recommendations. While CDSS improve workflow efficiency by prioritizing tasks, issues such as alert fatigue and poor interoperability with hospital systems hinder their potential. Patient outcomes benefit from CDSS-driven medication safety and risk prevention, yet adherence to recommendations varies among nurses. These findings underscore the need for user-centered CDSS that align with nursing values. Future research should explore long-term effectiveness, implementation strategies and best practices for integrating CDSS into nursing workflows.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.22](https://doi.org/10.18690/um.fov.4.2025.22)

ISBN

978-961-286-998-4

Keywords:

CDSS,
nurse,
data driven decision making,
healthcare,
workflow efficiency



University of Maribor Press

1 Introduction

In the evolving landscape of healthcare, Clinical Decision Support Systems (CDSS) have become essential tools designed to enhance clinical decision-making by providing healthcare professionals with evidence-based knowledge in real-time (Chen et al., 2023; Sutton et al., 2020). The integration of CDSS into healthcare settings has been associated with improved clinical practices and patient outcomes (Roshanov et al., 2013; Sutton et al., 2020; Taheri Moghadam et al., 2021). Typically embedded within Electronic Health Records (EHRs), CDSSs facilitate the access to patient data and clinical guidelines, allowing healthcare providers to make more informed decisions (Chen et al., 2023).

While CDSS have been broadly studied in the context of physician decision-making, their specific impact on nursing practice remains underexplored (Chen et al., 2023; Meunier et al., 2023). Nurses play a critical role in patient care, as a first contact to patients, engaging in (complex) decision-making processes that encompass medication administration, patient monitoring, and risk assessment (Kwon & Lee, 2024). Nurses also interact with CDSSs differently than physicians, often facing unique challenges that can affect adoption and effectiveness. Unlike physicians, whose primary focus is on diagnosing diseases and formulating treatment plans, nurses are responsible for executing these plans while continuously monitoring patient conditions (Deshpande, 2009). The dynamic nature of nursing workflows, particularly in high-acuity settings, adds further complexity to CDSS implementation. For instance, in intensive care units, nurses frequently override CDSS alerts because many systems rely on static algorithms that fail to account for rapidly evolving patient conditions (Stevens et al., 2024).

CDSS have the potential to support nursing decision-making by improving efficiency and patient safety, however, their implementation presents several challenges. Nurses encounter issues such as alert fatigue, workflow disruptions, and usability concerns, which can hinder the effective utilization of these systems (Shah et al., 2022). A systematic review analyzing CDSS use in intensive care units found that while these systems show promise in enhancing care quality, their integration into nursing workflows remains difficult due to poor usability and misalignment with real-world clinical needs (Sarıköse & Şenol Çelik, 2024). Earlier studies show that CDSS hold promise for improving key aspects of nursing practice, particularly in

three critical areas identified by earlier studies: the decision-making process, workflow efficiency, and patient outcomes (Chen et al., 2023; Kwon & Lee, 2024; Sutton et al., 2020).

However, the integration of CDSS into nursing practice involves not only technical and procedural aspects, as defined by regulatory frameworks such as the Medical Device Regulation, but also socio-technical factors that influence everyday clinical use. Therefore, it is also connected to professional values. Patient-centered care, clinical autonomy, and ethical responsibility shape how nurses interact with decision-support technologies. A fundamental challenge in CDSS adoption is ensuring that these systems support, rather than undermine or take over, these core values. Transparency in how CDSSs generate recommendations is critical, as nurses must trust the system to make informed decisions without feeling that their professional judgment is being replaced (Elgin & Elgin, 2024; MacIntyre et al., 2023). Ethical considerations, such as bias in algorithmic recommendations and accountability for decision-making, further highlight the need for CDSS to align with nursing principles. Understanding how these values influence CDSS effectiveness is essential for evaluating its impact on hospital nurses' decision-making, workflow efficiency, and patient outcomes. This study does not evaluate CDSS compliance or certification but instead examines real-world outcomes and challenges in nursing use.

Therefore, an exploration into the role of CDSS in nursing practice is warranted to find better integration and utilization strategies. This study seeks to address the following research question: *How does the use of Clinical Decision Support Systems (CDSS) impact the decision-making processes, workflow efficiency, and patient outcomes in hospital nursing practice?*

The remainder of this paper proceeds as follows. The next section provides the methodology of the study. The third section describes the included papers. Section four presents the discussion of the results. The final section concludes the study and presents relevant future research directions.

2 Methodology

This study follows a rapid review methodology, applying systematic review principles while optimizing time and resource constraints (Hamel et al., 2021). The methodology adheres to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, with modifications to accelerate the review process while maintaining rigor and transparency (Page et al., 2021). To give structure to the aim of the review, we used the PICO framework (Richardson et al., 1995), as seen in Table 1. With the created PICO we built our search strategy.

Table 1: PICO framework

Population	Hospital Nurses (including Registered Nurses, Nurse Practitioners, and other nursing professionals).
Intervention	CDSS
Comparison	Nurses without CDSS or pre-implementation vs. post-implementation studies.
Outcomes	Nursing decision-making, workflow efficiency, and patient outcomes.

1

Search Strategy

This review followed a rapid review methodology, which necessitates prioritization of high-yield databases. We selected PubMed, Embase, and CINAHL due to their comprehensive indexing of nursing and healthcare, literature, domains most relevant to our PICO framework. The following search terms were used for each database: “Clinical Decision Support System” and alternatives, “Nursing” and alternatives, and decision making, Workflow Efficiency, and patient outcomes. The specific search strings for each database can be found in Appendix A. We also found that the volume of nursing CDSS studies increased since 2014, with an average of 154 publications indexed in Pubmed alone per year (compared with an average of 70 per year from 2000-2013) (Akbar et al., 2021). Therefore, we chose to only search for publications after 2014.

Inclusion & Exclusion Criteria

Table 2: Inclusion & Exclusion criteria

Inclusion Criteria	Exclusion Criteria
- Population: Hospital Nurses using CDSS in clinical practice. - Intervention: Studies examining CDSS use for nursing decision-making. - Outcomes: Nursing decision-making, workflow efficiency, patient safety, or patient outcomes. - Study Types: peer reviewed academic journal papers, peer reviewed academic conference papers - Language: English-language studies. - Publication Date: 2014–2024 (last 10 years). - Settings: Hospitals, ICU, emergency departments.	- Focused only on physicians or other healthcare professionals without nurse involvement. - Examined CDSS used exclusively by patients (e.g., patient-facing decision aids).

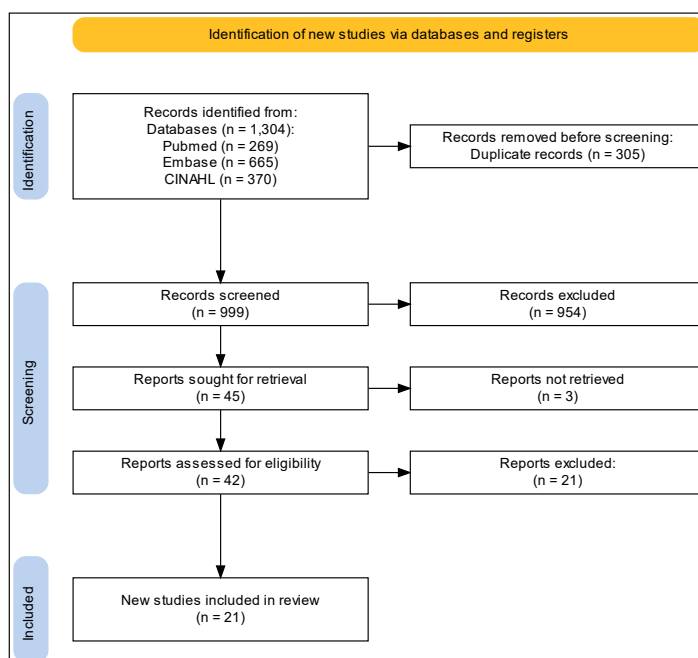


Figure 1: PRISMA flowchart (created in PRISMA Flow Diagram tool (Haddaway et al., 2022))

The PICO was also used to create inclusion and exclusion criteria for the found articles. The inclusion and exclusion criteria can be found in Table 2.

The study selection followed the PRISMA framework. First, a Title & Abstract Screening was performed. After that, the full text of the papers was screened for final inclusion. The included papers were examined to extract key variables related to CDSS usage, nursing decision-making, workflow efficiency, and patient outcomes. The PRISMA flow diagram documented the number of records identified, screened, excluded, and included in the final analysis, see Figure 1.

2 Findings and Synthesis

This synthesis integrates findings from 21 included studies to provide an analysis of how the use of CDSS impacts nursing practice. The results are categorized into three main themes: Nursing Decision-making, Workflow Efficiency, and Patient Outcomes.

2.1 CDSS & Nursing Decision-making

Effective clinical decision-making in nursing requires rapid assessment, prioritization, and risk evaluation, often under conditions of uncertainty. CDSS have been introduced to reduce variability in clinical judgment, ensuring that decision-making is standardized and based on established protocols. Studies have demonstrated the effectiveness of CDSS in triage and emergency decision-making. Chang et al. (2017) and Agnihotri et al. (2021) found that CDSS-assisted triage systems improved the accuracy of patient prioritization, leading to more timely treatment for high-risk patients. Their findings indicated that CDSS reduced inconsistencies in triage classifications between different nursing staff, suggesting an improvement in standardization. In these settings, CDSS facilitated the identification of critical cases through algorithmic risk assessments, which were found to enhance clinical workflow and optimize resource allocation.

Beyond emergency care, CDSS have been shown to contribute to risk assessment and preventative care. Dos Santos Diogo et al. (2023) examined the role of CDSS in diagnostic support and reported that nurses using these systems demonstrated higher diagnostic conformity. Their study highlighted that CDSS-assisted clinical

assessments were more consistent across multiple practitioners, reducing variability in risk identification. Similarly, Jacobsohn et al. (2022) investigated the use of CDSS in fall-risk detection, reporting that automated alerts enabled earlier interventions and a reduction in patient falls. Their findings indicated that CDSS-supported risk assessments aligned with existing prevention protocols, improving adherence to intervention strategies. CDSS have also been evaluated in the context of clinical protocol standardization and guideline adherence. Mebrahtu et al. (2021) found that nurses using CDSS demonstrated higher compliance with infection control protocols, medication safety procedures, and ICU workflows. Their study reported that CDSS improved the consistency of protocol adherence by integrating clinical guidelines into real-time decision-making processes. Kim et al. (2023) examined CDSS-guided pressure ulcer prevention programs and found that their implementation improved adherence to patient repositioning schedules and skin integrity monitoring, leading to a measurable reduction in pressure ulcer incidence rates.

Despite these observed benefits for decision-making, studies have also identified challenges in CDSS implementation for decision-making. Liberati et al. (2017) and Schwartz et al. (2022) reported that nurses frequently overrode CDSS recommendations. Their findings attributed this to a lack of alignment between CDSS-generated alerts and nurses' clinical assessments, leading to a lower rate of system adherence. Hants et al. (2023) analyzed instances in which CDSS recommendations conflicted with nurses' evaluations and found that these conflicts influenced decision-making patterns, sometimes leading to a preference for manual assessments over automated guidance.

2.2 CDSS and Workflow Efficiency

Efficiency in nursing workflows is fundamental for trying to reduce the cognitive overload, optimizing time management, and prioritizing patient care. CDSS have shown promising results in automating documentation, reducing workload, and streamlining care coordination by integrating evidence-based recommendations into clinical workflows. One of the key areas where CDSS have proven effective is reducing administrative redundancy and cognitive burden. Sarıköse & Çelik (2024) found that CDSS implementation in ICUs reduced documentation time, allowing nurses to allocate more time to direct patient care rather than data entry. Their study

reported that automated data entry and structured clinical templates within CDSS contributed to a more efficient documentation process. A similar effect was observed in general hospital workflows, where electronic nursing care plans integrated with CDSS helped eliminate redundant administrative tasks, creating more available time for direct patient interactions (Wong et al., 2015). The study suggests that CDSS can facilitate workflow optimization by automating routine documentation and ensuring that relevant patient information is easily accessible.

Beyond administrative efficiency, CDSS have demonstrated effectiveness in optimizing clinical task prioritization. Akbar et al. (2021) found that CDSS-driven automation enabled nurses to identify and prioritize high-risk patients more effectively, ensuring that urgent situations were addressed in a timely manner. Their study highlighted that automated decision-support algorithms helped reduce delays in intervention by directing nursing attention to the most critical cases. Similarly, Jacobsohn et al. (2022) reported that CDSS-generated fall-risk alerts led to earlier interventions, reducing manual workload and allowing nurses to allocate their time better. These findings indicate that CDSS not only streamline workflow but also support patient management by ensuring that high-risk cases receive prompt attention. CDSS have also been shown to improve workflow coordination within surgical and emergency settings. Quindroit et al. (2024) found that CDSS-supported perioperative workflows led to better coordination between surgical teams, reducing post-operative complications and enhancing overall surgical efficiency. In trauma care, Sanderson et al. (2023) reported that CDSS-assisted transfusion protocols reduced the response times, facilitating faster access to blood transfusions. Similarly, Donsa et al. (2016) found that CDSS-supported diabetes medication management reduced dosage errors, contributing to better treatment accuracy and reducing the time required for dosage adjustments.

Despite these benefits, workflow integration challenges have been identified. Stevens et al. (2024) reported that high alert frequencies within CDSS contributed to alert fatigue, leading to instances where nurses overlooked or overrode critical alerts. This issue was caused to lots of system notifications that disrupted the workflow. Furthermore, Gartrell et al. (2023) found that interoperability issues between separate CDSS and existing EHRs led to workflow disruptions, requiring additional training and adaptation to ensure seamless system use. Additionally, Hants et al. (2023) highlighted that nursing workflows often involve complex, holistic

decision-making processes that cannot always be fully captured by a CDSS, limiting the system's ability to fully align with nursing practice. Both implemented within an EHR as well as in a separate system.

2.3 CDSS and Patient Outcomes

The integration of CDSS in nursing practice has been associated with improvements in patient outcomes, particularly in medication safety, early risk detection, and surgical recovery. One of the most significant impacts of CDSS regarding patient outcomes is in medication safety and error reduction. Stevens et al. (2024) found in a case study that CDSS-driven antibiotic stewardship programs led to a measurable reduction in medication errors, ensuring appropriate dosing and improved adherence to treatment guidelines. Their study highlighted that CDSS-supported prescribing reduced deviations from recommended protocols, minimizing adverse drug reactions. Similarly, Kim et al. (2023) demonstrated in a quasi-experimental study that CDSS-assisted medication administration led to fewer adverse drug events. Their findings suggested that automated medication verification systems improved administration accuracy and reduced human error in high-acuity settings.

CDSSs has also been evaluated in preventative care and early risk detection. Jacobsohn et al. (2022) and Insook Cho et al. (2023) found that CDSS-enabled early warning systems for fall prevention led to a reduction in patient injuries. Their studies reported that real-time alerts allowed nurses to intervene before falls occurred, improving patient safety in hospital environments. In ICU settings, Zhang et al. (2023) explored the role of CDSS-assisted delirium risk assessments and found that the use of these systems helped identify high-risk patients earlier, leading to preventive measures for patients. In surgical and emergency care, CDSS has been found to optimize treatment pathways and reduce complications.

Despite these observed benefits, compliance with CDSS recommendations varies. Paulsen et al. (2021) found that patient adherence to CDSS-generated recommendations was higher than nurse adherence. Their study reported that patients found CDSS guidance motivational in managing their health, whereas nurses sometimes viewed CDSS as an additional task rather than an integral part of workflow processes. These findings suggest that while CDSS can contribute to

improved patient outcomes, its effectiveness is influenced by user engagement and integration into clinical practice.

3 Discussion

The findings of this rapid review illustrate the potential benefits and challenges of CDSS in hospital nursing practice, specifically on the decision-making process, the workflow efficiency and patient outcomes.

Balancing Automation with Clinical Judgment in Decision-Making

Several studies, including those by Agnihotri et al. (2021) and Chang et al. (2017), have reported that CDSS can support more standardized decision-making in specific clinical contexts such as triage and emergency care.

However, this review also highlights a persistent issue, nurses frequently override CDSS recommendations due to a misalignment between system-generated alerts and real-time clinical assessments (Liberati et al., 2017; Schwartz et al., 2022). This raises concerns about the balance between automation and professional judgment. CDSS must be designed to support rather than replace clinical expertise, particularly in high-acuity settings where nursing assessments evolve rapidly. The reluctance to rely fully on CDSS suggests a need for systems that incorporate adaptive learning mechanisms, allowing recommendations to align with patient-specific contexts rather than static algorithms. Moreover, training programs should focus on enhancing nurses' ability to critically evaluate CDSS outputs rather than relying solely on the technology.

Workflow Efficiency: Resolving Disruptions and Enhancing Usability

CDSS have demonstrated the potential to streamline documentation and optimize task prioritization, yet efficiency gains are often hindered by alert fatigue and interoperability issues (Akbar et al., 2021; Sarıköse & Şenol Çelik, 2024). High-alert frequencies, as reported by Stevens et al. (2024), disrupt workflow and reduce adherence to critical notifications, undermining the intended benefits of CDSS. This indicates a need for prioritization algorithms that filter alerts based on clinical urgency rather than generating excessive notifications. Furthermore, the integration

of CDSS with existing EHR systems remains inconsistent, leading to workflow fragmentation (Gartrell et al., 2023). The literature suggests that CDSS effectiveness is highly dependent on seamless interoperability, meaning that future implementations should focus on harmonization with current hospital IT infrastructures to avoid creating additional burdens for nursing staff.

Patient Outcomes

The review findings reinforce that CDSS can contribute to improved patient safety, particularly in medication administration and risk prevention (Kim et al., 2023; Stevens et al., 2024). However, improved decision accuracy or workflow efficiency does not necessarily translate into better clinical outcomes unless these systems are integrated meaningfully into daily practice. For instance, Jacobsohn et al. (2022) found that fall-risk alerts enabled earlier interventions, yet their effectiveness ultimately depended on nurses' adherence to the recommendations. Paulsen et al. (2021) further identified discrepancies in adherence, with patients being more likely to follow CDSS-generated guidance than nurses. This suggests that CDSS success is contingent not only on technical accuracy but also on user engagement. To bridge this gap, further research should explore behavioral and organizational factors that influence CDSS adoption in the long term.

Implications for Nursing Practice and Future Research

Our findings suggest that CDSS will not reach their full potential in nursing practice unless they are further refined to align with the needs, workflows, and values of nurses. Future research should focus on understanding the contextual factors that determine the success or failure of CDSS implementations.

Another area for further exploration is why CDSS succeed in some contexts but fail in others. Differences in organizational culture, training, and system design likely play a role, but further research is needed to identify the specific mechanisms behind these variations. Additionally, the findings showed that many studies focus on how nurses use CDSS, but fewer have examined the fundamental characteristics that define an effective CDSS for nursing practice.

Another area for future research is the nature of CDSS recommendations and how they align with nursing practice. The question arises that a CDSS may be perceived as too rigid or detached from the clinical realities nurses face, particularly in high-acuity or rapidly changing environments.

Finally, future studies should explore how nurses perceive their role in relation to CDSS. Do nurses see CDSS as collaborative tools that support decision-making, or do they view them as systems that dictate actions, reducing clinical autonomy? Understanding this dynamic will be crucial for developing CDSS that nurses trust, use effectively, and integrate seamlessly into patient care.

Limitations

This study has its limitations. One key limitation of the chosen method, a rapid review design, is that it does not assess the methodological quality of the included studies in detail. While the findings suggest that CDSS improve nursing decision-making and workflow efficiency, the reliability of these outcomes is dependent on the rigor of the primary studies. We did not check for potential biases, study design flaws, or long-term patient outcome tracking. Checking the reliability can be important when evaluating patient outcomes.

Another limitation is the variability in CDSS types, nursing teams, and implementation strategies. Some studies examined triage-based CDSS (Agnihotri et al., 2021; Chang et al., 2017), while others focused on medication safety (Stevens et al., 2024) or perioperative care (Quindroit et al., 2024). This makes it hard to compare the CDSSs. The heterogeneity of CDSS applications makes it also difficult to determine generalizable best practices, as not all systems operate with the same level of integration, automation, or clinical relevance. Additionally, the effectiveness of CDSS is influenced by the experience and professional maturity of nurses, as their ability to interpret, adapt, and integrate system recommendations varies across different levels of clinical expertise.

4 Conclusion

This review suggests that CDSS may contribute to improvements in nursing decision-making, optimize workflow efficiency, and enhance patient safety. However, their success in clinical practice depends on proper system design, integration, and acceptance among nurses.

References

- Agnihotri, T., Fan, M., McLeod, S., Borgundvaag, B., Ovens, H., McCarron, J., & Trbovich, P. (2021). Impact of an electronic decision-support system on nursing triage process: A usability and workflow analysis. *Canadian Journal of Nursing Research*, 53(2), 107–113. <https://doi.org/10.1177/0844562119893514>
- Akbar, S., Lyell, D., & Magrabi, F. (2021). Automation in nursing decision support systems: A systematic review of effects on decision making, care delivery, and patient outcomes. *Journal of the American Medical Informatics Association*, 28(11), 2502–2513. <https://doi.org/10.1093/jamia/ocab123>
- Chang, W., Liu, H.-E., Goopy, S., Chen, L.-C., Chen, H.-J., & Han, C.-Y. (2017). Using the five-level taiwan triage and acuity scale computerized system: Factors in decision making by emergency department triage nurses. *Clinical Nursing Research*, 26(5), 651–666. <https://doi.org/10.1177/1054773816636360>
- Chen, Z., Liang, N., Zhang, H., Li, H., Yang, Y., Zong, X., Chen, Y., Wang, Y., & Shi, N. (2023). Harnessing the power of clinical decision support systems: Challenges and opportunities. *Open Heart*, 10(2), e002432. <https://doi.org/10.1136/openhrt-2023-002432>
- Deshpande, S. P. (2009). A study of ethical decision making by physicians and nurses in hospitals. *Journal of Business Ethics*, 90(3), 387–397. <https://doi.org/10.1007/s10551-009-0049-5>
- Donsa, K., Beck, P., Höll, B., Mader, J. K., Schaupp, L., Plank, J., Neubauer, K. M., Baumgartner, C., & Pieber, T. R. (2016). Impact of errors in paper-based and computerized diabetes management with decision support for hospitalized patients with type 2 diabetes. A post-hoc analysis of a before and after study. *International Journal of Medical Informatics*, 90, 58–67. <https://doi.org/10.1016/j.ijmedinf.2016.03.007>
- Dos Santos Diogo, R. C., Silva Butcher, R. D. C. G. E., & Peres, H. H. C. (2023). Diagnostic concordance among nursing clinical decision support system users: A pilot study. *Journal of the American Medical Informatics Association*, 30(11), 1784–1793. <https://doi.org/10.1093/jamia/ocad144>
- Elgin, C. Y., & Elgin, C. (2024). Ethical implications of AI-driven clinical decision support systems on healthcare resource allocation: A qualitative study of healthcare professionals’ perspectives. *BMC Medical Ethics*, 25(1), 148. <https://doi.org/10.1186/s12910-024-01151-8>
- Gartrell, K., Warren, J. I., & Fontelo, P. (2023). PubMed4Hh: A point-of-care mobile app for evidence-based clinical decision support for nurse residents in maryland. *CIN: Computers, Informatics, Nursing*, 41(12), 983–992. <https://doi.org/10.1097/CIN.0000000000001063>
- Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA42020: An R package and shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Systematic Reviews*, 18(2), e1230. <https://doi.org/10.1002/cl2.1230>
- Hamel, C., Michaud, A., Thuku, M., Skidmore, B., Stevens, A., Nussbaumer-Streit, B., & Garritty, C. (2021). Defining rapid reviews: A systematic scoping review and thematic analysis of

- definitions and defining characteristics of rapid reviews. *Journal of Clinical Epidemiology*, 129, 74–85. <https://doi.org/10.1016/j.jclinepi.2020.09.041>
- Hants, L., Bail, K., & Paterson, C. (2023). Clinical decision-making and the nursing process in digital health systems: An integrated systematic review. *Journal of Clinical Nursing*, 32(19–20), 7010–7035. <https://doi.org/10.1111/jocn.16823>
- Insook CHO, MiSoon KIM, & Mi Ra SONG. (2023). Evaluation of a machine learning-based decision support intervention for inpatient falls...19th world congress on medical and health informatics, july 8-12, 2023, new south Wales, australia. *Studies in Health Technology and Informatics*, 310, 1406–1407. [rzh. https://doi.org/10.3233/SHIT231217](https://doi.org/10.3233/SHIT231217)
- Jacobsohn, G. C., Leaf, M., Liao, F., Maru, A. P., Engstrom, C. J., Salwei, M. E., Pankratz, G. T., Eastman, A., Carayon, P., Wiegmann, D. A., Galang, J. S., Smith, M. A., Shah, M. N., & Patterson, B. W. (2022). Collaborative design and implementation of a clinical decision support system for automated fall-risk identification and referrals in emergency departments. *Health Care*, 10(1), 100598. <https://doi.org/10.1016/j.hjdsi.2021.100598>
- Kim, M. S., Ryu, J. M., & Choi, B. K. (2023). Development and effectiveness of a clinical decision support system for pressure ulcer prevention care using machine learning: A quasi-experimental study. *CIN: Computers, Informatics, Nursing*, 41(4), 236–245. <https://doi.org/10.1097/CIN.0000000000000899>
- Kwon, H., & Lee, D. (2024). Clinical decision support system for clinical nurses' decision-making on nurse-to-patient assignment: A scoping review protocol. *BMJ Open*, 14(1), e080208. <https://doi.org/10.1136/bmjopen-2023-080208>
- Liberati, E. G., Ruggiero, F., Galuppo, L., Gorli, M., González-Lorenzo, M., Maraldi, M., Ruggieri, P., Polo Friz, H., Scaratti, G., Kwag, K. H., Vespignani, R., & Moja, L. (2017). What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implementation Science*, 12(1), 113. <https://doi.org/10.1186/s13012-017-0644-2>
- MacIntyre, M. R., Cockerill, R. G., Mirza, O. F., & Appel, J. M. (2023). Ethical considerations for the use of artificial intelligence in medical decision-making capacity assessments. *Psychiatry Research*, 328, 115466. <https://doi.org/10.1016/j.psychres.2023.115466>
- Mebrahtu, T. F., Skyrme, S., Randell, R., Keenan, A.-M., Bloor, K., Yang, H., Andre, D., Ledward, A., King, H., & Thompson, C. (2021). Effects of computerised clinical decision support systems (CDSS) on nursing and allied health professional performance and patient outcomes: A systematic review of experimental and observational studies. *BMJ Open*, 11(12), e053886. <https://doi.org/10.1136/bmjopen-2021-053886>
- Meunier, P.-Y., Raynaud, C., Guimaraes, E., Gueyffier, F., & Letrilliart, L. (2023). Barriers and facilitators to the use of clinical decision support systems in primary care: A mixed-methods systematic review. *Annals of Family Medicine*, 21(1), 57–69. <https://doi.org/10.1370/afm.2908>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>
- Paulsen, M. M., Varsi, C., & Andersen, L. F. (2021). Process evaluation of the implementation of a decision support system to prevent and treat disease-related malnutrition in a hospital setting. *BMC Health Services Research*, 21(1), 281. <https://doi.org/10.1186/s12913-021-06236-3>
- Quindroit, P., Testier, M., Regis, D., Biayimujinga, S., Schricke, O., & Cartron, E. (2024). Clinical decision support systems for operating room nurses: A rapid review. In G. Strudwick, N. R. Hardiker, G. Rees, R. Cook, & Y. J. Lee (Eds.), *Studies in Health Technology and Informatics*. IOS Press. <https://doi.org/10.3233/SHIT240218>
- Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built clinical question: A key to evidence-based decisions. *ACP Journal Club*, 123(3), A12-13.
- Roshanov, P. S., Fernandes, N., Wilczynski, J. M., Hemens, B. J., You, J. J., Handler, S. M., Nieuwlaat, R., Souza, N. M., Beyene, J., Spall, H. G. C. V., Garg, A. X., & Haynes, R. B.

- (2013). Features of effective computerised clinical decision support systems: Meta-regression of 162 randomised trials. *BMJ*, 346(feb14 1), f657–f657. <https://doi.org/10.1136/bmj.f657>
- Sanderson, B. J., Field, J. D., Kocaballi, A. B., Estcourt, L. J., Magrabi, F., Wood, E. M., & Coiera, E. (2023). Clinical decision support versus a paper-based protocol for massive transfusion: Impact on decision outcomes in a simulation study. *Transfusion*, 63(12), 2225–2233. <https://doi.org/10.1111/trf.17580>
- Sanköse, S., & Şenol Çelik, S. (2024). The effect of clinical decision support systems on patients, nurses, and work environment in ICUs: A systematic review. *CIN: Computers, Informatics, Nursing*. <https://doi.org/10.1097/CIN.0000000000001107>
- Schwartz, J. M., George, M., Rossetti, S. C., Dykes, P. C., Minshall, S. R., Lucas, E., & Cato, K. D. (2022). Factors Influencing Clinician Trust in Predictive Clinical Decision Support Systems for In-Hospital Deterioration: Qualitative Descriptive Study. *JMIR Human Factors*, 9(2), e39960. <https://doi.org/10.2196/33960>
- Shah, N. R., Khetpal, V., & Erqou, S. (2022). Anticipating and addressing challenges during implementation of clinical decision support systems. *JAMA Network Open*, 5(2), e2146528. <https://doi.org/10.1001/jamanetworkopen.2021.46528>
- Stevens, E. R., Xu, L., Kwon, J., Tasneem, S., Henning, N., Feldthouse, D., Kim, E. J., Hess, R., Dauber-Decker, K. L., Smith, P. D., Halm, W., Gautam-Goyal, P., Feldstein, D. A., & Mann, D. M. (2024). Barriers to implementing registered nurse-driven clinical decision support for antibiotic stewardship: Retrospective case study. *JMIR Formative Research*, 8, e54996. <https://doi.org/10.2196/54996>
- Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digital Medicine*, 3, 17. <https://doi.org/10.1038/s41746-020-0221-y>
- Taheri Moghadam, S., Sadoughi, F., Velayati, F., Ehsanzadeh, S. J., & Poursharif, S. (2021). The effects of clinical decision support system for prescribing medication on patient outcomes and physician practice performance: A systematic review and meta-analysis. *BMC Medical Informatics and Decision Making*, 21(1), 98. <https://doi.org/10.1186/s12911-020-01376-8>
- Wong, C. M., Wu, S. Y., Ting, W. H., Ho, K. H., Tong, L. H., & Cheung, N. T. (2015). An electronic nursing patient care plan helps in clinical decision support. *Studies in Health Technology and Informatics*, 216((Wong C.M.) Health Informatics Section, Hospital Authority, Hong Kong SAR), 945. Medline.
- Zhang, S., Ding, S., Cui, W., Li, X., Wei, J., & Wu, Y. (2023). Impact of clinical decision support system assisted prevention and management for delirium on guideline adherence and cognitive load among intensive care unit nurses (CDSSD-ICU): Protocol of a multicentre, cluster randomized trial. *PLOS One*, 18(11), e0293950. <https://doi.org/10.1371/journal.pone.0293950>

Appendix A:**Pubmed**

("Clinical Decision Support Systems"[MeSH] OR "Decision Support Systems" OR "Clinical Decision Support" OR "CDSS")

AND

("Nursing"[MeSH] OR "Nurses"[MeSH] OR "Nursing Care" OR "Nurse" OR "Nursing Practice" OR "Nursing Informatics")

AND

("Decision Making"[MeSH] OR "Clinical Reasoning" OR "Efficiency" OR "Workflow" OR "Patient Outcome Assessment"[MeSH] OR "Patient Safety" OR "Quality of Care")

Embase

('clinical decision support system'/exp OR 'decision support system' OR 'clinical decision support' OR 'CDSS')

AND

('nursing'/exp OR 'nursing care' OR 'nurse'/exp OR 'nurses' OR 'nursing practice' OR 'nursing informatics')

AND

('clinical decision making'/exp OR 'workflow' OR 'efficiency' OR 'patient outcome'/exp OR 'patient safety' OR 'quality of care')

CINAHL

("Clinical Decision Support Systems" OR "Decision Support Systems" OR "Clinical Decision Support" OR "CDSS" OR "Computerized Decision Support")

AND

("Nursing" OR "Nurses" OR "Nursing Care" OR "Nurse" OR "Nursing Practice" OR "Nursing Informatics")

AND

("Decision Making" OR "Clinical Reasoning" OR "Efficiency" OR "Workflow" OR "Patient Outcome Assessment" OR "Patient Safety" OR "Quality of Care")

AI-BASED BUSINESS MODEL INNOVATION: A CHATGPT INTEGRATION APPROACH

FLORIAN BONGARD, JÜRGEN MOORMANN

Frankfurt School of Finance & Management, Management Department, Frankfurt am
Main, Hesse, Germany
j.moormann@fs.de, florianbongard13@gmail.com

This paper addresses a significant gap in the current literature by exploring the configuration of AI to support the process of business model innovation. Design science research is applied, and a ChatGPT-based prototype is developed to generate business models by recombining industry-independent business model patterns. The aim is therefore to investigate how such an AI assistant can be effectively configured. The study uses case studies from various industries to demonstrate the configured chatbot and evaluate its ability to adapt to different industry contexts and provide customized, innovative business models. The findings were used for refining the prototype, demonstrating its capabilities and evaluating the results. The assessment of the usability and performance of the AI assistant indicates the potential of the prototype and future developments. The study shows that the explorative configuration process can serve as a blueprint for all researchers and practitioners who want to develop innovative business models.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.23](https://doi.org/10.18690/um.fov.4.2025.23)

ISBN

978-961-286-998-4

Keywords:

AI,
business model,
business model innovation,
chatbot,
ChatGPT



University of Maribor Press

1 Introduction

There is broad consensus in both the literature and in practice that the long-term competitive success of companies depends on their ability to develop innovative business models (Hedman & Kalling, 2003). However, the nature of Business Model Innovation (BMI) remains inherently complex and unpredictable (Chesbrough, 2010; Lindgardt et al., 2009). Organisations have to deal with considerable ambiguity and uncertainty, which requires a substantial investment of time and energy to explore various possibilities before an optimal solution can be found (Osterwalder & Pigneur, 2010). Therefore, creativity and divergent thinking are crucial parts of developing an innovative Business Model (BM) (Gassmann et al., 2014).

In the past, it was assumed that this creativity could only be achieved by humans. However, in 2020 Amabile proposed that creativity could also be achieved by Artificial Intelligence (AI) in the future. In 2021 Haefner et al. discussed AI's potential to undertake traditional human tasks in organizations, particularly in innovation. They highlighted recent advancements in AI algorithms, indicating AI's potential to address the more challenging aspects of innovation management, such as generating completely novel ideas. A mere year later, in November 2022, the release of OpenAI's ChatGPT significantly improved AI's capabilities in producing human-like work (Dell'Acqua et al., 2023; OpenAI, 2022). Given these capabilities, the integration of human work with AI presents both fundamental challenges and opportunities, especially in knowledge-intensive domains, such as BMI.

These Generative AI (GenAI) models, such as ChatGPT, create new data samples based on learned patterns (Feuerriegel et al., 2023). This presents a significant opportunity to use GenAI for BMI, as it can learn similarities in BMs, known as Business Model Patterns (BMPs), including characteristics, arrangement of building blocks and behaviours (Osterwalder & Pigneur, 2010). Based on these patterns, GenAI should then be able to create new BMs.

Nevertheless, every GenAI model is incomplete and necessitates additional fine-tuning for specific tasks (Feuerriegel et al., 2023). Currently, there is no research on how those AI assistants can be effectively configured to utilize BMPs for developing innovative BMs. Thus, this paper seeks to answer the research question: *How can AI*

assistants, particularly ChatGPT, be configured to utilize business model patterns for developing innovative business models across industries?

This paper contributes to the existing literature by providing insights into the integration of AI in innovation management, particularly in BMI. In addition, the paper offers a guideline for practitioners to utilising AI capabilities to enhance BMI.

2 Theoretical and Technological Foundation

2.1 Business Model

Osterwalder and Pigneur (2010, p. 14) define a BM as the rationale underlying how an organization “creates, delivers, and captures value”. That is analogous to a blueprint for a strategy, which is executed through organizational structures, processes and systems. A BM comprises nine essential components: customer segments, value proposition, channels, customer relationships, revenue streams, key resources, key activities, key partnerships and cost structure. Collectively, these elements constitute the Business Model Canvas (BMC), a tool designed to illustrate and assess both existing and prospective BMs (Osterwalder & Pigneur, 2010).

To improve the understanding of BMs, an additional framework known as the Magic Triangle can be integrated into the BMC (Gassmann et al., 2014). This framework introduces four critical dimensions that align with the nine components of the BMC: who the customer is, what the customer is purchasing, how the product or service is produced, and why the business is profitable. The “who-what-how-why” framework (Gassmann et al., 2014) is a descriptive approach to BMs. The initial two dimensions (“who” and “what”) pertain to external aspects, whereas the latter two dimensions (“how” and “why”) relate to internal aspects.

The arrangement of the building blocks of a BM can be represented by BMPs. Furthermore, empirical studies have shown that the use of BMPs enables companies to reduce the effort required to develop BMs, while the pattern library reliably increases the opportunities for innovation (Lüttgens & Diener, 2016). This is consistent with the findings of Gassmann et al. (2014), who discovered that 90% of all new BMs are not entirely original but are derived from 55 existing patterns. The key to success lies in the ability to identify and apply successful patterns from other

industries to one's own. This process entails comprehending the patterns, translating them into a contextually relevant framework, recombining them in a manner that is pertinent to one's industry, and transferring them in an efficacious manner (Enkel & Mezger, 2013). By creatively imitating BMs from other industries, a business can therefore position itself as an innovation leader within its industry.

Based on these BMPs, Gassmann et al. (2014) developed the Business Model Navigator (BMN), a tool for streamlining BMI through the recombination of the 55 patterns. It is designed to be used in a workshop setting and follows four steps: initiation, ideation, integration and implementation.

2.2 Large Language Models

The advent of Large Language Models (LLMs) has significantly changed the interaction dynamics between chatbots and users. LLMs are trained on huge and diverse language datasets in advance, enabling them to understand and generate texts that closely mimic human communication in a variety of contexts (Pappula & Allam, 2023). Furthermore, it is possible to fine-tune LLMs for specific applications, industries or user bases, allowing for a more personalized user experience. LLMs serve as the basis for GenAI models (Chui et al., 2023). The term “generative” is employed to describe the capacity of the model to generate content that is both novel and meaningful, based on the training data (Feuerriegel et al., 2023).

A prevalent family of LLMs are generative pre-trained transformers (GPTs) that are used in the conversational agent ChatGPT (Feuerriegel et al., 2023). The latest version (as of October 2024), GPT-4o, shows superior performance in Massive Multitask Language Understanding (MMLU), a crucial benchmark test for evaluating the multitask accuracy of text models (Hendrycks et al., 2021), compared to other language models (OpenAI, 2024).

GPT Builder is a tool that allows users to create customized versions of ChatGPT tailored for specific tasks or domains. Since the GPT Builder was only announced by OpenAI in November 2023, there is little academic research and best practices for customising GPTs. So far, only a few studies have been published that focus on LLM technology for BMI (e.g. Collins et al., 2024; Aagaard & Tucci, 2024; Machucho & Ortiz, 2025). However, the authors do not document the process, but

focus on the results of the customised GPT or address the topic from a more general perspective. The present paper aims to fill this research gap by examining and detailing the configuration process.

3 Methodology

Our research focusses on the prototype development of the AI assistant in order to answer the research question. The procedure used in the study adheres to the Design Science Research Methodology (DSRM) as proposed by Peffers et al. (2007). The DSRM was selected due to its systematic approach to the creation and evaluation of information system (IS) artifacts, which aligns with the objectives of this study. The research design follows in particular the adaption by Fayoumi and Loucopoulos (2021) in order to ensure academic rigor. It consists of five steps:

The first step, *problem identification*, addresses the inherent complexity and unpredictability of BMI. Secondly, the primary *objective of our study* was to improve the BMI process and thereby increase its reliability, efficiency and straightforwardness. To achieve this, the solution involves the implementation of an AI assistant that automates the key aspects of the BMI process. Thirdly, in terms of *design and development*, the GPT-based AI assistant should comprehend and utilize various BMPs, providing a structured framework for generating novel BMs in different industries. The fourth step was the *demonstration* of the configured GPT. This step focused on a technical assessment of the prototype, i.e. the GPT's ability to accurately follow its configuration and to create coherent BMs based on the BMPs. The demonstration was carried out on the basis of case studies from different industries. After completing this step, we returned to the design and development phase to improve the prototype's effectiveness. Then we proceeded to the fifth step, *evaluation*, by presenting the revised prototype of the chatbot to eight high-profile experts and to assessing its applicability based on the resulting interviews.

4 Design and Development

This section is about creating the artefact, i.e. using the GPT Builder to develop a customised GPT model for business model innovation using BMPs. The tool was used because it showed the best performing LLM at the time of our investigation, as described in subsection 2.2.

4.1 Creation of the Training Data

Initially, a comprehensive set of BMPs was identified and curated from existing literature. These patterns were integrated into the GPT Builder. To facilitate this process, data was primarily gathered from the website *businessmodelnavigator.com*, as it provides more recent information than the book by Gassmann et al. (2014). An Excel spreadsheet was set up that encompasses the total of 55 (plus 5 additional) BMPs. The data provided for each pattern includes several critical elements: a) pattern name serving as the identifier for each BMP, b) pattern description containing a summary of the key characteristics and principles of the pattern, c) BM components including the “how”, “what”, “who” and “why” details according to the Magic Triangle framework, d) industries in which each pattern is predominantly deployed, e) insights into the most frequently occurring patterns in relation to each other, providing an overview of potential synergies and common combinations, and f) examples of companies that have implemented these patterns.

Additionally, the GPT was provided with information about the process it should follow. Given that the objective of the chatbot is to develop new BMs, the GPT was trained exclusively on the design phase (i.e. initiation, ideation and integration), which involves conceptualizing the proposed BM, and not on the realization phase, which would include the implementation of the aforementioned BM.

4.2 Configuration of the Chatbot

The configuration process takes place in a conversational setting, wherein the user interacts with an AI (in this case ChatGPT) that helps him/her configure his/her own GPT model. The process starts with “Create” and is followed by the system’s question “What would you like to make?”. The process commenced with the following prompt to outline the GPT’s objective and included uploading the Excel spreadsheet with detailed information on the 60 BMPs:

“Make a strategy and innovation expert who creates innovative business models by recombining different business model patterns. It should use only the 60 business model patterns provided in the attached Excel spreadsheet. The Excel list also includes a column called “BM Components,” which refers to the Magic Triangle (Gassmann et al., 2014). A business model defines who your customers are, what you sell, how you produce your offering and why your business is profitable. Who-what-how-why describes a business model where the first two (who and what) address its external aspects

and the second two (how and why) address its internal dimensions. It should follow 3 steps: 1. initiation: It should analyze the ecosystem of the company (players and drivers of change) and also the current business model (if there is one); 2. Ideation: It should adapt the patterns, so it should challenge the current business model using the 60 business model patterns by applying the similarity principle and the confrontation principle; 3. Integration: It should now integrate these ideas into a new business model, making it internally and externally consistent.”

Based on this prompt, the AI suggested a name (“Innovative Strategist”) and a profile picture for the GPT, both of which were accepted. Next, the communication style of the Innovative Strategist was refined to ensure that it would act in a manner consistent with that of a consultant in their interactions with clients. This is because the GPT should emulate the depth and expertise of a subject matter expert or, in this case, a consultant, in the field.

After these first steps, a preview of the GPT was created to check the chatbot’s configuration so far. A new chat appeared to the right of the configuration window (i.e. the chat in which we had configured the GPT), in which we could test now our new (own) GPT. Accordingly, the so-called “Silly Cow Exercise” was employed. In accordance with the typical procedures employed in conventional BMI settings, the chatbot was instructed to illustrate a BM utilizing a cow (see details in Osterwalder & Pigneur, 2010).

The outcome of this preliminary trial was then employed to further improve the GPT. This process was highly iterative, entailing numerous cycles of testing and refinement. Moreover, the GPT explained certain aspects of the BMI process to the user, instead of applying them. For example, the GPT explained what “challenging the current business model (external/internal)” is instead of applying it itself. Therefore, another prompt was entered to ensure that the model adapts the patterns during the ideation phase according to the similarity and confrontation principle.

In conclusion, in addition to the initial prompt, 19 further prompts were employed to refine the behavior of the GPT. As a result of these refinements, the GPT was able to categorize the patterns exactly according to the dimensions of the Magic Triangle and to use only the provided patterns. Furthermore, the GPT has demonstrated the ability to follow the prescribed three-stage process (initiation, ideation and integration) and generate coherent BMs.

5 Demonstration

This phase of the DSRM is intended to assess the efficacy of the configured GPT in realistic settings. To this end, we have selected case studies from various industries and applied the chatbot to them to generate new BMs. This phase serves to validate the experimental configuration and training conducted in the previous phase, thus ensuring that the AI assistant can operate effectively in diverse industries.

5.1 Case Selection

To ensure a comprehensive analysis, five cases were selected from different industries, namely consumer goods, information technology, retail, financial services and media. The cases not only represent a range of different industries, but also different levels of complexity defined by the amount of data involved and the length of conversational dependencies. The reason for this was to thoroughly evaluate the chatbot's ability to handle and respond to different levels of complexity and long-term dependencies. In addition, both start-up and established companies were included to gain insights into the chatbot's performance at different stages of organizational maturity.

5.2 Application of the Configured GPT

For each of the cases we formulated a brief instruction which was individually entered into our GPT "Innovative Strategist". The instructions were formulated as a customer would speak to a consultant. All responses were documented.

The procedure can be shown at the example of the company "air up" from consumer goods industry. air up is a start-up company that was launched in 2019 (www.air-up.com). It offers a special water bottle that alters the flavor of the water simply by adding scents. Distinctive "aroma pods are affixed to the top of the bottle to achieve this effect. Based on that concept, the chatbot was queried with the following prompt:

“I’ve developed a game-changing product: a water bottle that changes the flavor of the water without adding any ingredients. All you have to do is stick different pods on top of the water bottle and it instantly changes the flavor. I’m looking for your help in creating a business model that will allow me to disrupt the market with my idea.”

Table 1 summarizes the results of this one of five case studies examined. A transcript of the complete result of this case is available in the Appendix A.

Table 1: Results of the demonstration phase (example of the consumer goods case)

Case	Recommended patterns	Business model
Consumer Goods: “air up”	- Razor and Blade - Subscription - Lock-in - Cross-Selling - Mass Customization - Experience Selling	Who: Health-conscious individuals; Fitness enthusiasts; Environmentally conscious consumers; Tech-savvy consumers looking for innovative products What: Core Product: A reusable water bottle that offers instant flavour changes without adding any ingredients; Supplementary Products: A variety of flavour pods catering to different tastes and dietary preferences; Unique Experience: An app or community platform that enhances user engagement and tracks hydration habits How: Design and Manufacturing: Collaborate with innovative flavour technology firms and high-quality manufacturers; Distribution Channels: Utilize online platforms, health and fitness stores, and direct partnerships with wellness brands; Customer Interaction: Develop a robust online presence with a user-friendly website and app, offering customization and subscription options Why: Razor and Blade Model: Generate recurring revenue through the sale of flavour pods; Subscription Service: Ensure steady income through subscription plans for regular pod deliveries; Premium Pricing for Customization: Charge a premium for personalized pods and bottle designs

6 Evaluation

Following the DSRM, the prototype of the GPT was presented to industry experts with the intention of assessing its usability (Rubin & Chisnell, 2008). For the study, semi-structured interviews were conducted, which are particularly suitable for usability-related questions. The interviewees were selected based on their specific knowledge and experience in areas relevant to the GPT’s objectives and who would potentially use the chatbot. Therefore, eight high-profile experts in the fields of venture building, strategy consulting, BMI, and business development across diverse

domains were selected (Rubin & Chisnell, 2008). Each expert (E) was provided with a concise overview of the objectives of the chatbot, followed by a demonstration of the aforementioned case studies. Appendix B provides a summary of each expert's expertise. The duration of each interview was on average 34 minutes including an extensive demonstration of the prototype.

The interviews were recorded and then transcribed using the speech recognition software Whisper to facilitate the analysis. For the interpretation of the material, inductive category formation was employed which included the prototype's usefulness, efficiency, effectiveness, satisfaction and learnability (Rubin & Chisnell, 2008). The analysis was carried out in eight steps, which were suggested by Mayring (2014). This resulted in four main categories (initial impression, functional capabilities, usability in practice, future developments) and nine subcategories. A summary of the interviewees' evaluations is provided in Appendix C.

7 Discussion

The first phase of prototype testing, based on the case studies, showed that the GPT was largely able to follow the structured BMI process, to analyze the request accurately and to identify relevant patterns from the provided list. Moreover, these patterns from different industries were effectively transferred and adapted to the specific case. This capacity is consistent with the innovation principles delineated by Gassmann et al. (2014). Furthermore, the GPT integrated these patterns into comprehensive and cohesive BMs, addressing all relevant dimensions in a systematic and coherent manner. These findings indicate that the exploratory configuration process presented in this study was both effective and reliable, thus confirming its efficacy.

However, the results varied across the case studies. Three pivotal issues were identified and subsequently addressed in the reconfiguration of the GPT. Two of these issues, namely the usage of patterns that are not included in the provided list and the omission of patterns after the ideation phase, can be attributed to hallucinations of the chatbot. In our study, the GPT utilized patterns not explicitly provided in the Excel spreadsheet. These instances, in which the model generated extrinsic hallucinations by drawing on information not provided before, are consistent with the behavior described by Maynez et al. (2020). This problem likely

stems from the nature of LLMs like GPT-4o, which are pre-trained on vast amounts of data. This pre-training can sometimes cause the model to produce hallucinated content if the data is inadequate (Zhang et al., 2023).

The third issue is related to the lack of coverage of the four dimensions of the Magic Triangle and was mainly observed in the first two cases. The chatbot did not suggest any patterns that covered all four dimensions of the BM, or rather, no patterns that considered the dimensions “who” and “how”. Although the chatbot also did not suggest a pattern for the “who” dimension in the third and fourth case, this omission was appropriate since a BM was already provided and not all dimensions needed to be addressed according to Gassmann et al. (2014). The missing “how” dimension can be attributed to the hallucinations as the omitted patterns would have covered this dimension.

The second phase of prototype testing was based on the assessments of professionals, who unanimously expressed a positive first impression of the prototype. This suggests that the experts, as potential users, would also use the GPT productively in actual application scenarios (Rubin & Chisnell, 2008).

A particularly noteworthy aspect was the tool’s capacity to save a substantial amount of time and effort for certain tasks. This indicates that the GPT is highly effective at automating seemingly creative but repetitive tasks and accelerating decision-making processes. This finding corresponds to Kumar et al. (2024), who found efficiency gains from integrating AI technologies into business automation. Furthermore, it is consistent with the findings of Candelon et al. (2023), who investigated the efficacy of AI assistants in product innovation.

Nevertheless, several areas for improvement were identified. One area for enhancement is the inclusion of supplementary examples. The repeated request for concrete examples to illustrate the presented patterns indicates that users find it challenging to translate theoretical models into practical applications. Thus, the incorporation of case studies and real-world examples could markedly enhance the GPT’s usability and effectiveness.

Another crucial area is customization, i.e. the tool should be more adaptable to the specific needs of individual users, suggesting that a universal approach may be inadequate. This means that the tool should include features that assess users' knowledge and customize their responses accordingly to improve the user experience for persons at different levels of expertise.

Although this paper offers important insights into the configuration and application of GPT models, there are several limitations to be aware of. First, further iterations of the prototype could have been used to assess the impact of different prompting variants on performance. Second, the lack of information about the source and methodology used by the GPT to obtain its data makes it impossible to determine whether the GPT merely replicated a BM or created an entirely new one. This aspect is closely linked to the black box problem of AI (von Eschenbach, 2021). Third, key variables such as effectiveness and efficiency were not quantitatively measured. Quantitative assessment could have strengthened the results (Rubin & Chisnell, 2008).

8 Conclusion

The results of this study reveal that tools like ChatGPT, when configured appropriately, can significantly improve the innovation process by providing structured, creative solutions that are both practical and adaptable. In this regard, the main contribution of our analysis lies in the area of BMI and how this can be supported by an appropriate AI configuration. It is crucial to note that the results of this study are time-limited and may quickly become outdated due to the rapid progress in LLMs. As AI continues to grow, its role in strategic management is likely to increase, making the insights gained from this research increasingly relevant for both academic research and practical application.

References

- Aagaard, A., & Tucci, C. (2024). AI-Driven Business Model Innovation: Pioneering New Frontiers in Value Creation. Aagaard, A. (eds) Business Model Innovation. Game Changers and Contemporary Issues. Palgrave Macmillan. Cham. 295-328.
- Amabile, T. M. (2020). Creativity, Artificial Intelligence, and a World of Surprises. *Academy of Management Discoveries*, 6(3), 351-354.
- Candelon, F., Kraymer, L., Rajendran, S., & Zuluaga Martínez, D. (2023). How People Can Create—and Destroy—Value with Generative AI. BCG Henderson Institute, Boston Consulting Group.

- Chesbrough, H. (2010). Business Model Innovation: Opportunities and Barriers. *Long Range Planning*, 43(2-3), 354-363.
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zimmel, R. (2023). The economic potential of generative AI: The next productivity frontier. McKinsey & Company.
- Collins, B. R., Black, E. W., & Rarey, K. E. (2024). Introducing AnatomyGPT: A customized artificial intelligence application for anatomical sciences education. *Clinical Anatomy*, 37(6), 661-669.
- Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayner, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: field experimental evidence of the effects of AI on knowledge worker productivity and quality. Working Paper, No. 24–013. Harvard Business School. Boston, MA.
- Enkel, E., & Mezger, F. (2013). Imitation Processes And Their Application For Business Model Innovation: An Explorative Study. *International Journal of Innovation Management*, 17(1), 1-34.
- Fayoumi, A., & Loucopoulos, P. (2021). Bridging the strategy execution gap of designing intelligent talent acquisition systems using enterprise modelling and simulation. *Enterprise Information Systems*, 17(6), 781-816.
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2023). Generative AI. *Business & Information Systems Engineering*, 66(1), 111-126.
- Gassmann, O., Frankenberger, K., & Csik, M. (2014). *The Business Model Navigator: 55 Models That Will Revolutionise Your Business*. FT Publishing. Upper Saddle River, NJ.
- Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change*, 162(C).
- Hedman, J., & Kalling, T. (2003). The Business Model Concept: Theoretical Underpinnings and Empirical Illustrations. *European Journal of Information Systems* 12(1), 49-59.
- Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D., & Steinhardt, J. (2021). Measuring Massive Multitask Language Understanding. *Proceedings of the 9th International Conference on Learning Representations*, virt.
- Lindgardt, Z., Reeves, M., Stalk, G., & Deimler, M. S. (2009). *Business Model Innovation – When the Game Gets Tough, Change the Game*. Boston Consulting Group.
- Lüttgens, D., & Diener, K. (2016). Business Model Patterns Used as a Tool for Creating (new) Innovative Business Models. *Journal of Business Models*, 4(3), 19-36.
- Machucho, R., & Ortiz, D. (2025). The Impacts of Artificial Intelligence on Business Innovation: A Comprehensive Review of Applications, Organizational Challenges, and Ethical Considerations. *Systems*, 13(4), 264.
- Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Klagenfurt, Austria.
- Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On Faithfulness and Factuality in Abstractive Summarization. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1906-1919.
- OpenAI (2022). Introducing ChatGPT. Retrieved July 5, 2024 from <https://openai.com/index/chatgpt/>.
- OpenAI (2024). Hello GPT-4o. Retrieved July 20, 2024 from <https://openai.com/index/hello-gpt-4o/>.
- Osterwalder, A., & Pigneur, Y. (2010). *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. John Wiley & Sons. Hoboken, NJ.
- Pappula, S. R., & Allam, S. R. (2023). LLMs for Conversational AI: enhancing chatbots and virtual assistants. *International Journal of Research Publication and Reviews*, 4(12), 1601-1611.
- Peffer, 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.

- Rubin, J., & Chisnell, D. (2008). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests* (2nd ed.). Wiley Publishing. Indianapolis, IN.
- von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology*, 34(4), 1607-1622.
- Zhang, Y., Li, Y., Cui, L., Cai, D., Liu, L., Fu, T., Huang, X., Zhao, E., Zhang, Y., Chen, Y., Wang, L., Luu, A. T., Bi, W., Shi, F., & Shi, S. (2023). Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models. *arXiv:2309.01219*. Cornell University, Ithaca, NY.

Appendix A:

Output of the GPT at the example of the consumer goods case:

1	USER: I've developed a game-changing product: a water bottle that changes the flavour of the water without adding any ingredients. All you have to do is to stick different pods on top of the water bottle and it instantly changes the flavour. I'm looking for your help in creating a business model that will allow me to disrupt the market with my idea.
2	CHATGPT: That sounds like an exciting and innovative product! Let's go through the process of developing a robust business model for your flavoured water bottle using the three phases: initiation, ideation, and integration. Initiation Phase 1. Ecosystem Analysis: – Players: ○ Direct Competitors: Other water bottle companies, flavoured water brands. ○ Indirect Competitors: Beverage companies, especially those focusing on flavoured drinks. ○ Partners: Pod manufacturers, distribution channels (online retailers, physical stores), marketing agencies. – Drivers of Change: ○ Technological Advances: Innovation in flavour technology and materials. ○ Consumer Trends: Increasing demand for health-conscious, sugar-free beverages. ○ Environmental Concerns: Growing preference for reusable products to reduce plastic waste. 2. Current Business Model (if any): Not applicable. Ideation Phase 1. Challenging the Current Business Model: Using the 60 business model patterns from the provided list, let's explore potential patterns that could innovate your business model: – Razor and Blade: Sell the water bottles at a lower margin or even at cost, while making higher margins on the flavour pods. ○ <i>Description:</i> Offer the primary product (razor) at a low price and earn profits from consumables (blades). – Subscription: Offer a subscription model where customers receive new flavour pods regularly.

	○ Description: Shift from selling products to providing a service, generating recurring revenue. – Lock-in: Develop proprietary pods that only work with your water bottles. ○ <i>Description:</i> Create dependency on your consumable products to ensure repeated purchases. – Cross-Selling: Partner with health and wellness brands to offer bundled products or discounts. ○ Description: Increase sales by offering related products to your customer base. – Mass Customization: Allow customers to personalize their flavour experiences through custom pod options. ○ Description: Provide tailor-made products to meet individual customer preferences. – Experience Selling: Create an immersive experience around the product, such as a community or app that tracks water intake and flavour preferences. ○ Description: Sell not just the product but the experience associated with it. 2. Application of Similarity and Confrontation Principles: – Similarity Principle: Analyse successful business models in the beverage industry (e.g., soda stream, coffee pods) and adapt their strategies. – Confrontation Principle: Challenge the norms in the reusable water bottle market by integrating high-tech features and unique value propositions. Integration Phase Magic Triangle of the New Business Model: – Who (Target Customer): ○ Health-conscious individuals ○ Fitness enthusiasts ○ Environmentally conscious consumers ○ Tech-savvy consumers looking for innovative products – What (Value Proposition): ○ Core Product: A reusable water bottle that offers instant flavour changes without adding any ingredients. ○ Supplementary Products: A variety of flavour pods catering to different tastes and dietary preferences. ○ Unique Experience: An app or community platform that enhances user engagement and tracks hydration habits.
--	---

	– How (Value Chain): ○ Design and Manufacturing: Collaborate with innovative flavour technology firms and high-quality manufacturers. ○ Distribution Channels: Utilize online platforms, health and fitness stores, and direct partnerships with wellness brands. ○ Customer Interaction: Develop a robust online presence with a user-friendly website and app, offering customization and subscription options. – Why (Profit Mechanism): ○ Razor and Blade Model: Generate recurring revenue through the sale of flavour pods. ○ Subscription Service: Ensure steady income through subscription plans for regular pod deliveries. ○ Premium Pricing for Customization: Charge a premium for personalized pods and bottle designs. This integrated business model should help you disrupt the market and create a strong, sustainable revenue stream.
--	--

Appendix B:

Overview of the participants of the expert interviews

ID of the interviewee	Brief description of relevant expertise
E1	Managing director of an IT company, responsible for the innovation of the current BM; co-founder of several companies
E2	Project manager core processes and digitalization at a large automotive company, responsible for internal start-up development and cooperations with start-ups
E3	Founder & CEO of a crypto service company; co-founder of digital transformation company; former strategy and innovation consulting experience at an international business consultancy
E4	Co-founder of a national consulting company, responsible for innovation management, digital strategy, and transformation
E5	Senior manager at an international business consultancy, responsible for finance transformation and innovation
E6	Partner at an international business consultancy, responsible for sales and marketing transformation
E7	Founder of a software development company, responsible for business and product strategy consulting; former product portfolio manager at a large international software company, responsible for BMI

E8	Venture builder at a large international logistics company, responsible for creating new BMs; co-founder of two start-up companies
----	--

Appendix C:

Summary of the interviewees' evaluations

Main categories and subcategories	Assessments of the experts
1. Initial Impression	
Positive Impressions	E1, E2, E3, E4, E5, E6, E7, E8
Neutral or negative Impressions	E6
2. Functional Capabilities	
a) Highlighted Features	
Time and Effort Savings	E1, E4, E7
Structured Process	E2, E3, E6, E8
Creativity and Idea Generation	E2, E4, E6, E7, E8
Quality and Content	E1, E2, E3, E4, E5, E6, E7, E8
b) Insufficient or Missing Features	
Example	E3, E4, E5
Customization	E2, E3, E4, E8
Guidance	E2, E3, E5, E7
Visualization	E1, E3, E4, E6
3. Usability in Practice	
Challenges in Application	E1, E6
Would use the GPT	E1, E2, E3, E4, E5, E6, E7, E8
4. Future Developments	
Evaluation of BM	E6, E8
Implementation Steps	E1, E4, E7
Market and Financial Analysis	E3, E4, E7

VIRTUAL CARE CENTERS AND THE EVOLVING ROLE OF NURSES

HENK SONNEVELD,¹ APPHIA TAN,² JOBBE P.L. LEENEN,^{3,4}
THIJS VAN HOUWELINGEN¹

¹ University of Applied Sciences Utrecht, Research Group Technology for Healthcare Innovations, Research Centre for Healthy and Sustainable Living, Utrecht, the Netherlands

henk.sonneveld@hu.nl, thijs.vanhouwelingen@hu.nl

² National Cancer Centre Singapore, Division of Supportive and Palliative Care, Singapore, Singapore

apphia.tan.jq@nccs.com.sg

³ Connected Care Center, Isala, Zwolle, the Netherlands

j.p.l.leenen@isala.nl

⁴ Windesheim University of Applied Sciences, Research Group IT Innovations in Healthcare, Zwolle, the Netherlands

j.p.l.leenen@isala.nl

New Virtual Care Centers (VCCs) within hospitals utilize information technology to remotely monitor and support patients with chronic diseases living at home. Nurses play a crucial role by providing remote coaching and guidance to help patients manage their conditions. Currently, there is a growing understanding regarding the evolving roles and responsibilities of nurses in VCCs, however studies have yet to establish connections with educational frameworks, which poses a challenge for nursing education programs to prepare students for this emerging professional role effectively. Our study aimed to provide insights into the evolving roles, tasks, and responsibilities of nurses providing remote care as per the CanMEDS framework. We conducted a qualitative content analysis of 15 interviews. Nursing work within VCCs is represented by the seven CanMEDS roles. Most tasks align with the roles of Leader and Collaborator, while Quality Promotor has the fewest. Our study maps the responsibilities and tasks of VCCs' care delivery to the core roles of nurses.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.24](https://doi.org/10.18690/um.fov.4.2025.24)

ISBN

978-961-286-998-4

Keywords:

telemonitoring,
nursing roles,
tasks,
digital health,
virtual care unit



University of Maribor Press

1 Introduction

As the global population ages and chronic diseases become more prevalent (UN, 2017), the demand for healthcare services increases, placing a significant strain on current healthcare systems (Hilderink & Verschuuren, 2018). In recent years, innovative care models for care delivery, such as Virtual Care Centers (VCCs), have emerged (Liljamo et al., 2021; Leenen et al., 2024). VCCs are novel hospital services, which, by use of information technology, enable remote patient monitoring, e.g., telemonitoring patients with chronic diseases living in their homes. This is supported by remote nurse-led coaching and guidance to help patients manage their conditions.

Telemonitoring is a complex, nurse-led intervention. It involves digital data collection by the healthcare professional and or by patients' self-measurements from a distance, adequate interpretation by the healthcare professional, and feedback phone calls, video calls, emails, and or texts (Thomas et al., 2021; HRSA, 2022). Telemonitoring is effective in enhancing the self-management of chronic diseases and reducing the impact on their need for care (Sanchez-Ramirez et al., 2022; Andersen et al., 2023). Key factors for safe and successful telemonitoring include: a practice-ready nursing workforce, patients with adequate self-management skills and sufficient technology literacy, nurses' competence in interpreting data, and awareness of the evolving role of nurses as care shifts from in-person to remote delivery (Eze et al., 2023).

VCCs focus on remote care for chronic diseases such as heart failure and chronic lung diseases. VCCs can also be harnessed for other types of care delivery, such as consultations for in- and outpatients. Although VCCs are emerging in the Netherlands with national guidelines for telemonitoring of patients at home (Ministerie van Volksgezondheid Welzijn en Sport, 2021), there may be variation in organization and execution of care between the hospitals involved. Nurses in VCCs are pivotal in providing care remotely in multidisciplinary teams, thus, they should remain actively engaged with emerging technologies for care delivery services.

Currently, there is a growing understanding regarding the evolving roles and responsibilities of nurses in providing care through new VCCs in the Netherlands (Leenen et al., 2024). However, existing studies have yet to establish connections with relevant educational frameworks, limiting the ability of modern nursing

programs to prepare students for their future roles and responsibilities in virtual care. There is an urgent need to define the roles and responsibilities of nurses in virtual care to enhance practice-readiness for safe and effective care delivery. Without role clarity, nursing education and clinical preparation may risk falling behind the expansion of virtual care, posing a risk to patient safety and care quality.

In the Dutch Bachelor of Nursing programs, the framework Canadian Medical Education for Specialists (CanMEDS) is used to outline roles and skills that nursing students need to develop during their training (see Figure 1) (Bouwes et al., 2023). The CanMEDS framework comprises the following key roles for nurses: 1. Nurse expert, 2. Communicator, 3. Collaborator, 4. Reflective Evidence Based Practice Professional, 5. Health Promoter, 6. Leader and 7. Quality Promoter. The first role focuses on high-quality patient-centered care provided by the nurse. The second role focuses on effectively communicating with patients, family, and other (healthcare) professionals. In the third role, the nurse collaborates with other healthcare team members for care delivery. Within the fourth role, nurses are expected to engage in continuous learning and apply evidence-based practice. The fifth role comprises health promotion amongst patients. The sixth role describes nurse leadership to improve healthcare (systems) and patient outcomes. Lastly, the seventh role involves the nurse ensuring the quality and safety of health care services (Bouwes et al., 2023). Nurses are expected to encompass these seven roles in their work, regardless of the healthcare setting.

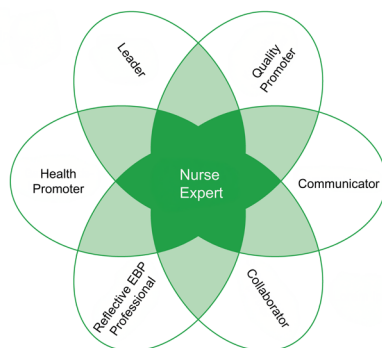


Figure 1: Core roles for nurses with a Bachelor's degree in the Netherlands

Source: (derived from Bouwes et al., 2023)

Our study aims to provide practice insights into the evolving roles, tasks, and responsibilities of nurses delivering care remotely from VCCs, mapped to the nursing roles within the CanMEDS framework. Through this qualitative study, we articulate the roles and responsibilities of nurses, which inform future development of competencies and guide the design of training programs in virtual care nursing.

2 Methods

2.1 Design

We adopted a qualitative content analysis, which allows the interpretation of meaning from textual data (Graneheim & Lundman, 2004). In our study, we conducted content analysis on interview data with nurses, which were collected as part of another larger study. Our study was conducted from November 2024 to January 2025. The Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist was utilized to guide and enhance the quality of our study. Our content analysis is a preliminary step in preparation for a Delphi study, in which we aim to further develop and validate the tasks, responsibilities, and roles of nurses in VCCs.

2.2 Setting

We conducted interviews with nurses who were working across three types* of hospitals providing telemonitoring: nurses from a general hospital, nurses from a top-clinical hospital, and nurses working in an academic hospital. Recruiting from multiple hospitals provided a more comprehensive understanding of various aspects of telemonitoring from VCCs, such as use of technology, diversity of chronic diseases involved in telemonitoring, and clinical expertise of the hospital (Faber et al., 2017).

Nurses were eligible to participate if they met the following criteria: 1) had a minimum of six months' experience in providing telemonitoring, ensuring adequate familiarity with the associated role and tasks, 2) held a registered nursing degree at European Qualification Framework level 4-6. Nursing students without a registered degree were excluded due to a possible lack of experience and scope of work.

**A general hospital provides a range of basic medical services, including surgery and medical care. A top-clinical hospital offers general medical care as well as more complex and specialized treatments, e.g., advanced cardiovascular care. An academic hospital delivers comprehensive medical services, including specialized and complex medical care, and also focuses on education, research, and training of professionals.*

2.3 Data collection

The (interview) data was collected between November 2023 and March 2024. An interview guide was developed and piloted, incorporating the following topics: motivation for working with telemonitoring, aspects concerning the training process, tasks and experiences with performing telemonitoring, patient-related aspects, and barriers and facilitators experienced. Face-to-face interviews were conducted by the first author on location in the hospitals. Written and verbal consent was obtained before commencing the interviews. The audio-recorded interviews (duration between 45-60 minutes) were then transcribed into textual data and subsequently uploaded to ATLAS.ti (*Scientific Software Development, Berlin, Germany*) as preparation for our qualitative content analysis.

2.4 Data analysis

Interviews were transcribed verbatim and content analyzed by two researchers (first author and colleague) to identify the evolving role and tasks of nurses providing care remotely for patients at home. The following steps were taken: preparation by selecting the unit of analysis (= fifteen written interview transcripts of nurses providing care within VVCs), and familiarization of interview data through multiple readings. Then the text data was organized by use of line-by-line, open coding, creating and assigning categories, and abstraction (Graneheim & Lundman, 2004). The researchers conducted their work independently; preliminary results were discussed collectively, and consensus was reached on the categories. This inductive approach enabled the exploration and identification of nursing tasks in the VCCs. Subsequently, a deductive approach was applied to map the identified tasks and roles to the seven CanMEDS core nursing roles. Nursing tasks were matched with corresponding nurses' responsibilities. These responsibilities were then aligned with the role descriptions in the CanMEDS framework. Finally, findings were presented to experts (nurse researchers) for validation and feedback.

2.5 Ethics

This study was approved by the Ethics Review Committee of the University of Applied Sciences Utrecht (HU) for a non-WMO application. The data was treated confidentially and processed anonymously, according to the European Union General Data Protection Regulation (EU GDPR).

3 Results

3.1 Participant characteristics

All fifteen participating nurses were female, and their age ranged from category [18-25 years] to category [56-65 years]. Three nurses worked at a general hospital, seven nurses in a top-clinical setting, and five in an academic hospital setting. Working experience as a graduate nurse ranged from 1.5 years to 40 years, and experiences with telemonitoring ranged from 1 year to a maximum of 3.5 years.

3.2 Nursing Tasks and Responsibilities Within Virtual Care Centers

Forty-six distinct nursing tasks were identified in the inductive analysis, which were subsequently categorized into key nurses' responsibilities for nurses providing care through VCCs. Their responsibilities varied from 'Remote care delivery', which includes tasks such as 'analyzing trends of e.g. vital signs from patients' self-administered measurements', to the responsibility of 'Collaboration and coordination' with tasks such as 'collaborating with other disciplines such as a nurse specialist and medical specialist', and 'collaborating with other type of professionals e.g. ICT professionals'. Most virtual care nursing tasks fell under the 'Remote care delivery' category. A summary of the responsibilities and related tasks is listed in Table 1.

3.3 Virtual Care by Nurses mapped to CanMEDS Nursing Roles

In the second part of the analysis, we mapped the identified nursing tasks and responsibilities to the seven core roles based on the CanMEDS framework. Table 1 displays the nurses' responsibilities, related tasks, and role(s).

Table 1: Nurses' Responsibilities and Tasks in Virtual Care Centers and Connection with Core Nursing Roles

Nurses' responsibility	Related task(s)	Roles ¹
Remote care delivery	- Collecting patients' self-administered measurements remotely - Analyzing and understanding data / trends without seeing patients live - Decision making and follow-up by use of nurse intervention(s) remotely - Evaluation of care with patients remotely - Reporting details in electronic patient files	Nurse Expert & Health Promotor
»An important task is to analyze incoming data and follow up. We aim to complete those tasks as quickly as possible...«		
Collaboration and coordination	- Working closely together with patients and healthcare professionals within and outside the hospital, such as a nurse specialist, medical specialist (e.g., GP). - Working closely together with types of professionals, e.g., in projects with ICT professionals, or when software crashes. - Coordinating patients' care remotely	Communicator & Collaborator
<i>»We work together with medical specialists like a cardiologist... We also work nicely together with the ICT department...« »You are working with a whole team, people from ehealth (ICT), medical specialists, and other people...«</i>		
Training and coaching	- Developing and providing (new) training programs, e.g., telemonitoring new disease groups - Coaching new colleagues on the job and supervising colleagues with specific telemonitoring-only tasks	Collaborator & Reflective EBP professional & Leader
<i>»...currently we are developing a new introduction program for new colleagues for new diseases...« »As a nurse, I train and supervise my colleagues who perform monitoring...«</i>		
Education	(Remotely) educating and instructing: e.g., patients and partner/family on how to use technology and how to perform self-measurements using technology;	Nurse Expert
<i>»... most of our patients are old... When they are included for monitoring, we have to help them install the monitoring application. Step by step, what button to click...« »In the beginning, patients may find it difficult to use, but after a few times, even the older patients succeed...«</i>		
Development	Initiate or participate in (new) or ongoing projects, e.g., new care pathways, new care protocols, new questionnaires in software/apps. E.g., conducting research, updating protocols, and evaluating new programs.	Reflective EBP Professional & Leader & Quality Promotor
<i>»As a nurse, we have a double role. We work as a project leader to develop new programs (monitoring new diseases)... »We evaluate protocols with our colleagues.... We also ask patients about their experiences with monitoring... «</i>		
Promotion	Promotion of the VCCs and the care services provided (telemonitoring) within and outside of the hospital, e.g., fairs or congresses.	Nurse Expert & Leader
<i>»...some medical specialists within our hospital have no clue what we are doing here (in VCC). So we sometimes provide tours in our VCC.«</i>		

¹ Based on existing CanMEDS nursing roles, displayed in Figure 1

All seven nursing roles are represented within the nursing practice in VCCs. Upon examining the 46 different tasks identified, the majority can be linked to the roles of Leader (9 responsibilities) and Collaborator (9 responsibilities). The fewest responsibilities (4) could be linked to the role of Quality Promoter.

4 Discussion

4.1 Key findings

We identified 46 distinct tasks for nurses working in VCCs, which could be categorized into six nurses' responsibilities. All seven Dutch nursing core roles are addressed in providing care services for patients through a VCC. Some of the responsibilities and tasks identified, are comparable to those of nurses providing care to patients in regular (hospital) wards. These include care delivery, in collaboration with patients and other disciplines, and coordination of care. However, a key distinction is that care within VCCs is provided remotely. No physical examination can be performed by nurses, and they rely solely on data for clinical interpretation.

Furthermore, within the responsibility of 'Promotion', several new nursing tasks were identified, such as advocating for the new VCCs, and promoting the new innovative care delivery services available. Promotion of VCCs by nurses is essential, as it improves healthcare accessibility by expanding remote care services for (new) patients at home. Thereby freeing up hospital beds for those patients requiring in-person treatment.

Additionally, the number and type of tasks associated with the various nurses' roles (Table 1) may indicate differing priorities. For example, the Nurse Expert role encompasses more tasks compared to the Health Promotor role. Whether this indicates a deliberate prioritization or a lesser focus on health promotion tasks is unclear.

4.2 Comparison with prior research

Our study findings align with the results described by Leenen et al. (2024) on nurses working in VCCs, particularly the evolving role of nurses in providing remote care. For instance, nurses in VCCs are increasingly involved in the development of new

care pathways and new protocols, and they play a more significant role in acquiring and promoting the new care services of the VCCs (nurse responsibility ‘Promotion’). Herein, the roles of Nurse Expert, Reflective EBP Professional, Leader, and Quality Promoter are more prominent among nurses in VCCs, suggesting that their focus extends beyond direct patient care provision. Compared to nurses in regular wards, they take on Leadership and Quality Promotion roles more often, which contributes to continuous improvement of (remote) health care services. This potentially supports optimization of patient outcomes, in line with evidence-based practices.

The theme ‘Cooperation with Other Healthcare Professionals’ (Leenen et al., 2024) aligns with our responsibility ‘Collaboration and coordination’, emphasizing the importance of working with a multidisciplinary care team. In our study, nurses highlighted the necessity of collaborating with various experts, including those from non-care and non-medical fields, such as (ICT) technicians. Leenen et al. (2024) also mentioned that the patient-professional relationship changed when providing care remotely. In our study, we did not explore that aspect in depth.

Furthermore, in our study, we found key responsibilities, including several tasks of nurses providing care remotely. Our findings differ somewhat from the six professional nursing activities as described in Tan et al. (2024), where entrustable professional activities for nurses were developed for teleconsultations for nursing home residents. Their study focuses on tasks that are required of nurses to facilitate teleconsultations with a remote physician, whilst ours focuses on nurses who remotely provide care to patients in their homes. Nevertheless, both studies can be used to develop virtual care curricula in nursing.

Additionally, van Houwelingen et al.’s (2016) study in home care, identified fourteen professional nursing activities and 52 aspects of required knowledge, skills, and attitudes. Although the study was conducted in the home care context, our findings are somewhat comparable. For instance, ‘Triaging incoming calls and alarms’, ‘Analyzing and interpreting incoming data’, and ‘Monitoring body functions and lifestyle’ align well with our nurses’ responsibility, ‘Remote care delivery’. This alignment underscores the relevance of these activities for nurses.

Finally, Jacob et al. (2025) outline a curriculum framework for higher education in telehealth for multidisciplinary healthcare providers. It consists of twelve domains, including ‘assessment, diagnosis, and treatment’, similar to our responsibility of ‘Remote care delivery’. In comparison, our findings offer more detailed insights, including the specific tasks of nurses providing remote care, while the domains in Jacob et al. (2025) offer a theoretical framework for curricula development. Together, these findings seem valuable for the development of virtual care nursing education programs.

4.3 Implications for Practice and Education

Within VCCs, there is an emphasis on responsibilities such as Remote care delivery, Development, and Promotion. Nurses who are new to VCCs are required to develop specialized knowledge and skills, such as using technology independently, guiding patients effectively through telehealth modalities, engaging in innovative projects, and also promoting VCCs. These types of knowledge and skills are not typically acquired through common care delivery services in a hospital setting.

Our results serve as a foundation for developing novel nursing education programs for both nursing students and graduate nurses working in settings outside VCCs. Responsibilities and tasks specific to virtual care, such as delivering care without physical contact, interpreting data trends, and collaborating with non-medical staff such as technicians, are not yet part of basic nursing training. With the implementation of new innovative care models such as VCCs, both nursing students and practicing nurses need to be equipped with additional knowledge and skills to meet emerging demands in virtual care.

Finally, the roles linked to the tasks and responsibilities identified in this study might not align with the expectations of both future nursing students and graduated nurses. It is well-documented that (digital) technology in healthcare influences nurses' professional identity (Knop et al., 2024). Without adequate focus on the professional identity, there is a risk that nursing students may face increased absenteeism or attrition during education and or later in nursing care practice (Rasmussen et al., 2021). It is crucial to provide future nursing students with opportunities to explore and determine if the evolving roles of nurses align with their expectations. We believe it is important for future nurses to be exposed to virtual nursing,

demonstrating that nurses in VCCs can still maintain close contact with patients and build strong professional relationships.

4.4 Future Research Into the Evolving Role of Nurses

Future research on the evolving role of nurses providing care through VCCs is appropriate. Currently, we are conducting a study focused on the broader experiences of nurses providing remote care through VCCs. Secondly, in a future study, we want to (further) develop and validate VCC nurses' roles, responsibilities, and tasks by using the framework of Entrustable Professional Activities (EPAs). EPAs are essential in nursing education, as they define key tasks or activities that nurses must be able to perform independently (Tan et al., 2024), e.g., the tasks and activities healthcare supervisors (experienced nurses) entrust their trainees (nursing students) with once they achieve adequate levels of competency (ten Cate et al., 2016).

Furthermore, it is important to investigate how our findings impact the professional identity of nurses. Innovations such as VCCs and the evolving role and tasks of nurses affect their professional identity in terms of self-perception, role, and context (Rasmussen et al., 2021). Without attention to professional identity, there is a risk of increased absenteeism or attrition, a potentially avoidable outcome given the current pressures on the healthcare system and increased demand for care.

4.5 Conclusion

Our study shows how the responsibilities and tasks of nurses providing care through VCCs are mapped to the core roles of nurses as outlined in the CanMEDS framework. Some responsibilities and related tasks differ from traditional ward-based care, which reflects the evolving nature of the nursing role. By highlighting emerging nursing practice in VCCs, our study offers a foundation to build future virtual care nursing education programs, which reflect the demand for modern healthcare.

References

- Allmér, H. (2018). *Servicescape for digital wellness services for the young elderly*. Åbo Akademi University Press, Turku, Finland
- Andersen, F.D., Trolle, C., Pedersen, A.R., Köpfli, M.L., Borgesen, S., Jensen, M.S., Hyldgaard, C. Effect of telemonitoring on readmissions for acute exacerbation of chronic obstructive pulmonary disease: A randomized clinical trial. *J Telemed Telecare*. 2023 Jan 23: 1357633X221150279. DOI: 10.1177/1357633X221150279
- Attig, C., Franke, T. (2020). Abandonment of personal quantification: a review and empirical study investigating reasons for wearable activity tracking attrition. *Computers in Human Behavior*, 102, 223-237 <https://doi.org/10.1016/j.chb.2019.08.025>
- Bouwes, A., Broekman, H., Dobber, J., Eisenberg, I., Hertog den, R., Rugers, A. (2023). Opleidingsprofiel Bachelor of Nursing 2030. Landelijk Overleg Opleidingen Verpleegkunde, oktober 2023
- Eze, C.E., Dorsch, M.P., Coe, A.B., Lester, C.A., Buis, L.R., Farris, K.B. (2023). Facilitators and barriers to blood pressure telemonitoring: A mixed-methods study. *Digit Health*. 2023 Jul 25;9:20552076231187585. doi: 10.1177/20552076231187585. PMID: 37529536; PMCID: PMC10387707
- Faber, S., van Geenhuizen, M., de Reuver, M. (2017). eHealth adoption factors in medical hospitals: A focus on the Netherlands. *Int J Med Inform*. 2017 Apr;100:77-89. DOI: 10.1016/j.ijmedinf.2017.01.009
- Graneheim, U.H., Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness, *Nurse Education Today*, Volume 24, Issue 2, 2004, Pages 105-112, ISSN 0260-6917, <https://doi.org/10.1016/j.nedt.2003.10.001>.
- Hilderink, H.B.M., Verschuuren, M. (2018). Public Health Foresight Report 2018 – A healthy prospect. Rapport#2018–0030. Rijksinstituut voor Volksgezondheid en Milieu. Bilthoven: 2018
- HRSA: Health Resources and Services Administration. (2022). Understanding telehealth. [Telehealth.hhs.gov](https://telehealth.hhs.gov/patients/understanding-telehealth). <https://telehealth.hhs.gov/patients/understanding-telehealth>
- Jacob, M.F.A., Fandim, J.V., Reis, F.J.J., Hartvigsen, J., Ferreira, P.H., Saragiotto, B.T. (2025). Defining core competencies for telehealth in healthcare higher education: A Delphi study. *Musculoskeletal Science and Practice*, Volume 75, 2025, 103244, ISSN 2468-7812, <https://doi.org/10.1016/j.msksp.2024.103244>
- Knop, M., Mueller, M., Kaiser, S., Rester, C. (2024). The impact of digital technology use on nurses' professional identity and relations of power: a literature review. *J Adv Nurs*. 2024 Nov;80(11):4346-4360. doi: 10.1111/jan.16178. Epub 2024 Apr 1. PMID: 38558440
- Leenen, J.P.L., Jordens, Y.J., Wegman, A., Heesink, L., Lenferink, A. (2024). Experiences of nurses working in novel virtual care centres in the Netherlands: a qualitative study. *BMC Digit Health* 2, 70 (2024). <https://doi.org/10.1186/s44247-024-00127-3>
- Liljamo, P., Wahlberg, J., Mikkonen, H., Reponen, J. (2021). A Digital Care Pathway to Access Healthcare Without Time and Place Restrictions. *Stud Health Technol Inform*. 2021;284:163–5. DOI: 10.3233/SHIT210690
- Ministerie van Volksgezondheid Welzijn en Sport. Inspectie Gezondheidszorg en Jeugd. (2021). Toetsingskader: Telemonitoring volwassenen thuis. Den Haag
- Rasmussen, P., Henderson, A., McCallum, J., Andrew, N. (2021). Professional identity in nursing: A mixed method research study, *Nurse Education in Practice*, Volume 52, 2021, 103039, ISSN 1471-5953, <https://doi.org/10.1016/j.nepr.2021.103039>
- Sanchez-Ramirez, D.C., Pol, M., Loewen, H., Choukou, M.A. (2022). Effect of telemonitoring and telerehabilitation on physical activity, exercise capacity, health-related quality of life and healthcare use in patients with chronic lung diseases or COVID-19: A scoping review. *J Telemed Telecare*. 2022 Aug 31:1357633X221122124 <https://doi.org/10.1177/1357633X221122124>

- Tan, A.J.Q., McKenna, L., Bramley, A., van Houwelingen, T., Tan, L.L.C., Lim, Y.J., Lau, S.T., Liaw, S.Y. (2024). Telemedicine entrustable professional activities for nurses in long-term care: A modified Delphi study, *Nurse Education Today*, Volume 140, 2024, 106264, ISSN 0260-6917, <https://doi.org/10.1016/j.nedt.2024.106264>
- Thomas, E.E., Taylor, M.L., Banbury, A., Snoswell, C.L., Haydon, H.M., Gallegos Rejas, V.M., Smith, A.C., Caffery, L.J. (2021). Factors influencing the effectiveness of remote patient monitoring interventions: a realist review. *BMJ Open*. 2021 Aug 25;11(8):e051844. doi: 10.1136/bmjopen-2021-051844. PMID: 34433611; PMCID: PMC8388293. <https://doi.org/10.1136/bmjopen-2021-051844>
- Ten Cate, O., Hart, D., Ankel, F., Busari, J., Englander, R., Glasgow, N., ... Snell, L. S. (2016). Entrustment decision making in clinical training. *Academic Medicine*, 91(2), 191–198 <https://doi.org/10.1097/acm.0000000000001044>
- United Nations. (2017). *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables*. Working Paper No. ESA/P/WP/248. Department of Economic and Social Affairs. Population Division. Geneva: 2017
- Van Houwelingen, C.T.M., Moerman, A.H., Ettema, R.G.A., Kort, H.S.M., ten Cate, O. (2016). Competencies required for nursing telehealth activities: A Delphi-study, *Nurse Education Today*, Volume 39, 2016, Pages 50-62, ISSN 0260-6917, <https://doi.org/10.1016/j.nedt.2015.12.025>

THE COST-BENEFIT ANALYSIS OF DIGITAL TIMESTAMPING IN MARITIME LOGISTICS: THE CASE OF AIKATIIETO

JUKKA HEIKKILÄ, MARIKKA HEIKKILÄ, MANJU KC DEUJA

University of Turku, School of Business, Turku, Finland
jups@utu.fi, marikka.heikkila@utu.fi, kc.manju@live.com

This study evaluates the costs and benefits of a digital timestamping service implemented in Finland for commercial maritime traffic. Using qualitative interviews and procurement document analysis, we examine its impact on various stakeholders, including ports, ship operators, and regulators. Our findings reveal that while the service's acquisition cost was relatively modest, the total cost across the stakeholder network was approximately ten times higher due to connected investments required to fully leverage the service. Benefits clustered around three primary areas: improved coordination among maritime stakeholders, enhanced data quality, and better system integration. Unforeseen consequences include a shift in balance between participants and ancillary actors. The study provides insights for cost-benefit analysis methodologies for public digital infrastructure investments by making positive network effects visible for multiple stakeholders.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.25](https://doi.org/10.18690/um.fov.4.2025.25)

ISBN
978-961-286-998-4

Keywords:

maritime logistics,
digital infrastructure,
cost-benefit analysis,
public investment,
network effects



University of Maribor Press

1 Introduction

Maritime logistics is a dynamic and interconnected domain involving a wide range of stakeholders whose activities must be closely coordinated. With the increasing adoption of digital technologies, the sector is seeing advancements in smart navigation, efficient port operations, and proactive infrastructure management (LVM, 2020; Tijan et al., 2021; Paulauskas et al., 2021; Heikkilä et al., 2022; 2024). One critical enabler of these improvements is the availability of reliable and precise Estimated Time of Arrival (ETA) data for vessels, which plays a central role in planning and operational efficiency (Elbert & Walter, 2014; Arbabkhah et al., 2024). However, conventional ETA records provided by ships are often inaccurate, limiting their usefulness for berth scheduling and real-time coordination at ports (Yoon et al., 2023).

Despite the clear need for improved ETA information, individual actors are often reluctant to invest, as the benefits appear too marginal at the individual level - resulting in a collective action problem. In Finland, this was addressed by Fintraffic Vessel Traffic Services Ltd (VTS), which launched the Aikatiето service to provide centralized, high-quality ETA estimates to Finnish ports.

Public investments in digital infrastructure, like Aikatiето, are difficult to evaluate using traditional cost-benefit analysis (CBA) methods. Benefits tend to be long-term, indirect, and spread across multiple stakeholders, often beyond the initial investors (Heikkilä et al., 2018; Korpela & Mäkitalo, 2008). In networked sectors such as maritime logistics, these limitations become particularly pronounced.

This paper examines Aikatiето as a case study to explore its implementation, costs, benefits, and the distribution of value across stakeholders. It addresses two key research questions:

1. What are the direct and indirect costs and benefits of the Aikatiето service, and how are they distributed among stakeholders?
2. What improvements are needed in public investment evaluation methods to better reflect the characteristics of networked digital infrastructure?

By analyzing this case, we aim to inform public authorities, researchers, and industry stakeholders on how to design, justify, and assess similar investments.

2 Public Investment Evaluation

Public investment evaluations, particularly in digital infrastructure, face growing criticism for relying on traditional methodologies like cost-benefit analysis (CBA), which often fail to capture the complex, dynamic, and networked nature of digital systems. Lau (2007) observes that CBA and similar tools typically emphasize short-term, quantifiable benefits, overlooking indirect, qualitative, and long-term public value outcomes that are critical for digital government initiatives. Hüging et al. (2014) argue that public officials often struggle with the monetization of intangible externalities, particularly when evaluating small-scale yet innovative urban infrastructure initiatives. Heikkilä et al. (2018) similarly argue that conventional evaluation models overlook network effects and dynamic feedback loops in digital ecosystems. Benefits in a digital infrastructure context tend to diffuse across many stakeholders making it hard for a single actor's CBA to reflect the collective and long-term value.

A deeper understanding of the underlying challenges is provided by classics of economic theory. Building on Coase (1937), Williamson (1985) proposed Transaction Costs Economics (TCE) detailing how governance structures - especially in high-uncertainty or asset-specific environments - can economize on transaction costs through tailored institutional arrangements; While simple, low-specificity transactions can be governed by markets through prices and competition, highly specialized investments and uncertain future contingencies require adaptive governance structures that facilitate renegotiation and trust-building. Hart's (1995) theory of incomplete contracts further reinforces the inadequacy of traditional evaluation tools. It asserts that contracts cannot specify all future outcomes in complex projects, particularly those involving innovation. Thus, governance models should allocate control rights and establish flexible decision-making frameworks to respond to unforeseen developments. Complementing this, Elinor Ostrom's (1990) research emphasizes the importance of adaptive rules and collective decision-making.

These insights are especially applicable in public investments in digital infrastructure, where rapid technological change and evolving policy landscapes make rigid, ex ante performance evaluations ineffective. Therefore, Frydinger et al. (2019) propose formal *relational contracting* to manage the uncertainty and complexity inherent in large-scale digital projects. Rather than specifying rigid outputs, these contracts focus on shared principles, collaborative processes, and mechanisms for renegotiation. Still, in the EU, public procurement is highly regulated with requirements for neutrality and transparency (VM, 2023). Traditional procurement procedures - like requiring fixed specifications and competitive tendering - can clash with the iterative, co-development practices often necessary for digital infrastructure projects. Also, Finland's processes for public information management investments follow government requirements, though these have been criticized for being vague and overly focused on data safety and security rather than architectural compliance with related services and supporting innovation. Recognizing this, recent EU reforms have introduced mechanisms like the Innovation Partnership and pre-commercial procurement (EC, 2021) to better support experimentation and long-term value creation. Nevertheless, challenges remain, including administrative burden, fragmented governance across EU and national levels, and limited flexibility for novel solutions (VM, 2023).

To understand these governance challenges in practice, we conducted an in-depth empirical investigation of the Aikatiето case. This case allows us to explore how costs and benefits of digital infrastructure investments are distributed across a network of stakeholders, and whether current evaluation and procurement models are adequate in capturing these dynamics.

3 Methodology

This case study (Yin, 2018) employed a mixed-methods approach to analyse the Aikatiето service, combining document analysis with qualitative interviews. We selected this approach to gain a comprehensive understanding of both the formal agreements and expectations surrounding the service and the experiences of stakeholders using the system. Our data collection process involved three primary methods:

1. **Document Analysis:** The team, consisting of all authors of this paper, examined non-public procurement documents, contracts, and proofs of concepts related to the Aikatiето service. This included reviewing the original tendering documentation, service specifications, and formal evaluation criteria. These documents and discussions with the person responsible for the tendering process provided insights into the formal expectations and requirements for the service.
2. **Semi-structured Interviews:** Team conducted site visits and face-to-face interviews with key stakeholders at five Finnish ports as well as the service vendor (Table 1). Ports were selected to represent a range of sizes, geographical locations, and use of Aikatiето. The interviews lasted between one and two hours and explored topics including service usage, benefits, costs, and implementation challenges.
3. **Follow-up Validation:** Interviewees validated the interview transcripts, ensuring accuracy and completeness of our analysis. The results of the analysis were presented to VTS.

Table 1: Interview Participants

Port	Interviewees	Port Size
HA	Harbor Master	Medium
UU	Managing Director, Traffic Manager	Small to
KO	Harbor Master, Development Manager	Large
HE	IT Manager, Traffic Manager	Large
HK	Deputy Manager	Large
Service provider	Service Manager, Director	N/A

All interviews were recorded, transcribed, and analysed using NVivo and Excel. Our analytical process was as follows: **Coding:** To ensure inter-rater reliability, two researchers coded the interview transcripts in NVivo to identify key themes related to costs, benefits, implementation challenges, and stakeholder relationships. After the initial round, the team compared coding results and discussed discrepancies to reach consensus (Barbour, 2014). This iterative process led to the refinement of our coding categories and ensured a consistent interpretation of key themes. **Thematic Analysis:** The team identified recurring themes and patterns across the interviews. **Cost-Benefit Mapping:** The team first conducted a cross-tabulation of stakeholder groups and associated benefits in Excel to visualize how value was distributed across the network. Based on this analysis, the first author then created a graphical

sociogram showing the expressed value of the associated stakeholders to each other.

To enhance the reliability and validity of our findings, we used methodological triangulation by combining document analysis with qualitative interviews. We validated interview transcripts through follow-up communications with participants and maintained transparency in our analytical process by documenting coding decisions systematically. Investigator triangulation - where multiple researchers independently analysed the data - helped minimize individual bias and enhanced the credibility of the findings (Archibald, 2016). This collaborative approach not only allowed confirmation of interpretations but also encouraged the integration of diverse perspectives, enriching the overall analysis (Cornish et al., 2013).

4 Findings

In Finland, VTS procured the Aikatiето service, which provides improved estimates of vessel arrival times by utilising machine learning algorithms that enhance ETA predictions.

4.1 The Aikatiето Service procurement process

The roots of Aikatiето originate from 2018 when the VTS examined the status of maritime logistics information sharing in Finland. **A working group** of stakeholders, established in 2019, identified potential for improvement in maritime information sharing, particularly regarding ETA information: 41,000 port calls from 2018 were analysed, finding that estimated time of arrival (ETA) deviated from actual arrival time (ATA) by an average of 33 minutes, with greater variations in tramp shipping compared to regular liner traffic. Following this analysis, **a pilot** was conducted at one Finnish port in 2020. Four individuals piloted a simple SMS-message service providing ETA data for two weeks. **Additional validation** was conducted in meetings with the port authority and the stevedoring company in 2021. Background **interviews** were conducted with maritime actors including icebreakers, pilots, and traffic coordinators to establish expected impacts (Pitkänen et al., 2021).

After successful piloting, VTS proceeded to a **public tendering** phase in 2021. An international tendering was conducted in two rounds, with eight competing offers in the final round. The ETA and timestamp data was to be redistributed by VTS at no additional cost to Finnish public authorities and the National Maritime Single Window (MSW) environment, as well as to port operators and their partners, who rely on ETA data for scheduling and coordination. While certain technical and service requirements were mandatory and established the minimum eligibility threshold, the final selection was primarily driven by price competitiveness (70%). Qualitative criteria - such as proposed added functionalities, implementation approach, and service continuity - accounted for the remaining 30%.

The contract was awarded to a company, that had participated in the working group and had already worked several years on a machine learning based ETA predictions. They delivered the service within one month of being selected. This service aggregates data from multiple sources to provide improved estimates of vessel arrival and departure times. Data sources include AIS (Automatic Identification System), port call predictions and realization, meteorological predictions and observations, and other maritime information on environment, cargo and traffic patterns. It is designed to be accessed either through an API interface (to be integrated with current port call and logistics systems) or as a Port Activity service, also available on mobile. The service adheres to relevant standards including the Port Call Optimization Task Force's Port Information Manual, the IALA S-211 Port Call standard, and The International PortCDM Council recommendations.

4.2 Costs and Investments

The costs associated with Aikatiето fell into three distinct categories: a) *Out-of-pocket costs*: Direct investment by VTS plus ongoing operation and maintenance costs (approximately 45% of initial per annum) b) *Project management costs*: Transaction costs of managing change across the stakeholder network, estimated at approximately €70,000 including staff time, travel, and preparation costs. c) *Connected investments*: The most significant category, representing vested investments made by stakeholders to fully leverage the service. These varied widely across ports (from €70,000 to €462,000) and were approximately ten times the direct investment in the service itself. This substantial ratio highlights the networked nature of benefits and

the need for complementary investments across the stakeholder ecosystem to fully realize the potential of such services.

4.3 Impacts of Aikatiето in ports

In Finland, there are approximately 50 ports, with 16 handling freight volumes exceeding 1 million tons. Based on feedback gathered by VTS in the spring of 2023, 12 of these major ports responded, and all of them were using Aikatiето and considered the service either useful or very useful.

However, our in-depth interviews revealed variations in how ports use the service highlighting the importance of flexibility in digital infrastructure design to accommodate different organizational needs and technical environments. It also demonstrates that the value of such services is perceived differently across organizations, influencing their willingness to invest in full integration. Table 2 shows the usage patterns, and an example of benefit charts can be found in the Appendix.

Table 2: Aikatiето Usage in Interviewed Ports

Port	Access	Integration Level	ETA Source	Improved by ETA data
Ha	API interface	High - Fully integrated into port's own system	System-generated dead reckoning ETA	Berth planning, operations scheduling, personnel alerts.
Uu	Port app (web & mobile)	Medium - alongside existing systems	Dead reckoning ETA, pilotage estimated	Berth planning, automated invoicing, transport scheduling, personnel alerts, geofencing.
Ko	Port app (web & mobile)	High - Used actively in existing systems	Dead reckoning ETA	Berth planning, Port-wide coordination, Staff coordination, real-time updates, optimized arrival times
He	Not using Aikatiето	None.	Port net system	Trucking companies complain that they don't get good enough ETA.
HK	Not using Aikatiето	A port activity system being developed	Marine Traffic AIS	Seeks a system to recommend optimized arrival times, not just ETA tracking

The two largest interviewed ports did not use Aikatiето, as it is not included in their current port activity systems. However, these systems are now outdated, and they are investing in new ones. For the others, the benefits of Aikatiето clustered around three primary areas: improved coordination, enhanced timestamp data quality, and system integration benefits.

Improved Coordination: Direct beneficiaries include ports, ground transport operators, vessels, and service providers. Ports gain advantages in better berth utilization planning and avoiding operational failures by sharing information: *“There is less confusion than before, which is why I do not get urgent phone calls as I used to before.”* Ground transport operators benefit from real-time coordination of their activities with the needs of arriving cargo and types of vessels. Ship operators, (i.e., vessels) can improve their route planning and coordination with smoother operations at harbour entry. Service providers such as bunkering and waste management companies have better situational awareness and improved planning capabilities. As one port noted: *“We see problems earlier, so we can solve them before they happen. If certain types of vessels arrive at the same time, we can warn them that ‘this is not a good time to come’ and suggest alternatives.”*

Indirect beneficiaries including pilots, icebreakers, tugboats, and linesmen can have significant improvements as well. Pilots gain from enhanced route planning and coordination capabilities. Icebreakers gain better situational awareness for their assistance operations. Tugboats benefit from improved service prioritization. Linesmen can have better work-life balance due to the improved predictability of vessel movements.

Improved Timestamp Data: The more precise timestamp information provides substantial benefits across stakeholder groups. Ports mentioned improved billing accuracy, better berth planning, enhanced vessel monitoring, and more optimized operations. They can also cut costs: *“with reliable ETA data, we can avoid calling in extra crew unnecessarily”*. Vessels benefit from smoother operations and reduced fuel consumption due to better planning capabilities. Agents’ situational awareness improves, and costs are optimised through reduced unnecessary waiting times. Linesmen gain better planning and resource allocation tools: *“We created an alert system so that when the vessel enters a specific area, personnel get a notification 1–2 hours before they need to be at the port.”*

System Integration: The integration of Aikatieto with existing systems provides several technological and operational benefits. Organizations reported more efficient billing processes through automation: *“Before, invoicing required a lot of manual labour. We had to check shipping information manually and input it into our invoicing system. Now, it’s automatically there.”* Berth planning becomes more data-driven and responsive. Information sharing across platforms improves significantly, leading to reduced manual data entry and fewer communication errors. The automated information exchange reduces the workload associated with monitoring vessel movements and updating stakeholders’ situational awareness.

4.4 Beneficiaries and their changing relationships

An important finding from our study was the complex network of beneficiaries and their relationships in the Aikatieto ecosystem. Figure 1 illustrates the key stakeholder groups and their connections, highlighting how Aikatieto has changed the relationships between parties by disintermediating some connections while concentrating or reinforcing others. Orange ovals illustrate that the new entrant Aikatieto and improved connection with the vessels was controversial among stakeholders despite the benefits, and while green ovals show that the ports and pilots were perceived positively thanks to more precise estimates and integration to the overall service network.

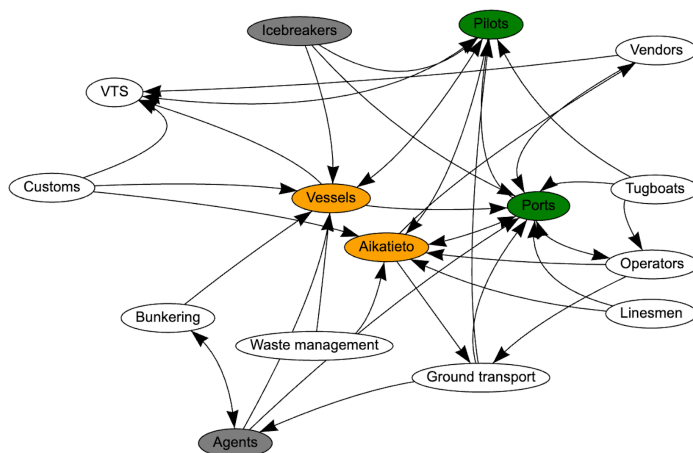


Figure 1: Beneficiaries Sociogram

The sociogram evidences the centrality of cooperation between vessels, pilots, ports, and the Aikatiето service. However, the agents' connection with other parties has become less critical and their relative position weakened (grey oval) when compared to the other parties. Ice-breaking services are still called for aid in the traditional manner – they were not part of Aikatiето design, which illustrates some planning integration problems in the operations.

5 Limitations of Current Public Investment and Procurement Approaches in Networked Digital Infrastructure

Our analysis identified several significant weaknesses in current public investment methods and procurement processes when applied to digital infrastructure projects in networked environments such as maritime logistics: The prioritization of digitalization projects is often based on *vague, quasi-economic criteria* that fail to capture the full range of benefits and costs across the stakeholder network. Comparison of alternatives becomes particularly difficult because typical methods cannot adequately identify *interdependencies* between necessary projects and investments. Investment models should explicitly account for complementary investments required by various stakeholders, not just the direct investment. Our finding that connected investments were approximately ten times the direct investment underscores the importance of this networked perspective.

Another limitation is the failure to consider *diffusion patterns* in benefit accrual. Current methods often assume immediate deployment and benefit realization, whereas in reality, benefits accrue gradually as adoption spreads across the network. This over-optimism can distort investment calculations. Additionally, *architectural priorities* such as compliance with initiatives like the European Maritime Single Window (MSW) and various interoperability requirements should be given sufficient weight in the prioritization process.

The current investment methods also inadequately account for different *development methodologies*. The increasing use of mixed development approaches - combining component-based, continuous integration and continuous delivery (CI/CD), agile, and waterfall systems development life cycles - requires different evaluation frameworks, yet these approaches are often treated uniformly in investment decisions. Furthermore, developing *open interfaces* for future uses is demanding and

requires cooperation between partners, stakeholders, and vendors, along with vested investments that may not be properly accounted for in initial evaluations, emphasizing the need for previous remedies.

Software procurement models add yet another layer of complexity. Whether the software is acquired as turnkey, Software as a Service (SaaS), or Platform as a Service (PaaS) requires different approaches to data governance and intellectual property agreements, as well as different ex-ante and ex-post risk management strategies and incentives between vendors and public agencies. These distinctions are rarely addressed adequately in current investment methods – even though they may make fundamental differences in intellectual property, updates and maintenance of the service.

Despite these challenges, our findings emphasize the growing importance of public investment in digital infrastructure for maritime logistics. Three factors illustrate this:

1. Capacity Preservation: Digitalization enables better utilization of existing capacity, preserving the status quo without requiring substantial physical infrastructure investment. As one port noted: *“With reliable ETA data we have been able to squeeze in more vessels in our schedule”*
2. Technology Affordability: Public investment has effectively utilized IT vendors' product development, maintaining affordable price levels that individual stakeholders might not have been able to achieve independently.
3. Coordination Benefits: The ability to leverage benefits in coordinating, planning, and operating ports across a network provides substantial value that would be difficult to achieve through uncoordinated individual investments.

The Aikatiето case also demonstrates how such digital infrastructure can catalyse broader shifts in business and operational models, extending value beyond the initial implementation context.

6 Conclusions

The Aikatiето case demonstrates both the potential and challenges of public investments in digital infrastructure for maritime logistics. While the direct investment was modest, the total investment across the stakeholder network was substantially higher -ten times the direct investment - demonstrating the networked nature of costs and benefits, and how traditional cost-benefit analyses fundamentally underestimate the true economic impact of such initiatives. This finding has significant implications for how such investments should be evaluated and managed. Second, the implementation reconfigured the maritime ecosystem's power dynamics and information flows—strengthening relationships between vessels, ports, and pilots while diminishing agents' centrality—highlighting how digital infrastructure can reshape industry structures beyond its intended operational improvements. Finally, it seems that conventional evaluation methods fail to adequately capture interdependencies, diffusion patterns, and architectural considerations that are essential for successful deployment in complex stakeholder networks.

These findings contribute to both transaction cost economics and incomplete contracts theory by demonstrating how digital infrastructure investments can reduce coordination costs across organizational boundaries while requiring flexible governance frameworks that accommodate evolving stakeholder needs. For policymakers and practitioners, our results suggest that evaluation models for digital maritime infrastructure should explicitly account for complementary investments, stakeholder network effects, and long-term architectural implications.

This study is not without limitations. It is based on a single case within the Finnish maritime logistics sector, which may limit generalizability. The qualitative nature of our data, while rich in detail, constrains the scope for quantitative validation. Furthermore, our analysis focused on one phase of implementation; future research could investigate the evolution of benefits and stakeholder dynamics over time. Comparative studies across different countries would help identify contextual factors that influence the results.

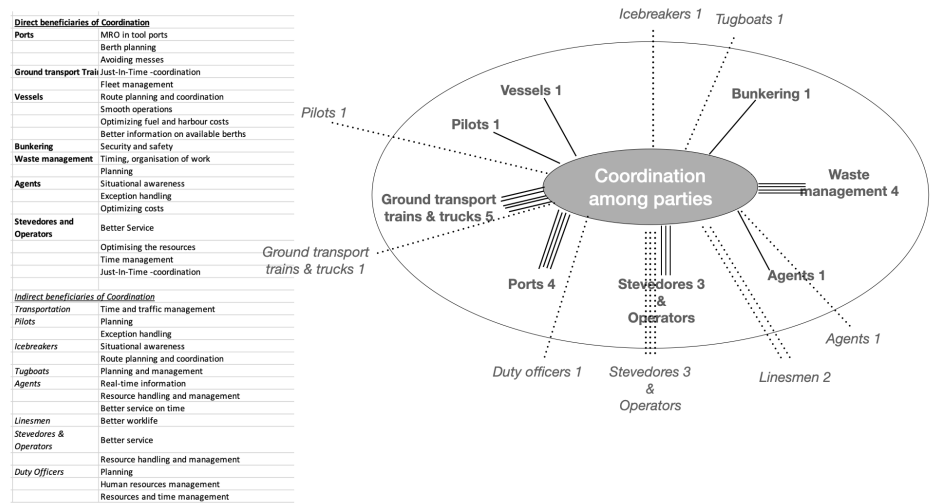
References

- Arbabkhah, H., Sedaghat, A., Kang, M. J., & Hamidi, M. (2024). Automatic identification system-based prediction of tanker and cargo estimated time of arrival in narrow waterways. *Journal of Marine Science and Engineering*, 12(2), 215. <https://doi.org/10.3390/jmse12020215>
- Archibald, M. M. (2016). Investigator triangulation. *Journal of Mixed Methods Research*, 10(3), 228-250. <https://doi.org/10.1177/1558689815570092>
- Barbour, R. S. (2014). Quality of data analysis. In: Uwe Flick, K. Metzler & W. Scott (Eds.), *The SAGE handbook of qualitative data analysis* (pp. 496-509). London: SAGE Publications Ltd. <https://search.datacite.org/works/10.4135/9781446282243.n34>
- Coase, R. H. (1937). The nature of the firm. *Economica*, 4(16), 386-405. <https://doi.org/10.1111/j.1468-0335.1937.tb00002.x>
- Cornish, F., Gillespie, A., & Zittoun, T. (2013). Collaborative analysis of qualitative data. In: U. Flick, K. Metzler & W. Scott (Eds.), *The SAGE handbook of qualitative data analysis* (pp. 79-93) https://search.credoreference.com/content/entry/sageukqda/collaborative_analysis_of_qualitative_data/0
- EC, (2021). Commission Notice: Guidance on Innovation Procurement. C(2021)/4320 final, 18.6.2021, 80 pages.
- Elbert, R. and Walter, F. (2014). Information flow along the maritime transport chain - a simulation-based approach to determine impacts of estimated time of arrival messages on the capacity utilization., 1795-1806. <https://doi.org/10.1109/wsc.2014.7020028>
- Frydlinger, D., Hart, O., & Vitasek, K. (2019). A new approach to contracts. *Harvard Business Review*, 97(5), 116-126.
- Hart, O. (1995). *Firms, contracts, and financial structure*. Oxford University Press.
- Heikkilä, J., et al. (2018). FEAR Governance Model for Public Service Development-From development initiative and goal setting to tendering and change management. University of Jyväskylä https://www.researchgate.net/publication/339055578_FEAR_Governance_Model_for_Public_Service_Development_From_development_initiative_and_goal_setting_to_tendering_and_change_management_FEAR_Project_15_March_2018_Version_21_Final_English_version_of_15_M.
- Heikkilä, M., Himmanen, H., Soininen, O., Sonninen, S., & Heikkilä, J. (2024). Navigating the future: developing smart fairways for enhanced maritime safety and efficiency. *Journal of Marine Science and Engineering*, 12(2), 324. <https://doi.org/10.3390/jmse12020324>
- Heikkilä, M., Saarni, J., & Saurama, A. (2022). Innovation in smart ports: Future directions of digitalization in container ports. *Journal of Marine Science and Engineering*, 10(12), 1925. <https://doi.org/10.3390/jmse10121925>
- Hüging, H., Glensor, K., & Lah, O. (2014). Need for a holistic assessment of urban mobility measures – review of existing methods and design of a simplified approach. *Transportation Research Procedia*, 4, 3-13. <https://doi.org/10.1016/j.trpro.2014.11.001>
- Korpela, J., & Mäkitalo, R. (2008). Julkishallinto murroksessa. Rohkeutta ja vauhtia muutokseen. Edita Prima, Helsinki (in Finnish).
- Lau, E. (2007). Electronic Government and the Drive for Growth and Equity. In Mayer-Schönberger and Lazer (eds.) *Governance and Information Technology: From Electronic Government to Information Government*. The MIT Press, London. pp. 39-58.
- LVM (2020). Logistiikan digitalisaatiostrategia, 5.10.2020, Liikenne- ja viestintäministeriön julkaisuja 2020:13. <https://julkaisut.valtioneuvosto.fi/handle/10024/162463>
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.
- Paulauskas, V., Filina-Dawidowicz, L., & Paulauskas, D. (2021). Ports Digitalization Level Evaluation. *Sensors*, 21(18), 6134. <https://doi.org/10.3390/s21186134>
- Pitkänen, J., Heikkilä, J., Pasanen, R., Kalatie, K., (2021). *Aikatielo-töryhmän työskentely ja rahtilaivojen aikatielönnön pilotointi 2019-2020: Raportti*. Traficom liikenne- ja viestintävirasto (in Finnish).

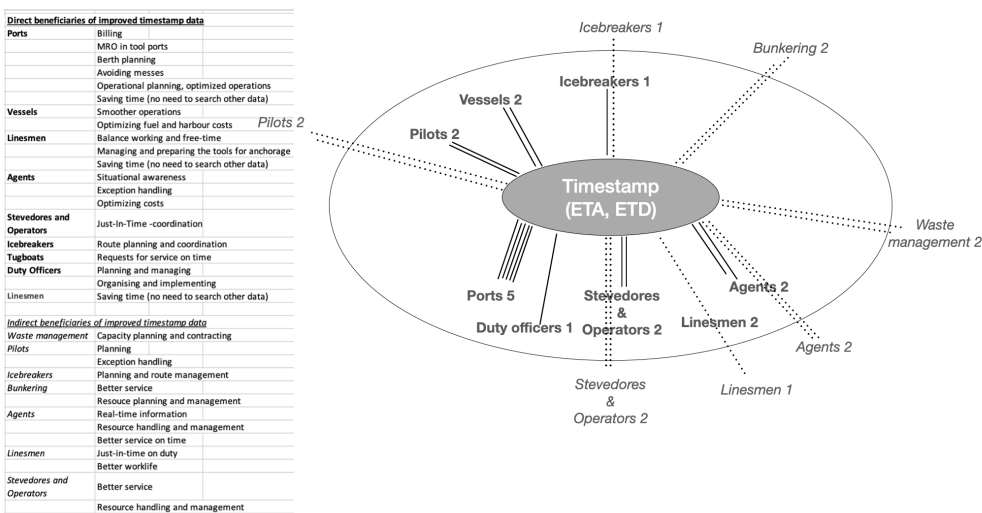
- Tijan, E., Jović, M., Aksentijević, S., & Pucihar, A. (2021). Digital transformation in the maritime transport sector. *Technological Forecasting and Social Change*, 170, 120879. <https://doi.org/10.1016/j.techfore.2021.120879>
- VM (2023). Julkisten Hankintojen Käsikirja, Valtiovarainministeriön julkaisuja 2023:60, 5.9.2023 <http://urn.fi/URN:ISBN:978-952-367-661-9>
- Williamson, O. E. (1985). *The economic institutions of capitalism: Firms, markets, relational contracting*. Free Press.
- Yin, R.K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). Thousand Oaks, CA: Sage
- Yoon, J., Kim, D., Yun, S., Kim, H., & Kim, S. (2023). Enhancing container vessel arrival time prediction through past voyage route modeling: a case study of Busan new port. *Journal of Marine Science and Engineering*, 11(6), 1234. <https://doi.org/10.3390/jmse11061234>

Appendix 1:

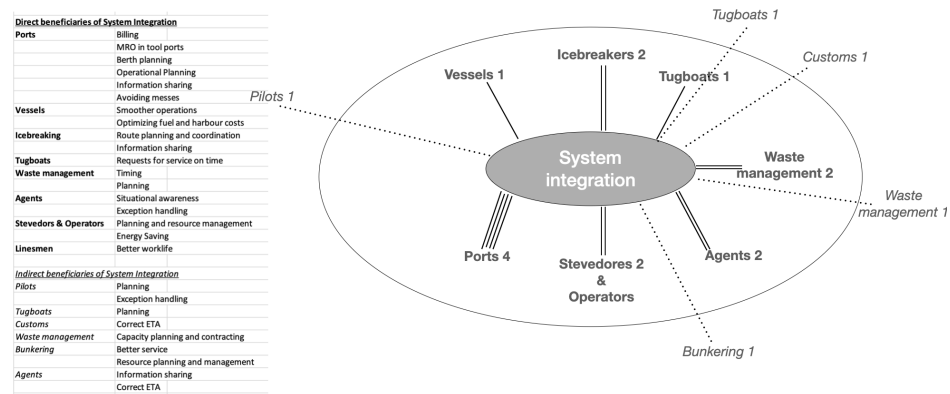
Direct and indirect beneficiaries of better coordination



Direct and indirect beneficiaries of improved timestamp data



Direct and indirect beneficiaries of increasing systems integration



GREEN(ING) OF TRANSPORT CHAINS: IMPLICATIONS TO BUSINESS MODELS

MARIKKA HEIKKILÄ, KRISTEL EDELMAN,
RIIKKA FRANZÉN

University of Turku, Turku School of Economics, Turku, Finland
marikka.heikkila@utu.fi, kristel.edelman@utu.fi, rimafr@utu.fi

This paper examines the transformation of business models in green transport through a case study involving cargo owners, land transport operators, and shipping companies. The study identifies four key themes driving green transport: commitment, collaboration, technology, and customer value, and examines how these themes shape emerging business models in the logistics sector. The findings reveal that leading companies are committing to better-than-regulation-standards. The study highlights the critical role of stakeholder collaboration in sharing infrastructure investments and risks, particularly in developing charging networks and renewable fuel supply chains. While market demand for green logistics is growing, especially among larger customers, the significant price premium for sustainable solutions remains a key challenge. The research indicates that business model transformation requires both technological innovation and new pricing strategies that balance cost with market acceptance. However, challenges such as inconsistent definitions of “green logistics” persist. This study contributes to the understanding of how logistics companies can effectively transition to sustainable operations while maintaining economic viability.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.26](https://doi.org/10.18690/um.fov.4.2025.26)

ISBN

978-961-286-998-4

Keywords:

business models,
transport chain,
green transition,
stakeholder collaboration,
logistics sector



University of Maribor Press

1 Introduction

The logistics sector is under increasing pressure to adopt practices that enable the measurement, analysis, and mitigation of the environmental impacts of logistics activities – referred to as green logistics (Blanco & Sheffi, 2024). Transportation is the 2nd largest source of global greenhouse gas emissions¹, contributing about 15% of emissions. As climate goals gain focus, transport providers have a big role in decarbonizing supply chains. Regulations such as CSRD and Fuel EU maritime and increasing customer expectations urge to address climate change (Bendig et al., 2023) - all of which demand transformative changes to existing business models. The different transport actors do not necessarily perceive green transport in a similar way; cargo owners and logistics service providers' perspectives are seldom aligned, even though variation exist based on e.g. cargo owner's industries and the size of logistics service provider (Jazairy et al, 2021; Huge-Brodin et al, 2020). Green transport entails more than operational adjustments towards less emissions; it requires re-evaluation of how value is created, delivered, and captured across the supply chain, impacting business models of multiple parties (Bankel & Govik, 2024). Prior research has explored emissions reduction technologies (Shah et al., 2021), logistics optimization (Cheng et al., 2024), and regulatory compliance (Bendig et al., 2023). Business model innovations have been examined in electric trucks (Lind & Melander, 2023), the maritime transport (Jović et al., 2022), shared mobility (Turoń, 2022), urban mobility (Bellini et al., 2019), and taxi services (Saqib & Satar, 2021; Niemimaa et al., 2019; Heikkilä & Heikkilä, 2019). However, little research explores how greening transport affects business models across a transport chain, a key challenge in balancing sustainability and competitiveness.

This paper analyses *how companies in a transport chain provide greener delivery options and what implications it has on their business models*. Through a qualitative study, the analysis identifies four themes - Commitment & Compliance, Collaboration & Value Co-creation, Technology & Innovation, and Market Value & Demand - highlighting key components and business model implications in the transition to green transport. The paper is structured as follows: Chapter 2 reviews literature on business models and green transport. Chapter 3 outlines the research method, Chapter 4 presents the

¹ <https://www.statista.com/statistics/241756/proportion-of-energy-in-global-greenhouse-gas-emissions/>

results, and Chapter 5 discusses the findings, connecting these with reviewed literature and limitations of the research.

2 Literature review

2.1 Business models

A Business Model (BM) defines how an organization creates, delivers, and captures value for its customers (Teece, 2010; Osterwalder et al., 2005). From an architectural perspective, it outlines the processes for delivering a value proposition and capturing the resulting benefits (Troisi et al., 2023). BMs consist of three interdependent dimensions: value proposition, value creation, and value capture (Bocken et al., 2014). The value proposition defines an organization's offering to its target market, addressing needs while considering economic, social, and environmental factors (Lüdeke-Freund et al., 2019; Geissdoerfer et al., 2017; Ranta et al., 2018). Value creation involves the activities, resources, and partnerships needed to develop and deliver the offering (Amit & Zott, 2001). Value capture refers to the mechanisms that enable a company to retain part of the value it creates, ensuring financial sustainability through pricing, cost structures, and revenue models (Teece, 2010; Schürtz et al., 2017). BMs are dynamic constructs that evolve in response to shifting market demands and emerging growth opportunities (Pohle & Chapman, 2006; Lindgardt et al., 2012; Bucherer et al., 2012; Marolt et al., 2018; Pucihar et al., 2019; Saebi et al., 2017). These adaptations can range from incremental modifications to comprehensive transformations of the entire BM (Saebi et al., 2017; Eriksson et al., 2021).

2.2 Green transport and business model transformations

A key driver of business model evolution in transport sector today is regulatory pressure, pushing companies to enhance sustainability, particularly by reducing carbon dioxide (CO₂) emissions (Dechezleprêtre et al., 2022). Regulations now extend beyond limiting emissions from direct production activities and energy consumption (i.e., Scope 1 and Scope 2 emissions) to addressing the entire value chain, including freight and supply chain operations (i.e., Scope 3 emissions). Compliance is becoming a baseline, with leading companies taking proactive actions in reducing their environmental impact (Bendig et al., 2023). The expanding scope

of regulations forces companies to rethink and reconfigure their business models (Salihi et al., 2024), traditionally built on low costs and reliability (Baştuğ & Yercan, 2021; Ellram et al., 2022). To stay competitive, they should offer greener transport options without compromising core performance. Environmental certifications have been suggested to show commitment to sustainability and to provide standardized frameworks for measuring and reporting environmental performance, though their credibility is debated (Chicca, 2024). Research indicates that sustainability is beginning to influence cargo owners' carrier selection (Davis-Sramek et al., 2020), bringing up a need for collaborative approaches where risk-sharing initiatives between carriers and cargo owners help overcome barriers to big sustainable investments. However, multitude of challenges exist related to business models, information sharing, and decision support systems and market structure. (Karam et al., 2021). Innovative technologies and digital solutions are essential for emissions reduction. Electrification, hydrogen, and biofuels reduce emissions (Folkson & Sapsford, 2022), while digital solutions improve emissions monitoring and logistics optimization (Noussan & Tagliapietra, 2020). Data transparency is now needed as stakeholders demand more detailed reporting throughout the transport chain (Haupt et al., 2024). Ultimately, market value and customer demand shape green business models. For example, green premium pricing has gained attention, though its implementation remains challenging in the cost-sensitive transport sector (Bertelè et al., 2024). Still, regulatory pressure and corporate sustainability commitments are fueling demand, creating possibilities for alternative green business models (Davis-Sramek et al., 2020). Trust-building is also becoming central in carrier selection. While trust traditionally takes time to develop (Dyer & Chu, 2000; Poppo et al., 2016), strong environmental performance can accelerate it in early stages of cargo owner-carrier relationships (Davis-Sramek et al., 2020). Green investments can serve as signals of trustworthiness (Blome & Paulraj, 2013) even before an initial relationship is established. Additionally, carriers who educate customers about sustainability practices demonstrate transparency and expertise, further enhancing trust (Bell et al., 2017). Therefore, trust between parties may build more rapidly than in traditional business relationships (Davis-Sramek et al., 2020), potentially giving carriers with green value propositions a competitive advantage.

3 Research method

Current research is part of the larger project “GreenConnect: Connecting green transition to competitive advantage in transport solutions.” The main focus of this paper is on the transport chain of green products considered as a specific case (Farquhar, 2012; Yin, 2018). The investigated transport in this case includes land and sea transport, covering upstream sea transportation and downstream multimodal transportation. The selected part of upstream transportation of materials included in the case is performed by one shipping company BulkerCo. The downstream transportation of the end product starts with land transport TruckCo from the cargo owner’s (ManuCo) premises, continues by shipping company RoroCo (184 km), and ends with land transport at the end customer, covering altogether 770 km. Timewise the transport takes about 17,8 hours, with an additional 1 hour of loading at the cargo owner’s premises and the land transport operator’s hub and waiting time at the port of departure.

To explore how companies in a transport chain aim to enhance green(er) delivery options within this specific case, a qualitative study (Creswell & Creswell Báez, 2021; Silverman, 2021; Hartley, 2004) was conducted from November 2023 to January 2025. The study consisted of 10 semi-structured interviews and four workshops with representatives of the companies of this specific case (see Appendix 1). These company representatives hold various leadership and specialist roles within their organisations, including positions such as directors, heads of units, managers, project coordinators, and specialists. With the first two workshops, the research team gained initial insights into the case and identified key aspects for further exploration (see Appendix 4). This phase supported the development of the interview framework, which was used in semi-structured interviews conducted in two stages. The first stage of the interviews concentrated on emission reduction goals, customer demand for emission data, and the role of data-driven and certified solutions for green(er) transport, as well as potential value and business benefits of offering green(er) transport solutions (see Appendix 2). The second interviewing stage focused on the impact of green transition on companies’ business models, pricing, and collaboration in logistics networks for facilitating the transition toward green(er) transport (see Appendix 3). All interviews were conducted online (via MS Teams or Zoom), recorded and transcribed, except for one interview, where notes were taken instead due to technical disruptions. The interviews lasted between 25 minutes and 1 hour,

workshops ranged from 1 to 3 hours. Collected qualitative data (104 pages transcribed text and 21 pages workshop notes) were analysed using thematic analysis method (Creswell & Creswell Báez, 2021; Lochmiller, 2021; Rapley, 2021) which included initial coding, grouping and merging the codes that enabled to identify themes and subthemes. This process was performed independently by three authors during the initial phase of coding, followed by mutual discussions to compare, (re)structure, and create meaningful themes and subthemes. The analysis of the collected data enabled to identify four key themes seen as drivers that shape the evolution of green business models in transport chains. These themes not only reflect the current state towards green(er) transport but also indicate how different elements of business models are being transformed to address environmental challenges.

4 Findings

Our analysis resulted in four key themes (commitment & compliance, technology & innovation, collaboration & value co-creation and market value & customer demand) driving the business model transformation of greener transport that will be discussed in following sections, together with the possible implications for the business models of the transport chain companies.

4.1 Commitment & compliance

ManuCo is committed to voluntary emission reduction to better serve its customers with high sustainability ambitions, and to enhance the transition towards net-zero economy. Through its emission targets, it is committed to reducing also the emissions of transporting materials and end products. The transport service providers are all committed to reducing emissions: shipping RoroCo in phase with the requirements of regulations, while shipping BulkerCo and TruckCo have set much tighter schedules of their own. *“[The company] has implemented initiatives like ISO certification and benchmarks... They emphasize the importance of meeting and exceeding regulatory requirements...”* (BulkerCo).

Adopting regulation such as ISO standards and validation of emissions and reduction targets support consistent and credible emissions reporting but also help communicate the commitment and proactive approach to decarbonisation. In

addition to own activities, certification systems and verification of emissions are required from other supply chain actors. These are perceived critical in enhancing transparency and allowing consistent and reliable reporting of emissions, supporting accountability and addressing concerns for greenwashing. *“We are asking transportation companies to provide emissions according to the 14083 ISO standard, which is the new one... This ensures structured and unified emission reporting, building trust and eliminating greenwashing”* (ManuCo).

4.2 Technology & Innovation

Companies in a current study are increasingly adopting digital solutions to enhance emissions tracking, transparency and reporting capabilities. For instance, ManuCo is testing a Digital Product Passport, which *“helps us make the whole supply chain more transparent and gives us possibilities to make it greener as well”*. This system enables precise emissions calculations per transport leg, integrating API data from transportation companies for improved reliability and accessibility. Additionally, ManuCo is implementing new Transportation Management Systems that embed emissions reporting requirements into procurement processes and tenders. In parallel, two out of three transport companies have developed customer-facing emissions portals, allowing customers to access trip-level emissions data, analyze historical trends, and model potential emissions reductions through alternative fuel scenarios.

Beyond tracking and reporting, companies are testing and implementing several emission reduction innovations across their operations. Alternative fuels, such as biofuels and electric power, are under exploration and investment, though current technical limitations hinder large-scale adoption. As TruckCo explained: *“Our electric vehicle pilot projects aim to gather knowledge and experience to shape our future business models and operational practices”*. BulkerCo has developed a smart fleet optimizer, which allows it *“to make both economically and environmentally sound decisions.”* Furthermore, BulkerCo has introduced virtual arrival, an operational practice where vessels adjust their speed and arrival times based on port availability rather than racing to port only to wait at anchor. Implemented profit-sharing contracts incentivize stakeholders to adopt this practice by equally distributing the resulting bunker fuel savings between parties. The demonstrated impact has been significant: *“We've adopted virtual arrival practices, saving 24% of emissions simply by implementing operational and technological changes.”*

Despite technological advancements, the research identified key implementation challenges. The most significant hurdle is the high capital investment required for transitioning to a greener fleet, which poses financial barriers to widespread adoption. Additionally, the pace of digitalization varies among the interviewed companies. While some have developed advanced digital emissions portals, others lag behind - RoroCo, for example, still relies on Excel for emissions reporting, highlighting inconsistencies in current reporting practices.

4.3 Collaboration & Value Co-Creation

The findings show that, to develop charging networks, as well as availability and reliability of renewable fuels - both essential for driving green transport - the collaboration among stakeholders is a prerequisite. Companies with a direct customer connection emphasise that customers expect end-to-end green solutions across the entire supply chain. This increases pressure on companies to collaborate, alongside regulatory requirements and their own emission reduction goals outlined in the previous section. *“We’re working with logistics and port operators to establish emission-free corridors because customers need a full chain solution, not just parts of it”* (ManuCo).

The way stakeholders collaborate is to establish risk-sharing initiatives where the contracts should be longer than one or two years. Longer contracts between stakeholders provide security for making big investments for green transport (e.g., investments related to renewable fuels). These long-term contracts are possible when in addition to strong partnerships, the companies that share the same values establish joint emission targets as part of a strategic collaboration benefiting both parties. Also, similar independent emission reduction targets between stakeholders can foster deeper partnerships, aligning business strategies with environmental goals. *“Some customers, like ManuCo, have clear emission reduction targets and collaborate closely with us to meet their sustainability goals”* (TruckCo).

The risk-sharing initiatives have made it possible to share financial benefits with customers, that is currently seen as an exception but might become more common in the near future. Companies who are taking proactive roles in emission reduction by anticipating future trends and customer needs (BulkCo and TruckCo), position themselves as leaders in green transport by leveraging financial incentives, such as subsidies or tax benefits, for investments. This positioning not only strengthens their

market profile but also encourages new collaboration partners to join, accelerating adoption for environmentally sustainable practices in its entire supply chain. Concerning land transportation, the price of electric trucks, as well as availability of charging infrastructure, are currently barriers to green transport. Customers who have set ambitious emission targets might be the ones who invest in charging infrastructure that the company could use. This way of collaborating with clients enables to tackle financial barriers to achieving greener transport.

4.4 Market Value & Customer Demand

Customer demand is already recognised as one of the drivers of stakeholder collaboration in delivering green transport across the entire supply chain. Moreover, when companies offer fossil-free products, customers expect the transportation to be green as well. However, the analysis indicates that when a company positions itself as an active player proactively driving the green transition, it seeks to shape customer demand by communicating its economic sustainability goals before customers begin requesting this information. *“We have tried to train and educate [customers] before they even know they need this information”* (BulkerCo).

Trust is the basis for creating and sustaining relationships with customers, especially with those who prioritise environmental sustainability. By communicating openly on environmental targets, companies can attract customers who share similar goals and are interested in collaborating which is easier with forerunner customers. Transparency, through open communication and data-sharing for assessing the environmental impacts of transportation, enables companies to build or deepen trusting relationships with clients. However, another approach to facilitating green transport would be to distribute the costs across all customers. In this model, which is currently in an idea phase and not implemented, those contributing a small premium would be allocated green transport options, which could encourage broader adoption. This would benefit the entire supply chain but is particularly seen as a driver for shipping company RoroCo involved in the study, as they rely on stable demand and long-term commitments to justify investments in greener technologies and infrastructure.

4.5 Business model implications

The biggest uncertainties the transport chain companies face are the choice and cost of energy, the availability of infrastructure, technological readiness, the heavy investments required, evolving regulations, and customers' willingness to pay for green transport services. The business models should be able to tackle these challenges to ensure the successful transition toward a sustainable transport sector. When transforming the offerings towards greener transport services, the companies may change their **value proposition** in following ways:

1. From a Cost to Value Proposition: Emissions tracking can be transformed from a mere obligation to a potential competitive advantage and a new value proposition. As one respondent noted, *"We have received much praise from customers for this, as it's a value-added service not commonly offered"* (BulkerCo). However, while early adopters perceive this as a competitive advantage, scalability remains a concern.

2. Integrated digital-physical offerings: Bundling physical transportation services with digital emissions optimization capabilities increases customer switching costs and deepens relationships through continuous digital engagement. Companies investing in advanced digital capabilities can attract environmentally conscious customers and position themselves as sustainability leaders in their industry.

3. Carbon Market Integration: Interviewees also highlighted emerging business models where transport services integrate with carbon credits markets, creating opportunities for bundled transportation and carbon offset packages or even transportation services that generate carbon credits.

The **value capture** component of the business model may include:

4. Green Premium Opportunities: Advanced emissions tracking enables companies to quantify and market the "green premium" of lower-emission transportation options. For example, *"The portal has been continuously developed, introducing new features such as the ability to see, based on historical data, the emission reduction if, for example, 30% bio-blend were used in transportation"* (BulkerCo). This capability supports tiered pricing models based on emissions intensity, and outcome-based pricing tied to emissions reduction targets of the customer. However, inconsistent

definitions of “green logistics” by companies belonging to the same logistics chain, make it more difficult to provide end-to-end green transport to the end customer.

5. Balancing green costs: A share of the costs for providing greener transport services can be allocated to customers ordering non-green services, thus balancing the burden of and supporting the demand for green services. However, subsidizing green transport costs could mean losing customers, particularly in price-competitive segments.

Our results highlight the following options for **value creation** of transport chain companies:

6. Value-Sharing Partnerships: The virtual arrival service fosters shared benefits and risks. As described by BulkerCo: *“We've added clauses in contracts to share 50/50 the bunker savings from virtual arrival practices, which incentivizes clients to adopt greener practices.”* However, this approach also requires collaboration with port: *“We had to push for port authorities to recognize virtual arrival to ensure compliance with the 'first-come, first-serve' principle while still optimizing emissions.”*

7. Risk-sharing partnerships: Long-term sustainability investments require risk-sharing mechanisms, but aligning the interests of various stakeholders remains a challenge. Transport providers, cargo owners, financiers, public actors and regulatory bodies need to establish clear agreements to ensure fair risk and reward distribution.

5 Discussion

Leading firms – such as our case transport chain – view compliance as an opportunity to differentiate through higher standards and sustainability certifications (Bendig et al., 2023; Dechezleprêtre et al., 2022). While past research (Noussan & Tagliapietra, 2020) emphasizes regulatory transparency, our study shows it can serve as a value-added service and deepen customer relationships. Digital solutions (Cheng et al., 2024) play an increasing role in value proposition and value creation, e.g. by making emission reduction more visible. However, their effective use depends on a firm’s digital capabilities - those with higher digital maturity are better positioned to develop AI and data-driven business models, while less mature firms may struggle

to align digital tools with strategic transformation (Eriksson & Heikkilä, 2023; Sjödin et al., 2021). Collaboration is essential in the green transition. Our findings highlight the need for long-term contracts and shared sustainability targets to mitigate investment risks. Unlike studies that frame sustainability as a cost challenge (Ellram et al., 2022) we find that co-investments and risk-sharing accelerate green transport adoption. Market demand is crucial. Our results demonstrate that leading companies proactively shape the demand by educating customers and sharing value from technological and operational solutions (Blome & Paulraj, 2013; Bell et al., 2017). The potential for green premium pricing aligns with prior research on customers' limited willingness to pay for sustainability (Bertelè et al., 2024). However, we find that emissions tracking enables tiered or outcome-based pricing. A key challenge is the inconsistent definition of "green logistics" across logistics chains, which hinders fully transparent, end-to-end solutions. Several of the presented business models focus on increasing commitment: bundling digital and physical offerings enhances customer lock-in (Teece, 2010), while risk-sharing partnerships enable investments (Bocken et al., 2014). Cost-balancing distributes sustainability costs between eco-conscious and mainstream customers (Turoń, 2022). Solutions like virtual arrival demonstrate how operational adjustments can generate both economic and environmental value.

The current study has certain limitations, as the case study presented covered one transport chain involving several companies, most of which are forerunners regarding their approach to emission reduction. However, not all of them showed proactiveness or leadership in all respects, which allowed to recognize also differences within the case, such as in the utilization of digital solutions. Expanding the analysis to other cases (Hartley, 2004; Yin, 2018), covering e.g. companies with less ambitious emission reduction targets or transport chains reaching outside Europe, where regulation for emission reduction is less consistent and stringent, would allow cross-case comparisons, and thus, could be a fruitful direction for the future study. Furthermore, the recent simplifications of the EU sustainability rules (the sustainability omnibus) together with the US government decisions have increased uncertainty among businesses, which is likely to hamper the development of green logistics (Jazairy et al, 2021). Our data indicate that especially follower companies might hesitate in their investments decisions and adopting a sustainability focused strategy. Further research is needed to understand the effects these changes will have on the transport market, e.g. is there going to be a clearer division between

companies linked to ‘green transport market’ and the others forming ‘black transport market’?

6 Conclusions

This study contributes to the understanding of how green transport affects business models within transport chains. It identifies four key themes - commitment & compliance, technology & innovation, collaboration & value co-creation, and market value & customer demand - as drivers of business model transformation. The paper presents seven alternative business models, which may help to push forward the green transport options.

Acknowledgements

This research was conducted as part of the project “GreenConnect: Connecting green transition to competitive advantage in transport solutions” co-funded by Business Finland, various companies and the University of Turku. We would like to express our gratitude to the participants of the study presented in the current paper for their time, willingness, and invaluable contributions.

References

- Amit, R., & Zott, C. (2001). Value creation in e-business. *Strategic management journal*, 22(6-7), 493-520. doi.org/10.1002/smj.187
- Bankel, A., Govik, L. (2024). Networked business models on a nascent market for sustainable innovation. *Supply Chain Management: An International Journal*, 29(7), 97-111. doi.org/10.1108/SCM-10-2023-0496
- Baştuğ, S., & Yercan, F. (2021). An explanatory approach to assess resilience: An evaluation of competitive priorities for logistics organizations. *Transport Policy*, 103, 156-166. doi.org/10.1016/j.tranpol.2021.01.016
- Bell, S. J., Auh, S., & Eisingerich, A. B. (2017). Unraveling the customer education paradox: when, and how, should firms educate their customers?. *Journal of Service Research*, 20(3), 306-321. doi.org/10.1177/1094670517691847
- Bellini, F., Dulskaiia, I., Savastano, M., & D’Ascenzo, F. (2019). Business models innovation for sustainable urban mobility in small and medium-sized European cities. *Management & Marketing*, 14(3), 266-277. doi.org/10.2478/mmcks-2019-0019
- Bendig, D., Wagner, A., & Lau, K. (2023). Does it pay to be science-based green? The impact of science-based emission-reduction targets on corporate financial performance. *Journal of Industrial Ecology*, 27(1), 125-140. doi.org/10.1111/jiec.13341
- Bertelè, A., Pacca, M., & Weber, B. (2024). Making green logistics services profitable, Article McKinsey & Company, March 26, 2024, <https://www.mckinsey.com/industries/logistics/our-insights/making-green-logistics-services-profitable>
- Blanco, E. E., & Sheffi, Y. (2024). Green logistics. In: Bouchery, Y., Corbett, C., Fransoo, J. C., & Tan, T. (Eds.) *Sustainable supply chains: A research-based textbook on operations and strategy*. 2nd Ed. Pp. 101-141. Springer, Cham, Germany.

- Blome, C., & Paulraj, A. (2013). Ethical climate and purchasing social responsibility: A benevolence focus. *Journal of Business Ethics*, 116, 567-585. doi.org/10.1007/s10551-012-1481-5
- Bocken, N. M., 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. doi.org/10.1016/j.jclepro.2013.11.039
- Bucherer, E., Eisert, U., & Gassmann, O. (2012). Towards systematic business model innovation: lessons from product innovation management. *Creativity and innovation management*, 21(2), 183-198. doi.org/10.1111/j.1467-8691.2012.00637.x
- Cheng, H., Chaudhry, H. R., Kazi, I., & Umar, M. (2024). Unlocking greener supply chains: A global innovative perspective on the role of logistics performance in reducing ecological footprints. *Journal of Innovation & Knowledge*, 9(4), 100612. doi.org/10.1016/j.jik.2024.100612
- Chicca, F. (2024). The role of environmental certifications in fostering changes. In *Sustainability and Toxicity of Building Materials*. https://doi.org/10.1016/B978-0-323-98336-5.00003-0
- Creswell, John W. and Johanna Creswell Báez. (2021). *30 Essential Skills for the Qualitative Researcher*. 2nd ed. Sage, London, United Kingdom.
- Davis-Sramek, B., Robinson, J. L., Darby, J. L., & Thomas, R. W. (2020). Exploring the differential roles of environmental and social sustainability in carrier selection decisions. *International Journal of Production Economics*, 227, 107660. doi.org/10.1016/j.ijpe.2020.107660
- Dechezleprêtre, A., Gennaioli, C., Martin, R., Muûls, M., & Stoerk, T. (2022). Searching for carbon leaks in multinational companies. *Journal of Environmental Economics and Management*, 112, 102601. doi.org/10.1016/J.JEEM.2021.102601
- Dyer, J. H., & Chu, W. (2000). The determinants of trust in supplier-automaker relationships in the US, Japan and Korea. *Journal of International Business Studies*, 31(2), 259-285. doi.org/10.1057/palgrave.jibs.8490905
- Ellram, L. M., Tate, W. L., & Saunders, L. W. (2022). A legitimacy theory perspective on Scope 3 freight transportation emissions. *Journal of Business Logistics*, 43(4), 472-498. doi.org/10.1111/jbl.12299
- Eriksson, T., & Heikkilä, M. (2023). Capabilities for data-driven innovation in B2B industrial companies. *Industrial Marketing Management*, 111, 158-172. https://doi.org/10.1016/j.indmarman.2023.04.005
- Eriksson, T., Heikkilä, M., & Nummela, N. (2021). Crafting SME business model for international expansion with data-driven services. In *Business Models and Firm Internationalisation* (pp. 137-151). Routledge, New York, USA. doi.org/10.4324/9781003204268
- Farquhar, J. D. (2012). *Case study research for business*. Sage, London, United Kingdom.
- Folkson, R., & Sapsford, S. (2022). *Alternative Fuels and Advanced Vehicle Technologies for Improved Environmental Performance: Towards Zero Carbon Transportation*. In *Alternative Fuels and Advanced Vehicle Technologies for Improved Environmental Performance: Towards Zero Carbon Transportation*. Elsevier Ltd. doi.org/10.1016/C2020-0-02395-9
- Geissdoerfer, M., Savaget, P., Bocken, N. M., & Hultink, E. J. (2017). The Circular Economy—A new sustainability paradigm? *Journal of cleaner production*, 143, 757-768. doi.org/10.1016/j.jclepro.2016.12.048
- Hartley, J. (2004). Case study research. In C. Cassell, G. Symon (Eds.) *Case study research* (pp. 323-333). SAGE Publications Ltd, London, United Kingdom. doi.org/10.4135/9781446280119
- Haupt, J., Cerdas, F., & Herrmann, C. (2024). Derivation of requirements for life cycle assessment-related information to be integrated in digital battery passports. *Procedia CIRP*, 122, 300–305. doi.org/10.1016/J.PROCIR.2024.01.044
- Heikkilä J., & Heikkilä, M (2019). Global service platforms in local markets: case taxi services. In *proceedings of Bled eConference – Humanizing Technology for a Sustainable Society*, June 16 - 19, 2019, Bled, Slovenia.
- Huge-Brodin, M., Sweeney, E., & Evangelista, P. (2020). Environmental alignment between logistics service providers and shippers—a supply chain perspective. *The International Journal of Logistics Management*, 31(3), 575-605.

- Jazairy, A., von Haartman, R., & Björklund, M. (2021). Unravelling collaboration mechanisms for green logistics: the perspectives of shippers and logistics service providers. *International Journal of Physical Distribution & Logistics Management*, 51(4), 423-448.
- Jović, M., Tijan, E., Vidmar, D., & Pucihar, A. (2022). Factors of Digital Transformation in the Maritime Transport Sector. *Sustainability*, 14(15), 9776. doi.org/10.3390/su14159776
- Karam, A., Reinau, K. H., & Østergaard, C. R. (2021). Horizontal collaboration in the freight transport sector: barrier and decision-making frameworks. *European Transport Research Review*, 13(1), 53. doi.org/10.1186/s12544-021-00512-3
- Lind, F., & Melander, L. (2023). Networked business models for current and future road freight transport: Taking a truck manufacturer's perspective. *Technology Analysis & Strategic Management*, 35(2), 167-178. doi.org/10.1080/09537325.2021.1970738
- Lindgardt, Z., Reeves, M., Stalk, Jr, G., & Deimler, M. (2012). Business model innovation: When the game gets tough, change the game. In: Deimler, M., Lesser, R., Rhodes, D., & Sinha, J. (Eds.). *Own the future: 50 ways to win from The Boston Consulting Group*. Pp. 291-298. Wiley, Hoboken, USA. doi.org/10.1002/9781119204084.ch40
- Lochmiller, C. R. (2021) Conducting thematic analysis with qualitative data. *The Qualitative Report*, 26(6), 2029-2044.
- Lüdeke-Freund, F., Gold, S., & Bocken, N. M. (2019). A review and typology of circular economy business model patterns. *Journal of industrial ecology*, 23(1), 36-61. doi.org/10.1111/jiec.12763
- Marolt, M., Lenart, G., Kljajić Borštnar, M., Vidmar, D., & Pucihar, A. (2018). SMEs perspective on business model innovation. In *Proceedings of the 31st Bled eConference Digital Transformation – Meeting the Challenges*, Bled, Slovenia, 17–20 June 2018. aisel.aisnet.org/bled2018/2
- Niemimaa, M., Järveläinen, J. Heikkilä, M. & Heikkilä, J. (2019). Business Continuity of Business Models: Evaluating the Resilience of Business Models for Contingencies, *International Journal of Information Management*, 49, 208-216. doi.org/10.1016/j.ijinfomgt.2019.04.010
- Noussan, M., & Tagliapietra, S. (2020). The effect of digitalization in the energy consumption of passenger transport: An analysis of future scenarios for Europe. *Journal of Cleaner Production*, 258. doi.org/10.1016/j.jclepro.2020.120926
- Osterwalder, A., Pigneur, Y., & Tucci, C. L. (2005). Clarifying business models: Origins, present, and future of the concept. *Communications of the association for Information Systems*, 16(1), 1. doi.org/10.17705/1CAIS.01601
- Pohle, G., & Chapman, M. (2006). IBM's global CEO report 2006: business model innovation matters. *Strategy & leadership*, 34(5), 34-40. doi.org/10.1108/10878570610701531
- Poppo, L., Zhou, K. Z., & Li, J. J. (2016). When can you trust "trust"? Calculative trust, relational trust, and supplier performance. *Strategic Management Journal*, 37(4), 724-741. doi.org/10.1002/smj.2374
- Pucihar, A., Lenart, G., Kljajić Borštnar, M., Vidmar, D., & Marolt, M. (2019). Drivers and outcomes of business model innovation—Micro, small and medium-sized enterprises perspective. *Sustainability*, 11(2), 344. doi.org/10.3390/su11020344
- Ranta, V., Aarikka-Stenroos, L., & Mäkinen, S. J. (2018). Creating value in the circular economy: A structured multiple-case analysis of business models. *Journal of cleaner production*, 201, 988-1000. doi.org/10.1016/j.jclepro.2018.08.072
- Rapley, T. (2021). Some pragmatics of qualitative data analysis. In: Silverman, D. (Ed.). *Qualitative research*. 5th Ed. Pp 341-356. Sage, London, United Kingdom.
- Saebi, T., Lien, L., & Foss, N. J. (2017). What drives business model adaptation? The impact of opportunities, threats and strategic orientation. *Long range planning*, 50(5), 567-581. doi.org/10.1016/j.lrp.2016.06.006
- Saqib, N. and Satar, M.S. (2021). Exploring business model innovation for competitive advantage: a lesson from an emerging market. *International Journal of Innovation Science*, 13(4), 477-491. doi.org/10.1108/IJIS-05-2020-0072

- Salihi, A. A., Ibrahim, H., & Baharudin, D. M. (2024). Environmental governance as a driver of green innovation capacity and firm value creation. *Innovation and Green Development*, 3(2), 100110. doi.org/10.1016/J.IGD.2023.100110
- Schüritz, R., Seebacher, S., & Dörner, R. (2017). Capturing value from data: Revenue models for data-driven services. *Proceedings of the 50th Hawaii International Conference on System Sciences*. hdl.handle.net/10125/41810
- Sjödín, D., Parida, V., Palmié, M., & Wincent, J. (2021). How AI capabilities enable business model innovation: Scaling AI through co-evolutionary processes and feedback loops. *Journal of Business Research*, 134, 574-587. https://doi.org/10.1016/j.jbusres.2021.05.009
- Shah, K. J., Pan, S. Y., Lee, I., Kim, H., You, Z., Zheng, J. M., & Chiang, P. C. (2021). Green transportation for sustainability: Review of current barriers, strategies, and innovative technologies. *Journal of Cleaner Production*, 326, 129392. doi.org/10.1016/j.jclepro.2021.129392
- Silverman, D. (2021). Introducing qualitative research. In: Silverman, D. (Ed.). *Qualitative research*. 5th Ed. Pp 3-16. Sage, London, United Kingdom.
- Teece, D. J. (2010). Business models, business strategy and innovation. *Long range planning*, 43(2-3), 172-194. doi.org/10.1016/j.lrp.2009.07.003
- Troisi, O., Visvizi, A., & Grimaldi, M. (2023). Digitalizing business models in hospitality ecosystems: toward data-driven innovation. *European Journal of Innovation Management*, 26(7), 242-277. doi.org/10.1108/EJIM-09-2022-0540
- Turoń, K. (2022). Open innovation business model as an opportunity to enhance the development of sustainable shared mobility industry. *Journal of open innovation technology market and complexity*, 8(1), 37. doi.org/10.3390/joitmc8010037
- Yin, R. K. (2018). *Case study research and applications: design and methods*. 6th Ed. Sage, Los Angeles, USA.

Appendix 1: Data collection

	Nov 2023	Feb 2024	Feb–Mar 2024	Sep–Oct 2024	Oct 2024	Jan 2025
	1st WORKSHOP	2nd WORKSHOP	1nd INTERVIEW STAGE	2nd INTERVIEW STAGE	3rd WORKSHOP	4th WORKSHOP
Cargo Owner ManuCo	4	2	2	2	4	3
Shipping Company RoroCo	1	1	-	1	1	1
Land transport operator TruckCo	1	1	-	-	1	1
Shipping Company BulkerCo	-	-	3	3	-	1
PARTICIPANTS TOTAL	6	4	5	6	6	6

Appendix 2: Interview framework, 1st stage

The aim of the interview is to understand the requirements and benefits of different stakeholders in green logistics. Some questions vary whether the interview is conducted with service provider or freight owner (questions 5-7, indicated after the specific question in an interview framework).

1. Could you first briefly tell about your current responsibilities in your company?
2. What are your company's goals in emission reductions? Do you think the goals are achievable in the timeframe? Why/ why not?
3. What kinds of means are used to reduce CO₂ emissions?
4. How do you estimate that your company can achieve value added from emission reductions?
5. How much your customer currently wants to know about CO₂ emissions of transportation (service provider)?/What solutions are you being offered by the logistics companies on transportation emission (freight owner)?
6. Do you have a systematic process how you offer emission data to your customers? What feedback have you received from them and how does it impact the emission reduction practices in your company (service provider)/What are your thoughts on the need to be verified or certified by an external party (freight owner)?

7. Does the emission data have to be certified or verified by a third party? What other ways does your company consider in providing reliable data to your customers (service provider)?/What are the key decision-making criteria in buying logistics services (freight owner)?
8. What gains is your company expecting from offering/using green transportation solution/services?
9. What bottlenecks or barriers (internal or external) do you face in offering/using green solution/services?
10. What are your thoughts on the twin transition? What type of data-driven solutions does your company provide to support green logistic solutions for your clients?
11. Anything else that you think is relevant in this theme that was not yet discussed?

Appendix 3: Interview framework, 2nd stage

In this discussion, we are particularly interested in the changes that the green transition enables or necessitates in the overall logistics chain, as well as in the business practices within your company. This discussion will support us to prepare for the upcoming workshops, where we aim to build an understanding of what changes the adoption of green transition will bring in logistic chain processes and business models.

Theme: Recent changes in green transition

Our GreenConnect project started last year in August (2023). What have been the main changes since then in terms of green transition?

Please provide a brief overview of your company's interests/situations in relation to the green transition.

- What do you see as the main opportunities and challenges currently?
- Why do these opportunities and challenges arise and what is needed to solve them?

Theme: Business models and practices – company perspective

How do changes in the transport sector towards more environmentally friendly directions shape your company's business?

- Can you provide some concrete examples of how these changes will affect business models and practices in the future?
- What impact in these changes have on pricing? How does it change?

Theme: Logistic chain – network perspective

Who are the other main actors in the logistics chain that can support the green transition in the transport sector?

- What is needed from these actors to facilitate this transition?
- How do you evaluate the role of your own company in relation to these other actors in the logistics chain for enabling the green transition?

Should the upcoming workshop² be company-specific or involve several companies?

Online or face-to-face? Can you suggest persons from your company who would be suitable for attending the upcoming workshop?

² At the end of each interview, we asked questions to better prepare for planning the fourth workshop with the same stakeholders interviewed in the second stage.

Appendix 4: Goals of the workshops

	Nov 2023	Feb 2024	Oct 2024	Jan 2025
	1 st WORKSHOP	2 nd WORKSHOP	3 rd WORKSHOP	4 th WORKSHOP
Workshop goal	The aim of the workshop was to make a final decision on the transport chain, which was used as a specific case in a project and in all following workshops. During the workshop, the transport chain in its full length was established and discussed in detail, along with the modes of transport, its restrictions, and the goals.	The aim of the workshop was to continue working on the transport chain selected in a previous workshop as a real-life case. Before the workshop, a detailed description of the case was shared with participants through a digital collaborative tool. This included a comprehensive overview of the entire transport chain, with specific information on each phase – such as responsible actors, detailed descriptions of the transport phases, calculated emissions, data ownership, applicable standards, and reporting practices, as well as future needs.	The aim of the workshop was to gather comments from participating stakeholders on the calculated emissions of the selected transport chain, as well as two future emission reduction scenarios and their impact to business.	The aim of the workshop was to establish a deeper understanding on the net-zero emission target for the transport chain, along with its perspectives and visions. First part of the workshop focused on identifying the emission reduction targets and timelines of the stakeholders involved. The second part concentrated on green logistics, exploring its current and future value propositions.
Workshop format	Online	Online	Online	Face to face

DIGITALIZATION MEETS CIRCULAR ECONOMY

HANS-DIETER ZIMMERMANN

OST Eastern Switzerland University of Applied Sciences, School of Management,
St. Gallen, Switzerland
hansdieter.zimmermann@ost.ch

This paper examines the role of digitalization in the implementation of the circular economy (CE). It shows that the circular economy can be seen as a tool for sustainability and that digitalization is a key driver for its implementation. The links between the circular economy and digitalization are shown and the concept of the digital circular economy is presented based on a narrative literature analysis. The Smart CE framework and the Digital Function for Circular Economy (DF4CE) framework are presented. The DF4CE framework links digital functions to the 10 R strategies of the circular economy. It derives seven mechanisms for how digital functions can improve the circular economy, including recycling, reuse and reduction. The paper emphasizes that further empirical research and practical guidelines are needed for the further development of the digital circular economy.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.27](https://doi.org/10.18690/um.fov.4.2025.27)

ISBN

978-961-286-998-4

Keywords:

circular economy,
digitalization,
sustainability,
smart CE framework,
DF4CE framework



University of Maribor Press

1 Introduction and Motivation

Driven by global challenges such as climate change and resource scarcity, sustainability has become a critical megatrend that aims to balance current and future environmental, social, and economic needs, often guided by the UN Sustainable Development Goals (SDGs). (*THE 17 GOALS | Sustainable Development*, n.d.) The circular economy is a specific approach that supports this goal. It's a transformative model that replaces the linear "take-make-dispose" system by focusing on minimizing waste and maximizing resource use through closed loops, keeping materials in circulation longer. The circular economy aims to reduce environmental impact, increase resource efficiency and promote sustainable economic growth without depleting finite resources, thereby supporting ecological balance and social equity. In fact, the circular economy is seen as a tool for achieving the SDGs as stated, e.g., in the EU circular economy objectives. (Rodríguez-Antón et al., 2022) Essentially, the circular economy serves as a critical tool for achieving broader sustainability goals, including the SDGs, by conserving resources and protecting the environment for future generations. While sustainability is the goal, circular economy is an important means to achieve it, complementing other approaches such as energy efficiency.

Digitalization, a key megatrend, is the use of digital technologies to improve or transform activities. More broadly, digital transformation involves fundamental changes in organizational culture, business models, and value creation. Crucially, this transformation can strategically advance sustainability and circular economy goals, making digitalization a key driver of the circular economy. The importance of digitization to support the implementation of the circular economy has been clearly addressed by several academic authors and many other organizations, such as the World Economic Forum (WEF) or the OECD. (e.g. (Barteková & Börkey, 2022; (Jensen, 2022) Kristoffersen et al., 2020; Neligan et al., 2023) Thus, businesses will “create new opportunities for innovation, differentiation, synergies and jobs.”(Jensen, 2022) For example, (Neligan et al., 2023) show that in the manufacturing sector, companies with a strong digital focus are more likely to adopt new business models that promote resource efficiency; other examples include the textile and the agricultural industries. (Heim & Hopper, 2022)(Farace & Tarabella, 2024).

The term "Digitainability" links the megatrends of sustainability and digitalization, emphasizing often-overlooked synergies. (Lichtenthaler, 2021) Scarce existing literature focuses on strategic perspectives, proposing frameworks like the Digitainability Assessment Framework (DAF) or conceptual models for integration. (Gupta & Rhyner, 2022)

As the concept of circular economy can be understood as a tool for sustainability, it can be assumed that applying digitalization to the circular economy would result in much more concrete and practical instructions for action. Therefore, in this paper, we would like to address the following research question: How can digitalization, e.g., through tools like the digital product passport, support the implementation and effectiveness of circular economy strategies?

To provide a broad overview of current knowledge, we use a narrative literature review. This method synthesizes research to provide a comprehensive, interpretive summary of a topic. Unlike systematic reviews with strict methodologies, narrative reviews are more flexible and subjective, incorporating different studies and perspectives. (Paré et al., 2015, p. 185 pp; Tate et al., 2015) However, narrative reviews face challenges such as subjectivity and bias in study selection and interpretation. Their less transparent and reproducible methods can hinder verification and lead to incompleteness. They often lack critical appraisal of study quality, which limits the generalizability of conclusions and makes them less suitable for making definitive recommendations than systematic reviews. Comprehensive analysis of complex, multidisciplinary topics is also inherently difficult for any literature review.

The paper is organized as follows: After the introduction, we will elaborate on the concept of circular economy as well as the corresponding implementation strategies, thus setting the stage for the discussion of the role of digitalization in the implementation of the circular economy. Then the role of digital technologies as drivers of the circular economy will be elaborated, followed by the presentation of the digital circular economy. Some conclusions will be drawn.

2 Circular Economy

In the following we first introduce the circular economy concept in more detail before presenting implementation strategies for the circular economy.

2.1 The Circular Economy Concept

The long-dominant linear economy operates on a "take-make-dispose" principle: extract raw materials, make products, and discard them as waste. While this model drives growth, it depletes finite resources and generates significant pollution. (Luban et al., 2025)

In stark contrast, the circular economy represents a paradigm shift, advocating a "close the loop" system. The concept was first developed in the 1970s. (Stahel, 2016) While definitions vary, Kirchherr et al. provide an in-depth discussion of the various understandings and conceptualization of the term. According to the authors the circular economy is an economic model that uses innovative business strategies to replace the "end-of-life" concept. It focuses on minimizing waste and promoting the reuse, recycling and recovery of materials throughout production, distribution and consumption. This model works at the micro (products, businesses, consumers), meso (eco-industrial parks) and macro (cities, regions, nations) levels. Its overarching goal is sustainable development, which aims to improve environmental quality, achieve economic prosperity, and ensure social equity for present and future generations. (Kirchherr et al., 2017, p. 224)

The circular economy challenges the linear model by prioritizing minimal waste and maximum resource use. It aims to eliminate waste and pollution, keep products/materials in circulation through reuse, repair and remanufacturing, and regenerate natural systems. This requires a shift from over-consumption to responsible resource management through reduction, reuse and recycling. The goal is a regenerative system that minimizes environmental impact while promoting long-term economic prosperity - a critical step toward a sustainable and resilient future.

The circular economy aims to decouple economic value from resource consumption, ultimately reducing material flows and their environmental impacts, including greenhouse gas emissions. Prioritizing resource prevention and

minimizing the use of raw materials is therefore key, surpassing recycling in importance. (Hofer et al., 2025)

The transition from linear ("take-make-consume-discard") to circular models represents a significant shift in resource management and production, driven by environmental degradation, resource depletion, and sustainability requirements. The circular economy offers a regenerative model that focuses on resource efficiency, waste reduction, and closed-loop systems. (Dabija & Năstase, 2024) Some of the key drivers of the transition include: (Dabija & Năstase, 2024; Dennison et al., 2024; Supanut et al., 2024)

- Environmental concerns: The depletion of natural resources and the environmental impact of industrial activities have necessitated a shift to sustainable practices. The circular economy provides a framework for reducing waste, emissions, and resource extraction.
- Resource scarcity: The finite nature of natural resources has prompted industries to explore ways to extend the life cycle of materials and products. Circular economy principles such as reuse, recycling, and remanufacturing address this challenge by promoting resource conservation.
- Economic opportunities: The circular economy offers opportunities for innovation, cost savings and competitive advantage. Businesses can create new revenue streams through product-as-a-service models, sharing platforms, and resource recovery.

2.2 Implementation Strategies for the Circular Economy

Concepts such as sustainability, digitalization and circular economy are all rather high-level and offer few concrete instructions for action. Therefore, the concepts need to be concretized and operationalized. One of the most frequently mentioned concepts in the scholarly literature and among practitioners are the so-called "R-strategies". (Hänggi et al., 2025)

The 'R-strategies', also known as circular strategies, are a fundamental operationalization of the circular economy. They provide a practical framework for businesses and policymakers to transition from a linear "take-make-dispose" model

to a circular one. They form a central framework of the circular economy that aims to keep products and materials in a continuous cycle, minimize the consumption of natural resources and reduce waste production.

Although there are different numbers of these so-called R-strategies, in the literature most authors refer to Potting et al. who propose 10 different R-strategies. (Potting et al., 2017). Based on an extensive review of the literature, Kirchherr et al. also identified these 10 R strategies. (Kirchherr et al., 2017) So the number of "R's" can vary, but the underlying principle remains the same: to create a closed-loop system. Since the 10 R's are well described in the literature, we will only briefly summarize them below.

To provide a clear flow of the 10 R-strategies, imagine a descending ladder, where each step represents a strategy, with the most desirable at the top and the least at the bottom. Here's a flow text description: At the top of this hierarchy are strategies that aim to make consumption and production more sustainable from the outset. These include Avoiding non-essential products (**Refuse**), Rethinking the efficiency of use (**Rethink**), and reducing the use of resources (**Reduce**). These approaches aim to reduce the need for new products and resources from the outset. The next stage involves strategies to extend the life of products and components. These include **Reuse**, **Repair**, **Refurbish**, **Remanufacture** and **Repurpose** of products. These measures help keep products in use longer and reduce the need for new products. The final stage is recycling (**Recycle**) and recovering energy from waste (**Recover**).

However, not all 10 strategies are implemented to the same extent.

While considerable practical and scholarly attention has been focused on understanding circular consumer practices, including reuse, recycling, and repair, the comprehensive research conducted by (Zimmermann et al., 2024) reveals a notable deficiency in examining the importance of consumer behavior with respect to other central R-strategies, particularly refuse, rethink, and reduce (strategies closely related to consumer attitudes and mindful consumption) and refurbish and remanufacture (strategies that involve the reintegration of a product at the end of its life cycle to be used again in its original capacity). In their research, the authors also analyzed the barriers to further adoption of circular R strategies. In their recent book (Podleisek

et al., 2025) present several case studies of successful circular market offerings, with details on each retention option.

These approaches are designed to make effective use of materials that can no longer be used in their original state. Although strategies such as reuse, recycling, and repair are at the bottom of the hierarchy, they play an essential role in the transition to a circular economy by ensuring that waste products and materials are used as much as possible and returned to the production cycle.

Based on an also extensive study of real-world cases as well as a literature review (Hofer et al., 2025) propose a different framework. The so-called “Value Hill” concept is used to illustrate the product life cycle (PLC) in both a linear and a circular context. In the linear value hill, the focus is on maximizing short-term profits, with little attention paid to the long-term sustainability of the product and the resources used. The authors differentiate the pre-use, use, and post-use phase. In the circular value hill, value is also added in the pre-use phase, and the highest value is reached in the use phase. However, if the product declines in the post-use phase, the Rs aim to slow down or close this cycle and capture the added value of the product beyond its use phase by retaining all or part of the resources already used.

3 Digital Technologies as Drivers of the Circular Economy

Digitization is the process of converting analog information into a digital format. Essentially, it's taking something physical or analog and turning it into digital data. Digitalization is about using digitized information to improve or enable existing processes. It's about using digital technologies to make processes more efficient. Digital transformation is about fundamentally changing the way an organization operates by integrating digital technologies across all areas of the business. It's about creating new business models, cultures and customer experiences. Relevant digital concepts and technologies include the Internet of Things (IoT), smart factory, cyber-physical systems (CPS), big data analytics, cloud computing, artificial intelligence (AI), distributed ledger technology (DLT) and blockchain, digital twin, digital product passport (DPP), digital platforms, and digital data spaces. Digital technologies have the potential to serve as central facilitators of the circular economy by closely monitoring the path of products, components, and materials, while making the resulting data accessible for improved resource management and

informed decision-making at various stages of the industrial life cycle. (Hariyani et al., 2024; Kristoffersen et al., 2020)

Data plays a pivotal role in digital transformation by serving as the foundation upon which organizations can build more efficient, innovative, and competitive operations. Integrating data into business processes enables better decision-making, improved customer experiences, and streamlined operations. This transformation is driven by the ability to collect, analyze, and leverage vast amounts of data to gain insights and drive strategic initiatives. At its core, data is the raw material that fuels digital transformation. It provides the insights, intelligence, and direction needed to drive meaningful change.

In this context, based on a comprehensive literature study Kristoffersen et al. come up with the “Smart CE framework”. (Kristoffersen et al., 2020) The framework consists of three main components: data transformation layers, resource optimization capabilities, and an intermediate layer that connects these components and consists of data flow processes, as shown in Figure 1. The various components were brought together using a hierarchical structure as the primary organizing principle, where each successive level is dependent on its predecessors. Specifically, regarding the data transformation level, resources must be connected via an Internet of Things (IoT) sensor to facilitate the generation of data. This data can then be transformed into information by integrating it with additional data sources and providing contextual relevance, culminating in the achievement of wisdom.

Similarly, in terms of resource optimization capabilities, diagnostic analytics provides insight into the underlying causes of events while building on descriptive analytics that elucidate the actual events. Similarly, data flow processes initially collect and integrate data to increase the effectiveness of data analytics.

In their paper Berg & Wilts state that implementing circularity is primarily an information problem. Numerous elements play a significant role in generating information asymmetries and market opacities that support a “take-make-dispose” of the linear economy paradigm. (Berg & Wilts, 2018) Examples are, among others, lack of information about used products and secondary materials, lack of data regarding product and material compositions, lack of data standards affect the comparability of environmental costs. (Staab et al., 2022) Both, Staab et al. as well

as Berg & Wilts argue that when it comes to implementing the circular economy, digital data platforms can offer several benefits.

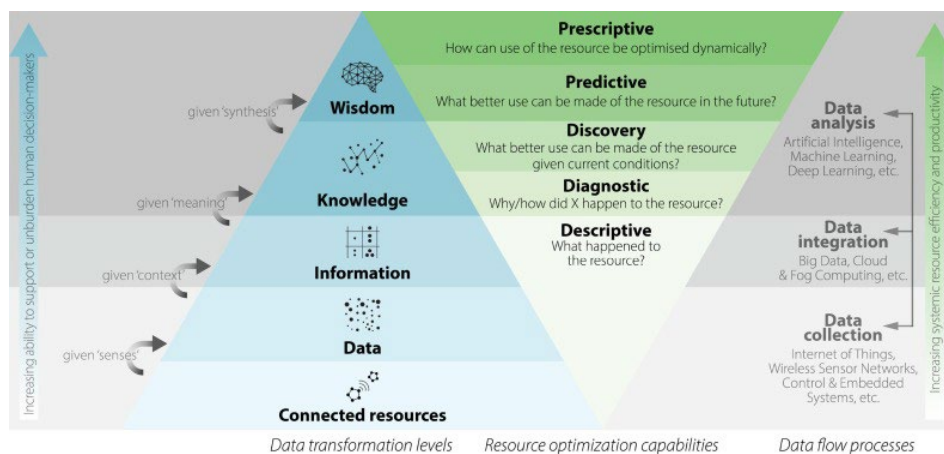


Figure 1: Smart Circular Economy Framework

Source: (Kristoffersen et al., 2020)

Data platforms play a critical role in advancing circular economy strategies by enabling the efficient sharing, management, and use of data across multiple sectors. These platforms support the transition to a circular economy by promoting sustainable practices such as reducing, reusing and recycling resources. They achieve this by providing a structured environment for data sharing and collaboration between stakeholders, thereby improving decision-making processes and fostering innovation. To this end, information-sharing platforms have been developed to facilitate the exchange of information such as the location of actors, the resources to be shared, and the costs of operating in such a network. In their recent research Möller et al. give an in-depth introduction and overview of data spaces and data space-enabled data ecosystems. (Möller et al., 2024)

In some cases, data sharing is now necessary to comply with regulations. In their recent paper Ströher et al. elaborate the case of adoption of automated data sharing for “carbon accounting” within and beyond organizations’ boundaries is imperative to comply with existing and upcoming regulations, to maintain competitiveness, and to foster sustainable practices for mitigating climate change. (Ströher et al., 2025)

However, real-world data platform implementation faces several challenges, ranging from jointly agreed data models to organizational orchestration mechanisms. (Acerbi et al., 2022; Blackburn et al., 2023)

Thus, data platforms are systems that collect, manage, and analyze data from various sources. They shall act as a central repository for storing data, making it easily accessible for users and applications. "Digital Product Passports" (DPPs) as a structured collection of product data throughout its entire lifecycle, rely on data platforms for their storage, management, and accessibility. DPPs use a unique identifier and electronic access to share product data, making them essential components of modern data platforms. DPPs are seen as concrete measure to implement circular economy strategies. (Plociennik et al., 2022; Ramesohl et al., 2022; Zhang & Seuring, 2024)

However, the implementation of Digital Product Passports (DPPs) in the context of the circular economy faces several challenges and barriers. These barriers are mainly related to technological, organizational and environmental factors that hinder the effective adoption and use of DPPs. Despite their potential to improve transparency, traceability and sustainability throughout the product life cycle, these challenges need to be addressed to fully realize the benefits of DPPs in promoting a circular economy. In line with the extensive research conducted by Chaudhuri et al., four primary barriers are identified: the need to formulate and articulate the business case for adopting DPPs; the increased demand for data; the establishment of data standards and ensuring interoperability between systems; the level of implementation effort required; and the need to ensure data security and integrity. (Chaudhuri et al., 2024)

In another study, the authors, while focusing on the building sector, identified challenges such as the lack of standardization of data templates and difficulties in collecting and tracking the data needed to create and maintain DPPs. Standardization would facilitate the implementation of passports, but aligning existing approaches and identifying overlaps remains a current challenge. (Honic et al., 2024) In addition to standardization and complex data management, Abedi et al. identified reluctance to share information between partners as a major obstacle to the adoption of DPPs in the circular economy. (Abedi et al., 2024)

In their recent work, Heeß et al. develop design principles that provide insight into how a Digital Product Passport should be designed to address the challenges and ensure the willingness of stakeholders to share their data, using the case of the low carbon hydrogen market. (Heeß et al., 2024)

A data-driven circular economy requires access to data throughout the product lifecycle. This requires breaking down data silos, facilitating data sharing across stakeholder boundaries, and fostering a culture of collaboration for collective benefit. Digitalization and data availability are key enablers, and a concrete measure to achieve this is the implementation of data platforms and digital product passports (DPPs).

4 The Digital Circular Economy

If we accept that digitalization plays a crucial role in the implementation of the circular economy, then the question arises as to how digital technologies or the digital functions identified by Kristofferson et al., for example - data collection, integration and analysis - can be applied specifically in the context of circular economy strategies. (Kristoffersen et al., 2020) In their study, the authors use examples from the literature to illustrate the relationship between digital technologies that support circular strategies in operational processes.

Based on an extensive study Liu et al. develop a framework of digital technologies for the circular economy. (Liu et al., 2022) To our knowledge, this is the only comprehensive study to explore the concept of the digital circular economy in detail.

In the framework the authors link digital functions to the 10 circular economy strategies (10 R) based on their study of the literature. In the study, the authors provide details on each feature, including examples drawn from the literature. The intersections of the x-axis (circular economy strategy) and the y-axis (digital functions) form different combinations of using a specific digital function to enhance a specific circular economy strategy. Using the relevant details from the study, the framework clarifies which digital technologies can be used to support which circular economy strategy.

Furthermore, the authors come up with seven so-called ‘mechanisms’. These mechanisms were derived by mapping the digital functions discussed in the papers examined to the corresponding circular strategies. The results of this mapping are presented in the “Digital Function for Circular Economy” (DF4CE) framework (Figure 2). The size of the circles in represents the number of mentions during the analysis and thus the frequency of discussion in previous studies. A higher frequency indicates a more comprehensive understanding of the implementation of digital functions for circular strategies. Thus, the seven derived mechanisms describe how digital functions can contribute to improving the circular economy.

It is evident that within the field of literature, these digital functionalities are predominantly employed for the purpose of *reduce*, succeeded by *rethink*, and subsequently, *recycle*. Again, the study provides several examples of the respective intersections based on the literature analysis. Furthermore, the authors claim that the seize of the circle can be interpreted as the respective maturity level of each specific digital function for each specific circular economy strategy. While this sounds understandable, one should be careful not to equate the number of use cases identified with the maturity of the solutions.

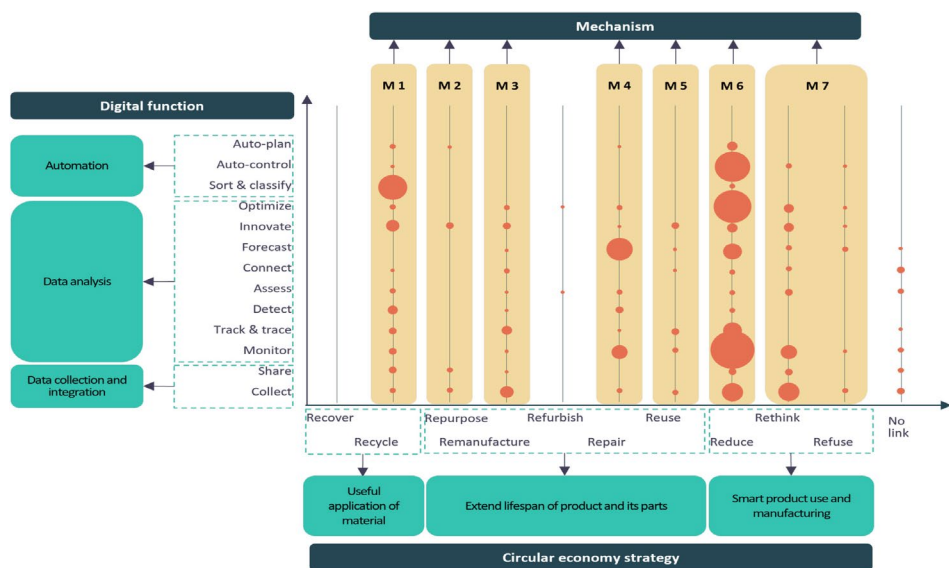


Figure 2: Structure of conceptual framework

Source: (Liu et al., 2022)

The research by Liu et al. certainly contributes to the theoretical understanding of how digitalization can support the implementation of circular economy strategies. Moreover, the paper provides some practical insights based on the examples identified in the literature. However, despite its valuable contribution to the field on the relationship between digital technologies and circular economy strategies, further research is needed to derive real-world guidelines to support the further development of the digital circular economy.

5 Conclusions

The aim of this research was to provide a comprehensive overview of existing knowledge on the relationship between how digitalization can support the implementation of the circular economy, based on a narrative literature review. It was found that the topic as such is quite complex. Not only are the concepts of digitalization and circular economy as such complex, but as a third dimension we also must consider the respective industry contexts.

However, this approach has several shortcomings. First, the methodological approach does not systematically cover the existing knowledge in the field. This would require a systematic literature review. Second, although some relevant findings could be developed, they are still insufficient. On the one hand, because of the non-systematic literature review, and on the other hand, from a practical point of view, the results need further investigation and further research.

From a practical point of view, it would be necessary not only to learn how to implement the R strategies in a specific business and industry context, but also to enable companies to prioritize the possible actions regarding solving existing challenges and/or deriving corresponding benefits. Therefore, a qualitative analysis such as the framework presented can only serve as a starting point. A quantitative assessment would be a necessary next step. Also, as the examples mentioned in Liu et al. are based on the analyzed literature, an analysis of real cases in different industries and thus empirical research would also be an important contribution. (Liu et al., 2022)

Another avenue for future research will be the role of integrated data platforms and DPPs in their respective contexts, as several challenges may arise, such as how to motivate actors in a given ecosystem to share data on a common platform. In addition, the issue of data security and privacy becomes relevant and important when collecting and sharing data.

Additional barriers and challenges to the adoption of digital technologies for circular economy implementation that should be further explored are identified in the research by Hariyani et al.: Costs of implementation, complexity and integration issues, data protection and security concerns, interoperability and standardization, risk management, resistance to change and cultural barriers, data quality and reliability, ethical and social implications, regulatory and political framework conditions, technological developments. (Hariyani et al., 2024) In addition, the authors, as well as Liu et al., based on their review of the literature, mention several future avenues of research in the area discussed in this paper. (Liu et al., 2022)

References

- Abedi, F., Saari, U. A., & Hakola, L. (2024, June). Implementation and Adoption of Digital Product Passports: A Systematic Literature Review. *Proceedings of the 30th ICE IEEE/ITMC Conference on Engineering, Technology, and Innovation: Digital Transformation on Engineering, Technology and Innovation, ICE 2024*. <https://doi.org/10.1109/ICE/ITMC61926.2024.10794320>
- Acerbi, F., Sassanelli, C., & Taisch, M. (2022). A conceptual data model promoting data-driven circular manufacturing. *Operations Management Research*, 15(3–4), 838–857. <https://doi.org/10.1007/S12063-022-00271-X>
- Barteková, E., & Börkey, P. (2022). *Digitalisation for the transition to a resource efficient and circular economy* (OECD Environment Working Papers, Vol. 192). <https://doi.org/10.1787/6F6D18E7-EN>
- Berg, H., & Wilts, H. (2018). Digital platforms as market places for the circular economy—requirements and challenges. *Sustainability Nexus Forum*, 27(1), 1–9. <https://doi.org/10.1007/S00550-018-0468-9>
- Blackburn, O., Ritala, P., & Keränen, J. (2023). Digital Platforms for the Circular Economy: Exploring Meta-Organizational Orchestration Mechanisms. *Organization & Environment*, 36(2), 253–281. <https://doi.org/10.1177/10860266221130717>
- Chaudhuri, A., Wahrens, B. V., Treiblmaier, H., & Jensen, S. F. (2024). Impact pathways: digital product passport for embedding circularity in electronics supply chains. *International Journal of Operations & Production Management, ahead-of-print*. <https://doi.org/10.1108/IJOPM-01-2024-0012>
- Dabija, D., & Năstase, C. (2024). Exploring the Evolution of Sustainability Paradigms: From Linear to Circular Economy Models. *Proceedings of the International Conference on Business Excellence*, 18(1), 754–769. <https://doi.org/10.2478/PICBE-2024-0066>
- Dennison, M. S., Kumar, M. B., & Jebabalan, S. K. (2024). Realization of circular economy principles in manufacturing: obstacles, advancements, and routes to achieve a sustainable industry transformation. *Discover Sustainability* 2024 5:1, 5(1), 1–32. <https://doi.org/10.1007/S43621-024-00689-2>

- Farace, B., & Tarabella, A. (2024). Exploring the role of digitalization as a driver for the adoption of circular economy principles in agrifood SMEs – an interpretive case study. *British Food Journal*, 126(1), 409–427. <https://doi.org/10.1108/BFJ-12-2022-1103>
- Gupta, S., & Rhyner, J. (2022). Mindful Application of Digitalization for Sustainable Development: The Digitainability Assessment Framework. *Sustainability (Switzerland)*, 14(5), 3114. <https://doi.org/10.3390/SU14053114>
- Hänggi, R., Podleisek, A., Luban, K., Hofer, N., & Podleisek, A. (2025). The Rapperswil Circular Economy Model – A Guide to Successful Implementation. In André Podleisek, Katharina Luban, & Roman Hänggi (Eds.), *Circular Economy: The Next Level of Company Success* (pp. 341–377). Springer, Cham. https://doi.org/10.1007/978-3-031-66406-9_19
- Hariyani, D., Hariyani, P., Mishra, S., & Kumar Sharma, M. (2024). Leveraging digital technologies for advancing circular economy practices and enhancing life cycle analysis: A systematic literature review. *Waste Management Bulletin*, 2(3), 69–83. <https://doi.org/10.1016/J.WMB.2024.06.007>
- Heeb, P., Rockstuhl, J., Körner, M. F., & Strüker, J. (2024). Enhancing trust in global supply chains: Conceptualizing Digital Product Passports for a low-carbon hydrogen market. *Electronic Markets*, 34(1), 1–20. <https://doi.org/https://doi.org/10.1007/S12525-024-00690-7>
- Heim, H., & Hopper, C. (2022). Dress code: the digital transformation of the circular fashion supply chain. *International Journal of Fashion Design, Technology and Education*, 15(2), 233–244. <https://doi.org/https://doi.org/10.1080/17543266.2021.2013956>
- Hofer, N., Podleisek, A., Hänggi, R., Luban, K., Hofer Zürich, N., Podleisek, A., & Hänggi, R. (2025). Approaches, Drivers and Enablers of Circular Economy - A Theoretical View. In André Podleisek, Katharina Luban, & Roman Hänggi (Eds.), *Circular Economy: The Next Level of Company Success* (pp. 21–61). Springer, Cham. https://doi.org/10.1007/978-3-031-66406-9_3
- Honic, M., Magalhães, P. M., Van Den Bosch, P., Magalhães, P. M., & Van, P. (2024). From Data Templates to Material Passports and Digital Product Passports. In Catherine De Wolf, Sultan Çetin, & Nancy M. P. Bocken (Eds.), *A Circular Built Environment in the Digital Age* (pp. 79–94). Springer, Cham. https://doi.org/10.1007/978-3-031-39675-5_5
- Jensen, H. H. (2022). *How digitalization can help build a circular economy ecosystem* | *World Economic Forum*. <https://www.weforum.org/agenda/2022/08/why-a-global-circular-economy-requires-a-digital-business-ecosystem/>
- Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling*, 127, 221–232. <https://doi.org/10.1016/j.resconrec.2017.09.005>
- Kristoffersen, E., Blomsma, F., Mikalef, P., & Li, J. (2020). The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies. *Journal of Business Research*, 120, 241–261. <https://doi.org/10.1016/J.JBUSRES.2020.07.044>
- Lichtenthaler, U. (2021). Digitainability: The Combined Effects of the Megatrends Digitalization and Sustainability. *Journal of Innovation Management*, 9(2), 64–80. https://doi.org/10.24840/2183-0606_009.002_0006
- Liu, Q., Trevisan, A. H., Yang, M., & Mascarenhas, J. (2022). A framework of digital technologies for the circular economy: Digital functions and mechanisms. *Business Strategy and the Environment*, 31(5), 2171–2192. <https://doi.org/10.1002/BSE.3015>
- Luban, K., Hofer, N., & Hänggi, R. (2025). Introduction to This Book. In André Podleisek, Katharina Luban, & Roman Hänggi (Eds.), *Circular Economy: The Next Level of Company Success* (pp. 1–15). Springer, Cham. https://doi.org/10.1007/978-3-031-66406-9_1
- Möller, F., Jussen, I., Springer, V., Gieß, A., Schweihoff, J. C., Gelhaar, J., Guggenberger, T., & Otto, B. (2024). Industrial data ecosystems and data spaces. *Electronic Markets*, 34(1), 1–17. <https://doi.org/https://doi.org/10.1007/S12525-024-00724-0>

- Neligan, A., Baumgartner, R. J., Geissdoerfer, M., & Schöggel, J. P. (2023). Circular disruption: Digitalisation as a driver of circular economy business models. *Business Strategy and the Environment*, 32(3), 1175–1188. <https://doi.org/10.1002/BSE.3100>
- 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>
- Plociennik, C., Pourjafarian, M., Saleh, S., Hagedorn, T., do Carmo Precci Lopes, A., Vogelgesang, M., Baehr, J., Kellerer, B., Jansen, M., Berg, H., Ruskowski, M., Schebek, L., & Ciroth, A. (2022). Requirements for a Digital Product Passport to Boost the Circular Economy. *Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft Fur Informatik (GI)*, P-326, 1485–1494. https://doi.org/10.18420/INF2022_127
- Podleisek, A., Luban, K., & Hänggi, R. (Eds.). (2025). *Circular Economy: The Next Level of Company Success*. Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-66406-9>
- Potting, J., Hekkert, M. P., Worrell, E., & Hanemaaijer, A. (2017). Circular Economy: Measuring Innovation in the Product Chain. In *Planbureau voor de Leefomgeving* (Issue 2544). PBL Publishers. <https://dspace.library.uu.nl/handle/1874/358310>
- Ramesohl, S., Berg, H., & Wirtz, J. (2022). *Circular Economy und Digitalisierung - Strategien für die digital-ökologische Industrietransformation*. Wuppertal Institut für Klima, Umwelt, Energie. <https://doi.org/10.48506/OPUS-7899>
- Rodríguez-Antón, J. M., Rubio-Andrada, L., Celemín-Pedroche, M. S., & Ruíz-Peñalver, S. M. (2022). From the circular economy to the sustainable development goals in the European Union: an empirical comparison. *International Environmental Agreements: Politics, Law and Economics*, 22(1), 67–95. <https://doi.org/10.1007/S10784-021-09553-4>
- Staab, P., Pietrón, D., & Hofmann, F. (2022). *Sustainable Digital Market Design: A Data-Based Approach to the Circular Economy*. <https://doi.org/10.14279/DEPOSITONCE-15014>
- Stahel, W. R. (2016). The circular economy. *Nature* 2016 531:7595, 531(7595), 435–438. <https://doi.org/10.1038/531435a>
- Ströher, T., Körner, M.-F., Paetzold, F., & Strüker, J. (2025). Bridging carbon data's organizational boundaries: toward automated data sharing in sustainable supply chains. *Electronic Markets*, 35(1), 33. <https://doi.org/10.1007/s12525-025-00779-7>
- Supanut, A., Maisak, R., & Ratchatakulpat, T. (2024). Circular Economy Strategies in Practice: A Qualitative Examination of Industry Adaptation and Innovation. *Revista de Gestão Social e Ambiental*, 18(3), e06723. <https://doi.org/10.24857/rgsa.v18n3-121>
- 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>
- THE 17 GOALS | Sustainable Development. (n.d.). Retrieved March 13, 2025, from <https://sdgs.un.org/goals>
- Zhang, A., & Seuring, S. (2024). Digital product passport for sustainable and circular supply chain management: a structured review of use cases. *International Journal of Logistics Research and Applications*, 27, 2513–2540. <https://doi.org/10.1080/13675567.2024.2374256>
- Zimmermann, R., Inês, A., Dalmarco, G., & Moreira, A. C. (2024). The role of consumers in the adoption of R-strategies: A review and research agenda. *Cleaner and Responsible Consumption*, 13, 100193. <https://doi.org/10.1016/J.CLRC.2024.100193>

DEVELOPING DIGIWISH: THE CO-CREATION OF A VISUAL STRENGTHS-BASED TRAINING TOOL FOR DIGITAL SKILLS

DIAN A. DE VRIES,^{1,2} ANDERS KROPVELD,¹
YURRIAN ROZENBERG,³ JANNIE GUO³

¹ HU University of Applied Sciences Utrecht, Organizations in Digital Transition,
Utrecht, the Netherlands
dian.devries@hu.nl, anders.kropveld@npuls.nl

² University of Amsterdam, Faculty of Social and Behavioral Sciences, Noord-Holland,
the Netherlands
dian.devries@hu.nl

³ YURR.Studio, Noord-Brabant, the Netherlands
yurr@yuur.nl, info@jannieguo.com

Digital skills have become a prerequisite for social inclusion in the era of digital transformation. However, many adults, particularly those with low reading literacy, struggle with digital skills. This paper describes the development and testing of DigiWish: a tool that coaches can use to help adults with low reading literacy identify personal and motivating digital learning goals and to work towards these goals. We followed an iterative, user-centered design methodology: An initial prototype of DigiWish was developed and evaluated with community coaches. Feedback from this co-creation session informed a second prototype: a visual and strengths-based tool. Subsequent field testing of the second prototype by coaches suggests DigiWish can help engage and motivate learners. Coaches and volunteers reported that the tool is useful, accessible, and motivating. The paper provides directions for future development and use of DigiWish and similar tools.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.28](https://doi.org/10.18690/um.fov.4.2025.28)

ISBN
978-961-286-998-4

Keywords:
digital skills,
digital inclusion,
reading literacy,
co-creation,
design-based research



University of Maribor Press

1 Introduction

Digital transformation is reshaping how we work, learn, and participate in society, making digital skills essential for all citizens (de Vries, Piotrowski, & de Vreese, 2025). Digital skills, and similar concepts such as digital literacy, have been defined and operationalized in different ways. In this paper, we refer to digital skills as the abilities required to use Information Society Technology for work, leisure, and communication (Carretero, Vuorikari, & Punie, 2017).

Governments have set ambitious targets regarding the digital skills of citizens. For example, the EU's Digital Decade aims for 80% of citizens to attain basic digital skills by 2030 (European Commission, 2025). However, a significant portion of the population currently lacks basic digital skills. In 2023, 44% of EU residents aged 16–74 did not have the basic digital skills required to participate in the digitalized society (Eurostat, 2023). Persons who have difficulty reading are particularly at risk of exclusion in the digital realm, as low reading literacy is a strong predictor of low digital skills (van Deursen & van Dijk, 2016). Unfortunately, digital skills training often fails to reach and involve people with low levels of reading literacy (Choudhary & Bansal, 2022). Yet, effective digital skills training tailored to people who have trouble reading is essential for ensuring that digital transformation is inclusive rather than widening social divides. The aim of this research is therefore to develop a tool that aids the learning of digital skills in a way that matches the wishes, preferences, and abilities of people who find reading difficult or burdensome. This paper describes the development and testing of this tool: DigiWish (DigiWens in Dutch).

2 Literature Review

Improving digital skills among low-literate adults is a multifaceted challenge that requires tailored tools and participatory development approaches (Choudhary & Bansal, 2022). One challenge is that conventional digital skills training and assessment tools often assume a baseline level of reading literacy. For example, the DigIQ online test in the Netherlands allows people (aged 10 and up) to evaluate their digital skills and receive personalized advice about how to improve (de Vries et al., 2025). However, text-based self-assessments like the DigIQ have trouble engaging adults who struggle with reading (de Vries et al., 2025). For people with

low reading literacy levels, the use of spoken communication, plain language and/or visual aids is a more useful alternative to text-based assessments (Kim & Lee, 2016).

Research indicates that effective digital skill programs for low-literate adults should incorporate tools that not only align with their reading literacy levels, but also with their learning preferences. For instance, the Digital Learning Ecology (DLE) framework emphasizes the importance of context, design, and motivation in creating mobile learning solutions that engage low-literate adults (Nedungadi et al., 2020). Similarly, a study on digital skill training for older adults highlights the effectiveness of personalized, volunteer-led programs that improve digital skills through one-on-one interactions, suggesting that tailored approaches can significantly enhance learning outcomes (Ngiam et al., 2022). Previous research thus emphasizes the need for interactive, personalized, and contextually relevant learning experiences.

Moreover, research shows that a participatory development approach is necessary for understanding what motivates and aids learners (Halvorsrud et al., 2021). Involving learners and trainers in the design process can make the tools developed more relevant and effective. For example, Smith and colleagues (2022) used co-design to create an Integrated Digital Literacy and Language Toolkit, which promotes meaningful learning among vulnerable migrant students in higher education. Furthermore, a study on mobile learning support tools for low-literacy adults found that involving users in the design process increased their motivation and confidence (Munteanu et al., 2014). We thus opted for a co-design approach to develop a highly visual tool for interactive, personalized, and contextually relevant learning in one-on-one interaction.

3 Methods

The project was a collaboration between a university research team and community partners, including a public library and adult education organizations. Our research followed a design science and participatory co-design methodology, consisting of two development iterations. The end-users targeted were adults who have difficulty reading (low literacy) and who likely have unmet digital learning needs. Because directly involving low-literate adults in early design can be challenging and demanding for them, we engaged intermediaries – namely, volunteers and coaches who work with the target group – in the co-creation of the prototypes.

The first prototype was developed together with a design company based on a literature review and explorative interviews. This prototype was then evaluated and built-upon in a three-hour co-creation session. Five experts (three youth social workers, one special education media coach, and one mental health support worker) participated in the co-creation workshop for Prototype 1. During this workshop, participants simulated using the first DigiWish prototype in pairs (one acting as coach, one as the learner) and provided feedback through guided discussion. We collected notes on their observations, experiences, and suggestions. This qualitative feedback was analyzed thematically to identify key issues and improvement opportunities. In essence, the workshop doubled as both a usability test and a design brainstorming session, consistent with participatory design methods.

Based on the feedback and ideas from the co-creation session, we developed a second prototype, together with the same design company. We then tested the second prototype in the field by distributing the Prototype 2 kit to volunteer digital coaches at two community organizations (a public library and a digital inclusion and employment coaching program). These practitioners used DigiWish with their clients during one-on-one coaching sessions. We conducted semi-structured interviews (in person and via email) with the coaches and volunteers. The interview questions covered how they used DigiWish, what the client's reaction was, and suggestions for improvement. The interviewees were people who had not joined the co-creation workshop.

For the co-creation session, written notes and audiotapes were reviewed and compared by two members of the research team and key points were grouped (e.g. comments about "language too difficult" were grouped under an Accessibility theme). For the field test interviews, we summarized written and verbal feedback from interviews and clustered comments to identify common sentiments (such as overall usefulness, specific liked features, or issues in certain contexts). We triangulated these insights with our observations from design meetings to formulate the results. The iterative nature of our methodology – design, test, redesign, and test – follows principles of agile development in educational technology.

Ethical procedures were observed: participants gave informed consent to participate in design sessions or interviews, and the focus was on tool utility rather than personal data of end-users. Ethical approval for these procedures was obtained at the University of Amsterdam.

4 Results

We present the findings in four parts: (1) the development of a first prototype; (2) feedback from the Prototype 1 testing and co-creation session; (3) the redesign resulting in prototype 2; and (4) evaluation of Prototype 2 in the field.

4.1 Prototype 1 Development

We began by (a) reviewing prior digital skills tools and literature on accessible design and (b) brainstorming with practitioners who work with digital skill learners with low reading literacy. Based on our findings, we concluded that the tool should be a **highly accessible (visual), motivating, personalized, non-threatening experience** that could be facilitated by a coach and **tailored to an individual's life context**. More specifically, the tool should be designed to help people articulate what they wish to learn to do online and motivate them to take first steps toward that goal. Experts also recommended that the tool should not be digital, but rather a **tangible conversational** tool, as people with low digital skills may avoid computers.

Incorporating these best practices from the literature and recommendations and experiences of practitioners, we conceptualized DigiWish Prototype 1 as a **tangible** card-based toolkit. To tap into **personal intrinsic motivation** and **tailor to an individual's life context**, the content of the cards was organized around seven life domains – drawing from the “Wheel of Life” framework (Swart, 2022) commonly used in coaching (domains included Health, Work/School, Family & Friends, Identity, Leisure, Community Participation, and Love/Romance). We created 40 prompt cards for **conversations**, each representing a topic within one of these domains (for example, a card for “Travel” under the domain of Leisure). On the front of each card, we printed a relatable picture depicting the topic, to serve as a **visual** “talking piece”. On the back, we listed a few simple questions that a coach

could ask to explore the person's interests or needs in that area, including apps or online tools related to the topic.

The intended use was that a coach and learner would go through the cards together: The pictures would spark discussion about what the learner wishes to do or improve in their life, and the coach could then help identify a digital wish that is **personally meaningful** to the learner (e.g. "I want to video chat with my grandchildren" or "I'd like to learn to use online maps.") The aim was to have a **non-threatening** tool that was not quizzing their current skills, but rather uncovering a motivation that could drive learning.

4.2 Co-Creation Workshop (Prototype 1)

We tested and discussed the first prototype in a co-creation workshop. Participants in this co-creation workshop responded favorably to the concept of using visual life-domain cards. They agreed that this approach "has potential." The images acted as icebreakers, helping practitioners start a conversation in a learner-centric way. Despite this promise, the session uncovered several critical issues: First, the textual prompts on the back of the cards were too complex and abstract for the target group. The coaches felt that questions like "What would you like to learn online in this area?" were not concrete enough to elicit specific answers from learners, especially those who may have trouble articulating needs. In some cases, the wording was too difficult. One coach remarked that: "Spoken language can also be a problem for persons who have a preference for working in a visual way," indicating that even if the learner isn't reading, the coach's reliance on reading complicated questions aloud could be a barrier. The participants suggested simplifying language drastically and relying even more on imagery or examples to guide conversation.

Second, participants highlighted a lack of a clear pathway to action. While Prototype 1 helped identify a goal (e.g. the person wants to be able to shop online safely), it did not help break that goal into actionable learning steps. One participant remarked about the questions on the back of the cards: "How does it feel if someone comes to you for help and you answer their question with another question? Why not a solution?" Participants expressed the need for bridging to actual learning opportunities, such as referencing appropriate courses, apps, or tutorials once a wish is identified.

Third, another key theme was the importance of a strengths-based approach. Co-creation attendees stressed the importance of beginning with what someone already can do, instead of emphasizing what they cannot. They shared that many low-literate adults have accumulated “splinter skills” – partial digital skills developed informally – and that acknowledging these can empower the learner. For example, one person might not use email but is proficient at sending WhatsApp voice messages; another can navigate by a transit app but not by general map apps. DigiWish should capitalize on these existing competencies, celebrating any skill the person already has, however small, to build confidence before introducing new steps.

Lastly, flexibility was identified as crucial. The group noted that different facilitators have different styles and different clients have unique preferences and existing skills. One participant gave an example: “One person plans a trip with Google Maps, another with the train company app. The tool must be flexible to accommodate multiple ways to reach similar goals.” Another participant said: “Wouldn’t it be great if someone else can draw and paste alongside you?” This emphasized that DigiWish should not enforce a single learning sequence but rather allow adaptability. In summary, in the words of one participant, the ideal tool would be “Something that is accessible, positive, not too difficult, and with not too much text.”



Figure 1: Testing the first prototype

4.3 Prototype 2 Development

Based on the co-creation findings, we developed Prototype 2, aiming to address the shortcomings of the initial version: simplify and reduce text; add mechanisms to go from wish to plan; use a format that highlights user strengths; and ensure the tool

can be adapted to various contexts. Prototype 2 was developed in collaboration with the same design partner, ensuring continuity and professional polish in the materials.

Table 1: Evolution of DigiWish prototypes through co-creation feedback.

Aspect	Prototype 1	Prototype 2
Purpose	Help coach discover the learner's digital wish (goal) by exploring life domains and interests. Emphasis on inspiration and motivation rather than on testing skills.	Help coach and learner define a wish and outline a learning path with concrete steps to achieve it. Emphasis on both motivation and planning next actions.
Format & Content	40 physical cards across 7 life domains (Wheel of Life categories). Each card has a picture on the front and several prompting questions on the back. The cards serve as visual conversation starters about what the person finds important or wants to improve in life.	Card-based toolkit including "wish" cards (with pictures, fewer words), new "step" cards to write down or select actions, and a "board" to provide insight into existing skills and steps for which actions learning and/or help is needed. Strong visual design with images and icons, significantly reduced text. Provides a flexible template to jointly create a sequence of learning steps toward the user's goal while also emphasizing existing strengths .
Facilitation	Coach-driven questioning using text prompts (learner does not need to read but must process spoken questions). The interaction is relatively open-ended: coach and learner discuss interests sparked by the cards, aiming to identify a concrete digital skill learning goal.	Collaborative : coach and learner both select cards. The learner can choose or draw images that resonate; coach uses simple prompts or examples if needed. Once a digital wish is chosen, they <i>co-create</i> a plan by ordering step cards or drawing a path. More hands-on engagement for the learner (e.g., physically placing cards) to accommodate creative and visual thinkers.
Feedback from Testing	Potential seen , but several issues: (a) Questions on cards were still too abstract and language-heavy for some – making the tool "too verbal" despite images. Some users might find spoken questions daunting if they prefer non-verbal learning. (b) Lack of clarity on translating wishes to actionable learning (c) More emphasis needed on acknowledging what learners <i>can already do</i> to boost confidence	Highly positive reception: coaches/volunteers found Prototype 2 very useful, helpful and motivating in guiding conversations and planning. The tool was flexible to different contexts. Overall, it addressed most issues from Prototype 1: less language-focused, concrete learning actions as output, and a strengths-based and motivating process. Minor suggestions were made and incorporated in a free download , so users can print the DigiWish themselves or in a copy shop.

The second prototype's design emphasized even more minimal text and more visual guidance, based on the feedback of the coaches that the first prototype was too language-based. For example, we removed the back-of-card questions and instead provided short prompts or icons to guide the conversation. Furthermore, while Prototype 1 was about inspiration (finding a wish), participants noted it lacked

support for translating that wish into action. To this end, Prototype 2 introduced a new element: “step cards” that help break down a goal or wish into smaller steps or learning actions. The step cards show different applications, websites, and devices, which can be used to reach a desired goal or wish. In prototype 2, once a learner’s goal has been identified using the “wish cards” of prototype 1, the coach and learner together outline an action plan. They might select or write down a first step (e.g. Using the camera app”) on a step card, then a next step (e.g. “Attach a photo to email”), and so on, creating a visual sequence or “roadmap” of the steps required to reach the goal. The steps can differ per goal and also per person even with the same goal or wish. There are also empty step cards to draw additional steps on. This makes the tool flexible and allows learners to make use of their “splinter skills.” Furthermore, the steps can be structured on a newly developed “board” into what a person can do already by themselves, what they can do with help, and what they can learn to do. The “learn to do” category can offer inspiration for concrete learning goals that are in line with the learner’s personal goals and motivation. Also, by including a category with what one can do already, we incorporated the explicit recommendation from the co-creation session to use positive, strengths-based framing. In the guiding manual, the DigiWish also explicitly prompts coaches to start by acknowledging what the learner already can do, avoiding a deficit-focused approach. Table 1 summarizes the two prototype iterations of DigiWish and how the design evolved.



Figure 2: The second prototype

4.4 Field Evaluation Results (Prototype 2)

The second prototype of DigiWish was met with enthusiastic feedback from the coaches and volunteers who integrated it into their sessions. Evaluators reported that the tool helped them engage clients in discussing digital skills in a “playful but effective” manner. Coaches found DigiWish valuable in uncovering learners’ needs and motivations. “It’s a handy and simple tool that we will definitely use more often,” said one library volunteer. The tool was helpful particularly for people who did not know what they wanted to learn, which according to interviewees happens frequently. “The theme cards helped to start the conversation.” Furthermore, the tool was found to be easy to use and accessible. As one volunteer remarked: “It’s not complicated. After a sort learning moment, you can start straight away.” DigiWish was also found to fit well with the organizations’ usual way of working as “It fits better than a questionnaire, which puts visitors off.”

The introduction of step cards and board were praised for bringing structure and motivating action. Coaches reported that clients could literally see their path forward laid out on the table, which made the learning process more tangible. “It was perfect to make the priorities and learning order clear,” noted one volunteer, referring to how the step-by-step plan provides a shared understanding of where to start and what to do next. He continued: “A nice overview for the client that we keep expanding with relevant step cards... I was very satisfied with the result, and our client was too. We immediately scheduled a follow-up appointment to work on clearing her phone and then moving her photos to her laptop. She also took a picture of the board herself.” Other coaches mentioned as a positive point “the three strokes on the board, which make clear what can be learned.” This illustrates how DigiWish not only identifies a goal but actively facilitates the planning of learning activities, which is a critical outcome.

Respondents also highlighted the positive tone and strengths-based approach of DigiWish as a benefit. One coach remarked: “*I can do this*’ is good because it works from a positive strengths analysis.” The field test also confirmed that the second version of the DigiWish was flexible, as intended. Different facilitators adapted DigiWish to their style, which the tool allowed. For example, some took out all the step cards at once, whereas others made a pre-selection. DigiWish was flexible enough to accommodate both approaches.

While the overall feedback was very positive, participants also provided suggestions. For example, because we used stock images and icons, a couple of coaches mentioned that not every image was immediately clear or culturally perfect for every client. Furthermore, one volunteer expressed preferring to “start doing straight away, rather than just talking.” Other suggestions included: making cards in multiple languages, including a way to categorize learning goals into learn now versus for the future, a method to gain more insight into the network of persons who can help the learner with digital skills, and a box or pouch to keep the DigiWish cards tidy and structured. Importantly, no one suggested the need for more text or instructions, implying our minimalist approach to text was on target.

In summary, Prototype 2’s evaluation suggests that DigiWish is a user-friendly tool for engaging low-literate adults in learning digital skills. It is promising for creating a positive experience for learners and provides coaches with a clear method to identify needs and plan training. The findings also point to potential enhancements. Some of the more minor alterations were made before the DigiWish was shared as a free download for printing. Other suggestions may be incorporated in potential future iterations.

5 Discussion

5.1 From assessment to empowerment

Traditional digital literacy programs often begin by assessing what learners do not know, which can inadvertently discourage those with low self-confidence. DigiWish instead starts with a learner’s aspirations and existing strengths to chart a path forward. Rather than merely testing skills at program intake, facilitators can use tools like DigiWish to have a constructive dialogue with learners, thereby personalizing learning goals and pathways. The enthusiastic response from coaches to this approach suggests that empowerment-based approach may be a better fit than assessment for this audience. This finding aligns with theories of adult learning, which emphasize the importance of self-directed goals and immediate relevance for adult learners (Howard et al., 2020).

5.2 Limitations and recommendations

It is important to also acknowledge the limitations of our study and tool. First, our evaluation was qualitative and involved a relatively small number of participants. While the feedback is encouraging, further research with both coaches and learners is needed to assess effectiveness for learning on the short and long-term. A next step could be a longitudinal study tracking learners who have gone through the DigiWish process, to see if it leads to higher involvement and retention in digital education programs or greater self-efficacy and usage of digital services. Second, as with any tool, one size may not fit all. The feedback hinted that slight variations of DigiWish might better serve different contexts (e.g. an adaptation for jobseekers versus one for senior citizens, each with context-specific imagery and example tasks). We hope further testing and use by other organizations will lead to further improvements and upscaling of DigiWish, potentially also in other languages and contexts.

One question open to exploration is whether the DigiWish methodology may also be useful as a tool in digital inclusion and digital skills research. Researchers studying digital skills and learning could explore if (an adaptation of) DigiWish offers a useful substitute for text-based assessments of digital skills. This could be a less threatening and more motivating alternative to “testing”, particularly for certain groups in society who are currently underrepresented in digital skills research. We can also envision that a variation of the DigiWish methodology, i.e., visualizing pathways to reach a goal and identifying where a person gets “stuck,” may be useful in uncovering (in)accessibility issues of digital tools and processes.

6 Conclusion

We have developed and tested DigiWish, an innovative co-created tool aimed at improving digital skills among adults with low reading literacy. Through two cycles of design and feedback, DigiWish evolved from a concept to a practical aid that community coaches have evaluated as accessible and motivating for learners. The tool shifts the focus from testing knowledge to inspiring a desire to learn, and it helps translate that desire into a concrete learning path. DigiWish in its current form is a prototype and there are still improvements to be made. However, as we have seen the usefulness of DigiWish in its current form, we have developed a free download version, which anyone can print themselves or at a copy shop (including

print instructions). We hope this work inspires further research and action toward inclusive digital empowerment.

Acknowledgements

This project was funded by Stichting Go Fonds. We also thank: Jessica Piotrowski and Frank Huysmans from the University of Amsterdam; Marieke van der Kruijs and Babette Zwitterlood of Utrecht Library; Maja van Eijndthoven and Femke Reijnen of BiSC Utrecht; Kiki Bathoorn, Ronald Ferdinandus, and colleagues and visitors of the Nieuwegein Library; Jacqueline Oude Hesselink, clients, and colleagues of DigiConsulent®; Elmira Leijten of VSO Rotterdam College; the participants of the co-design session; the reviewers of this paper; and all others who provided their invaluable ideas and feedback.

References

- Carretero, S., Vuorikari, R. and Punie, Y. (2017). DigComp 2.1 – The digital competence framework for citizens with eight proficiency levels and examples of use. Publications Office, <https://data.europa.eu/doi/10.2760/38842>
- Choudhary, H., & Bansal, N. (2022). Addressing digital divide through digital literacy training programs: A systematic literature review. *Digital Education Review*, 41, 224-248. <https://doi.org/10.1344/der.2022.41.224-248>
- de Vries, D. A., Piotrowski, J., de Vreese, C. H. (2025, March 3). Test your digital knowledge and skills with DigIQ. University of Amsterdam. <https://www.uva.nl/en/about-the-uva/organisation/faculties/faculty-of-social-and-behavioural-sciences/collaboration/research/digiq.html#:~:text=The%20DigIQ%20tool%20offers%20the,insight>
- de Vries D. A., , Piotrowski J. T., de Vreese, C. (2025). Developing the DigIQ: A measure of digital competence. *PLoS One* 20(5): e0322995. <https://doi.org/10.1371/journal.pone.0322995>
- European Commission. (2025, March 3). Europe's Digital Decade: digital targets for 2030. European Commission. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en
- Eurostat. (2025, March 3). 56% of EU people have basic digital skills. Eurostat. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20231215-3#:~:text=In%202023%2C%2056,least%20basic%20overall%20%2075>
- Halvorsrud, K., Kucharska, J., Adlington, K., Rüdell, K., Brown Hajdukova, E., Nazroo, J., ... & Bhui, K. (2021). Identifying evidence of effectiveness in the co-creation of research: A systematic review and meta-analysis of the international healthcare literature. *Journal of Public Health*, 43(1), 197-208. <https://doi.org/10.1093/pubmed/fdz126>
- Howard, J., Chong, J., & Bureau, J. (2020). The tripartite model of intrinsic motivation in education: A 30-year retrospective and meta-analysis. *Journal of Personality*, 88, 1268-1285. <https://doi.org/10.1111/jopy.12570>
- Kim, S., & Lee, A. (2016). Health-literacy-sensitive diabetes self-management interventions: A systematic review and meta-analysis.. *Worldviews on Evidence-Based Nursing*, 13, 324-333. <https://doi.org/10.1111/wvn.12157>
- Nedungadi, P., Devenport, K., Sutcliffe, R., & Raman, R. (2020). Towards a digital learning ecology to address the grand challenge in adult literacy. *Interactive Learning Environments*, 31, 383 - 396. <https://doi.org/10.1080/10494820.2020.1789668>
- Ngiam, N., Yee, W., Teo, N., Yow, K., Soundararajan, A., Lim, J., Lim, H., Tey, A., Tang, K., Tham, C., Tan, J., Lu, S., Yoon, S., Ng, K., & Low, L. (2022). Building digital literacy in older adults of low socioeconomic status in Singapore (Project Wire Up): Nonrandomized Controlled Trial. *Journal of Medical Internet Research*, 24, e40341. <https://doi.org/10.2196/40341>

- Munteanu, C., Molyneaux, H., Maitland, J., McDonald, D., Leung, R., Fournier, H., & Lumsden, J. (2014). Hidden in plain sight: Low-literacy adults in a developed country overcoming social and educational challenges through mobile learning support tools. *Personal and Ubiquitous Computing*, 18, 1455-1469. <https://doi.org/10.1007/s00779-013-0748-x>
- Smith, A., Rubio-Rico, L., McClelland, G. T., Monserrate-Gómez, S., Font-Jiménez, I., & de Molina-Fernández, I. (2022). Co-designing and piloting an Integrated Digital Literacy and Language Toolkit for vulnerable migrant students in higher education. *Education and Information Technologies*, 27, 6847-6867. <https://doi.org/10.1007/s10639-021-10845-0>
- Swart, J. (2022). The Wheel of Life as a Coaching Tool to Audit Life Priorities. <https://doi.org/10.13140/RG.2.2.25897.67688>
- van Deursen, A. J., & van Dijk, J. A. (2016). Modeling traditional literacy, Internet skills and Internet usage: An empirical study. *Interacting with Computers*, 28(1), 13-26. <https://doi.org/10.1093/iwc/iwu027>

STUDENTS' USAGE OF GENAI IN UNIVERSITIES OF APPLIED SCIENCES: EXPERIENCES AND DEVELOPMENT NEEDS FOR GUIDANCE AND SUPPORT

MINNA SILVENNOINEN,¹ SATU AKSOVAARA,¹
MERJA ALANKO-TURUNEN²

¹ Jamk University of Applied Sciences, Jyväskylä, Finland
minna.silvennoinen@jamk.fi, satu.aksovaara@jamk.fi

² Haaga-Helia University of Applied Sciences, Helsinki, Finland
merja.alanko-turunen@haaga-helia.fi

This article explores University of Applied Sciences students' awareness and use of GenAI, as well as their experiences with institutional guidelines and educator support. At the time of data collection, these institutions were in the early stages of GenAI adoption. A survey of 160 students revealed varied uses of GenAI and suggested that a notable proportion of students perceived the guidelines, and their delivery as unclear or inconsistent. Our results indicate that GenAI usage is not yet fully addressed in course-level practices, which may contribute to fragmented and inconsistent use of the technology among students. Moreover, students were more likely to seek help from peers than from educators when encountering difficulties with AI applications, which may reflect students' perceptions of the limited availability of educator support. The findings support previous research that there is a need for increased awareness of GenAI use in higher education among both students and educators. The study highlights the need for strengthening the role of educators in encouraging and influencing how their students perceive and adopt GenAI technology.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.29](https://doi.org/10.18690/um.fov.4.2025.29)

ISBN
978-961-286-998-4

Keywords:
students,
GenAI,
higher education,
guidelines,
educator support



University of Maribor Press

1 Introduction

The rapid development of Generative Artificial Intelligence (GenAI) is reshaping higher education (HE) pedagogy, requiring continuous adaptation from both educators and students (Crompton & Burke, 2023; Kasneci et al., 2023; Peres et al., 2023). This adaptation is particularly relevant for Universities of Applied Sciences (UAS), which focus on partnering with industry and integrating work experiences to meet workforce needs more effectively than traditional universities (Virolainen et al., 2024; Carvalho et al., 2019). UASs' incorporate GenAI technologies into their education, supporting their mission to train students to meet the needs of future workplaces. This prepares students more effectively for society (Arene, 2024). While increasing numbers of students and faculty are using GenAI (e.g. An et al., 2025), clear strategies and guidelines for the use and integration of GenAI into learning practices are thus needed to ensure effective and safe integration into their curricula and practices (Moorhouse et al., 2023). As a key driver of digital transformation, AI is increasingly shaping teaching, learning, and assessment, highlighting the need to examine the evolution of GenAI and its potential to transform traditional pedagogical practices (Ruiz-Rojas et al., 2023, 2024; Crawford et al., 2024; Eager & Brunton, 2023; Atchley et al., 2024). GenAI is a branch of artificial intelligence (AI) that generates human-like content—including text, images, video, music, and code—by analyzing patterns in training data (Baidoo-Anu & Ansah, 2023). A defining feature of these applications is their ability to generate contextually appropriate responses to user prompts (Peres et al., 2023). These systems use sophisticated deep-learning techniques to identify and replicate underlying linguistic patterns and structural elements in the text. In practice, GenAI enables personalisation of learning, the automation of content production, and the support of students' creative processes (Barrett & Back, 2023; Tlili et al., 2023; Hu, 2024). Its potential goes beyond student engagement and interaction - it is reshaping teaching methods and educational content creation (Chen et al., 2020; Kasneci et al., 2023; Onesí-Ozigagun et al., 2024) enabling customized learning materials thus promoting interactive and pedagogically engaging learning experiences (Crawford et al., 2023).

Previous research calls for more education-focused AI studies, noting that most AI researchers' engineering backgrounds lead to a focus on technology rather than educational aspects (Chiu et al., 2023). While GenAI has significant potential, its implementation in HE is still in its early stages, facing several challenges that require

research, such as the readiness of students, educators, and higher education institutions (HEIs) to utilize GenAI (UNESCO, 2023). This study aims to understand how students use GenAI tools in their studies as well as students' experiences with their institutional guidelines and support received with the use of GenAI. The context, two Finnish University of Applied Sciences (UAS) institutions engaged in the early adoption phase, and at the same time, Finland's UAS sector was updating recommendations for GenAI use in education (Arene, 2024) making this investigation timely and relevant as institutions aligned their GenAI policies. This study specifically focuses on the following research questions:

RQ 1 What is the level of GenAI's awareness and usage among UAS students?

RQ 2 What are the experiences of UAS students regarding their institutional guidelines and the support received concerning the use of GenAI?

RQ3 How are students' experiences with institutional guidelines and educator support related to whether they have studied GenAI?

2 Background

Many studies on students' expectations and experiences of using GenAI have already been carried out in different HE contexts (e.g. Almaraz-López et al., 2023; Hwang & Tu, 2021; Lemke et al., 2023) particularly exploring the use of GenAI applications such as chatGPT (Johnston et al., 2024; Ravšelj et al., 2025). Systematic scoping reviews have revealed that students use ChatGPT as a personal tutor for various learning purposes, including content explanation, feedback, and structuring writing (Deng et al., 2025). In higher education, there have been discussions about how students perceive GenAI merely as a tool, and how educators in HE categorize its use as plagiarism (Ibrahim et al., 2023). Even though HE students could recognize the potential of GenAI tools, a significant number of them possibly avoid the use of GenAI in academic contexts due to the fear of sanctions or uncertainty regarding its permissibility (Baidoo-Anu et al., 2024; Luo, 2024a; Luo 2024b). A common vision of how GenAI can best serve HE and what its role should be is needed.

Research highlights the importance of upskilling both educators and students to effectively integrate GenAI into teaching and learning processes (Chan & Lee, 2023; Chan & Tsi, 2024). Educational programs have been used to enhance students' GenAI skills, with promising results suggesting that those who attend university-led training feel more confident and motivated to adopt GenAI tools. (e.g. Southworth et al., 2023). However, it seems that much still is needed in students' proper adoption and understanding of the possibilities of GenAI. A comparative study by Abdelwahab et al. (2023) on business students revealed that the awareness of GenAI is insufficient for fully understanding its implications for themselves or their future careers. Almaraz-López et al. (2023) found that while university business students were aware of GenAI and eager to learn more, their knowledge remained limited due to inadequate training. Despite students' expectations for strong GenAI expertise and clear guidance from educators, limited staff competence and the lack of university-wide policies may hinder its adoption (Michel-Villarreal et al., 2023; Chiu et al., 2023; Bissessar, 2023; Johnston et al., 2024). Similarly, a preliminary Finnish study found that UAS students expect educators to provide guidance and permissions regarding GenAI use (Suonpää et al., 2024). Bearman & Ajjawi (2023) note that clear institutional GenAI guidelines set boundaries and standards, fostering awareness of its potential and limits in education, which in turn guides ethical and effective usage by students. Baidoo-Anu et al. (2024) thus highlight frequent concerns about inconsistent GenAI messaging across various departments in HE, emphasizing the need to explore the situation from students' perspectives.

Many academic institutions and stakeholder groups closely associated with HE have developed recommendations and principles for using GenAI, and in particular, how assessment practices should be reformed (Lodge et al., 2023). Sabzalieva & Valentini (2023) authored UNESCO's Quick Start Guide to ChatGPT and AI in HE, designed to quickly orient educators, though its recommendations target institutional leaders rather than students and educators. (Pratschke, 2024). According to Moorhouse et al. (2023) HEI's GenAI use guidelines for educators mainly cover three main areas: academic integrity, advice on assessment design and communicating with students, however under half of the institutions had developed these kinds of publicly available guidelines by 2023 (Moorhouse et al. 2023). It seems today that the majority of institutional guidelines of HE is rather embracing the GenAI technology than banning it (Xiao et al. 2023; McDonald et al., 2024). Arene (2024), as the coordinator of Finnish UAS rectors, released guidelines for students and educators on using

GenAI in studies, encouraging HEIs to guide its use in student learning activities (Arene, 2024).

There is a lack of systematic knowledge on how HEI's adopt mechanisms to promote GenAI among students. Additionally, there are few comprehensive studies on how institutional guidance shapes students' GenAI experiences and usage (Almassaad et al., 2024). Research is needed particularly in the context of UAS. Both students and educators may often struggle to interpret and apply GenAI guidelines consistently, leading to uncertainty and variation in practice (An et al., 2025). Unclear institutional policies on academic misconduct even allow students to justify behaviors they might otherwise deem unethical (Mulder et al., 2015). There is also a lack of knowledge about how these recommendations are translated into actual teaching practice and guidance activities. It is therefore important to not only identify the various ways in which students use GenAI, but also to understand the factors driving its use, such as the guidance from educational institutions and support from educators.

3 Context and methodology

Data collection targeted two UASs in Finland which shared a need for a baseline study to investigate students' use of GenAI and to monitor changes over two years., aligning with ongoing policy updates and GenAI implementation strategies in the HE. The survey included multiple-choice questions and Likert-type scale items. The items covered GenAI awareness and usage, experiences with institutional guidelines and educator support, as well as students' learning and support-seeking behavior.

The survey instrument was partially informed by the TEK [Trade union for academic engineers and architects in Finland] Student Survey: AI in Studies (2023), which explored technology students' experiences with GenAI tools in HE. Selected statements adapted from the TEK survey and analysed in this study focused on the usage of GenAI among UAS students (Table 2), its applications (Figure 1), and students' experiences with institutional guidelines and practices (Table 3). The survey was also partially informed by established frameworks of studies by Yilmaz et al. (2023) and Wang et al. (2023). Three statements selected for analysis were adapted from Yilmaz et al. (2023) and focused on students' experiences of support and help: "Educators important to me think that I should use AI tools in my studies", "I get

help from educators when I have difficulties using AI applications.” and “I get help from peers when I have difficulties using AI applications.” The data collection occurred during the Spring semester 2024 when the online survey was distributed to students through two primary channels: 1) Compulsory Course Distribution ensuring broad participation across various study programs. 2) Instructor-Led Dissemination in selected courses. Students were informed about the purpose of the study, participation was voluntary, and informed consent was requested within the questionnaire. The demographics of the participants are presented in Table 1.

Table 1: Demographics

Category	Group	N=160	%
Age Distribution	Under 20 years old	9	6
	20–24 years old	85	53
	25–30 years old	29	18
	Over 30 years old	37	23
Gender Distribution	Female	113	71
	Male	46	29
	Missing		<1
Education Level	Vocational Education (Secondary Level)	33	21
	High School	64	40
	University of Applied Sciences	48	30
	Other	15	9

3.3 Data analysis

The statistical analysis was conducted using descriptive analysis, cross-tabulation, and the chi-square test (χ^2). The descriptive analysis included means, standard deviations, medians, and percentage distributions. Cross-tabulation was used to examine students' experiences with AI tools and the factors influencing their usage. The statistical significance of differences between groups was tested using the chi-square test (χ^2 , $p < .05$), and effect sizes were assessed with Cramér's V. Statistical analyses were conducted using IBM SPSS v. 28.0 in a pseudonymized form.

4 Results

The level of GenAI's awareness and usage among UAS students (RQ1) was measured using a Likert scale ranging from 1 to 5. Responses were classified into 'users' and 'non-users' in Table 2.

Table 2: Awareness and usage of GenAI among UAS students

Usage of AI	Statement	N=160	%
Non-users	AI tools are new to me (1)	15	9
	I am aware of AI tools but have not used them (2)	39	24
	I am afraid to use them because I don't know if it's allowed (3)	8	5
Users	I occasionally use AI tools (4)	74	46
	I frequently use AI tools (5)	24	15

The survey results indicate that ChatGPT (free version) was the most used GenAI tool (97%), followed by Copilot (Bing) (15%) and Google Gemini (free version) (10%). Other tools had lower usage rates (1–9%). Figure 1 illustrates the applications of AI reported by UAS students.

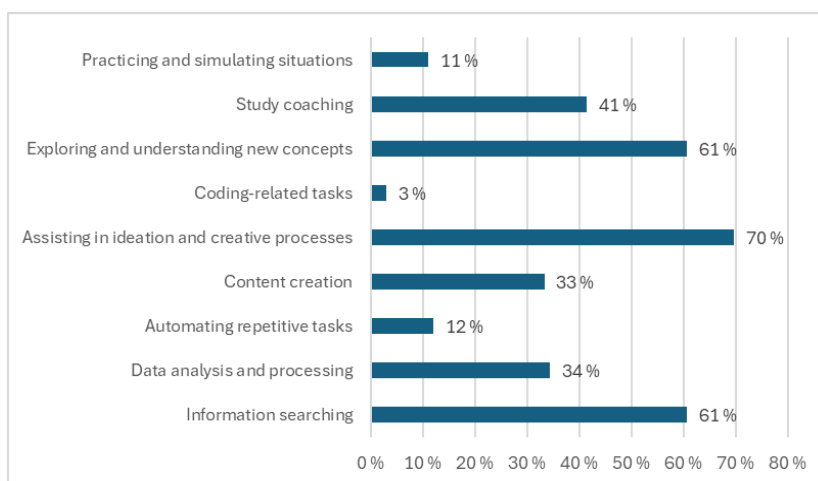


Figure 1: Applications of GenAI among UAS Students

The three most common uses of GenAI among students were assisting in ideation, information searching, and exploring new concepts, while content creation being only 33% (see Figure 1). Among those who reported using AI for other purposes (14%), the main applications included grammar and text proofreading, translation, generating self-assessment questions, simplifying tasks, and text formatting.

In line with **RQ2, student experiences regarding institutional guidelines and received support were examined**. Firstly, the focus was on institutional guidelines and practices (see Table 3). The responses were based on a Likert scale, where 1

represented "Strongly Disagree" and 5 represented "Strongly Agree". The largest standard deviations are highlighted.

Table 3: Student responses on institutional guidelines and practices regarding GenAI use

Statement	%						Mean	Std. Dev.
	Don't Know	1	2	3	4	5		
1. My HE institution has a clear policy on the use of AI tools in studies.	19	2	8	18	$\frac{2}{9}$	24	3.82	1.048
2. My HE institution encourages the use of AI tools in studies.	15	7	11	33	$\frac{2}{4}$	10	3.23	1.082
3. I know what to do if I use AI tools to assist with my learning tasks (e.g., in referencing).	16	13	17	19	$\frac{1}{9}$	15	3.07	1.336
4. I know how the use of AI applications affects the assessment of my learning tasks.	28	18	11	16	$\frac{1}{6}$	12	2.9	1.42
5. My HE institution limits the use of AI tools in studies.	28	7	18	25	$\frac{1}{4}$	8	2.97	1.135
6. The use of AI tools in studies is prohibited at my HE institution.	19	39	23	12	4	3	1.88	1.09

Secondly, we explored support in relation to GenAI use (see Figure 2).

A Pearson Chi-Square test was conducted to examine whether there was a significant difference in students' experiences of receiving support from educators versus friends. The results revealed a significant association between the source of support and the distribution of responses, $\chi^2(4) = 46.51$, $p < .001$. The effect size (Cramér's $V = 0.38$) indicates a moderate association, suggesting a meaningful difference in support experiences.

A series of Pearson Chi-Square (χ^2) tests were conducted to explore **relationships between students' experiences with institutional guidelines, educator support, and whether they have studied GenAI (RQ3)**. Crosstabulation analysis was first conducted to examine the relationship between whether students had or had not studied GenAI and their experiences with institutional guidelines. Chi-square tests ($p < 0.05$) revealed no statistically significant association, indicating that having studied GenAI does not influence students' experiences with the institutional guidelines presented in Table 2, (statements 1–6). Crosstabulation analysis was thereafter used to explore the relationships between the educators' encouragement

to use AI tools “Educators important to me think that I should use AI tools in my studies” and the same variables of statements concerning institutional guidelines and practices regarding GenAI use (see Table 3 statements), with Pearson Chi-Square tests conducted to assess the statistical significance of these associations.

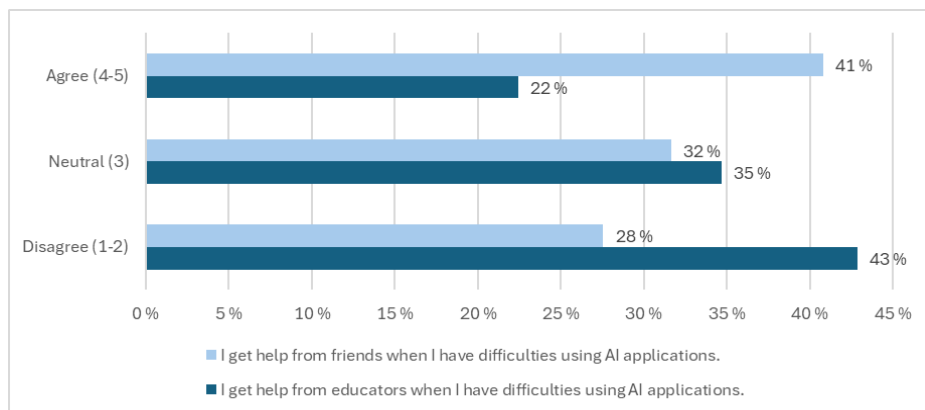


Figure 2: Student responses on receiving help from educators versus friends in using GenAI applications

The results revealed significant associations between students' perceptions of educators' encouragement to use AI tools and institutional policies. Specifically, students who believed that educators important to them supported/encouraged GenAI tool usage were more likely to report that their institution had a clear policy on AI tools, statement 1 ($\chi^2(5) = 17.130$, $p = .004$) and encouraged their use, statement 2 ($\chi^2(5) = 18.214$, $p = .003$). Additionally, there was a significant relationship between educators' encouragement of AI tool usage and students' understanding of how to proceed when using AI tools in assignments, statement 3 ($\chi^2(5) = 12.217$, $p = .032$). However, no significant association ($p < 0.05$) was found between educators' encouragement in relation to assessment and whether institutional limits or prohibits the use of AI tools (statements 4, 5 and 6).

Third crosstabulation was conducted to examine the relationship between students' perception of receiving help from educators "I get help from educators when I have difficulties using AI applications" when encountering difficulties with GenAI and the statements concerning institutional guidelines (see Table 3, statements). The findings indicate that students who perceived help from educators were significantly more likely to report knowing how to proceed when using AI tools in assignments, statement 3 ($\chi^2(5) = 14.753, p = .011$) and understanding how AI tool usage affects assignment assessments, statement 4 ($\chi^2(5) = 12.805, p = .025$). However, no significant association ($p < 0.05$) was found between educator support and students' perception of institutional guidelines, statements 1, 2, 5 or 6.

5 Discussion and conclusions

This study aimed to deepen the understanding of the current state of UAS students in using GenAI in Spring 2024. At that time, these organizations were still in the early stages of adoption, and guidelines and recommendations had only just begun to be developed. Findings showed that students are indeed using GenAI in their studies, with ChatGPT as the most used tool applied in diverse ways, including conceptual exploration and content generation (see also Digital Education Council 2024; Suonpää et al. 2024). The findings suggest that a notable proportion of students experience uncertainty regarding GenAI-related guidelines and policies at their institution. A considerable number of respondents selected neutral or 'Don't know' options, particularly in relation to policy clarity, encouragement, and implications for assessment, indicating a potential lack of clarity. Additionally, there is a noticeable absence of encouragement for students to use GenAI (see also Barrett & Pack, 2023; Hamerman et al., 2025). This underscores the need for clearer, more consistent, and more accessible instructions with no room for varying interpretations (see also M; Xiao et al. 2023; Dabis & Csáki, 2024).

Our results indicate that knowledge of handling GenAI usage in studies is incomplete, and when facing difficulties, students more often report turning to their friends than to educators, which may suggest greater accessibility or comfort with peer support. This may reflect students' perception of limited availability of timely educator support or inconsistencies in the delivery of institutional guidelines (see also Michel-Villarreal et al., 2023). Student hesitation may also be caused by fear of academic dishonesty being exposed (e.g. Luo, 2024b) since lack of clear institutional

policies contributes to ethical ambiguities (e.g. Duah & McGivern, 2024). Clear policies and guidelines ensure students' ethical and proficient use of GenAI. Our findings suggest that students who receive encouragement from educators tend to have a clearer understanding of GenAI guidelines and greater awareness of appropriate use. Guidelines are primarily communicated through educator interaction or institutional websites (see e.g. Duah & McGivern 2024). Instructions posted on intranet platforms alone appear insufficient; active educator involvement is essential for developing students' procedural knowledge and understanding of GenAI. Moreover, the findings suggest that formal study of GenAI may have limited impact if it is not clearly linked to practical use, since studying GenAI did not significantly influence students' experiences with guidelines.

Given the need for clear guidelines and the influential role of educators, enhancing their GenAI competence is essential for supporting students effectively. It is known that the extent to which educators perceive technology as effective or relevant can influence how they discuss, encourage, or guide students (see e.g. Kizilcec 2024). Guidance is essential for differentiating between 'what' the guidelines dictate and 'how' to apply these guidelines effectively. Institution guidelines are important during the transition phase when GenAI is new, as they support educators' confidence and understanding that GenAI use is pedagogically justified (see e.g. McDonald et al., 2024; Bates et al., 2020). HE educators are thus encouraged to move beyond basic use of GenAI toward more active and creative engagement, grounded in a deep understanding of its potential (Zhai, 2024; Cha et al., 2024). At a minimum, educators are expected to stay up to date with AI literacy, including both subject-specific applications of GenAI and its pedagogically sound use in teaching, learning, and assessment. To foster meaningful learning experiences and support students in navigating and managing GenAI, educators must learn to co-exist and collaborate with it, rather than resist its presence in educational contexts (see e.g. Chan & Tsi 2024).

Overall, this study relates the need for improved GenAI training and awareness among students and educators in HE (see e.g. Farrelly & Baker, 2023; Michel-Villarreal et al., 2023; Chiu et al., 2023). As a limitation, the total sample size offers only a moderate basis for future comparative analyses. The survey's focus on business and technical students may limit the applicability of the results to other fields. Further research is needed to examine the implementation of GenAI

guidelines, and the quality of support students receive—areas that will form the next phase of our analysis. Future studies should also explore educators' perceptions of GenAI and their role in its pedagogical use (also e.g. Kizilcec 2024; Chiu et al. 2023). It is important to evaluate the overall effectiveness of educational strategies in enhancing GenAI awareness across HE.

While institutional GenAI guidelines have largely focused on academic integrity, student involvement in their development has been limited. Collaborative guideline development involving all stakeholders should occur to align with the needs of both students and educators (see e.g. Cacho, 2024). HE institutions should not let the varying competencies of students, educators or academic dishonesty challenges impede GenAI adoption. Instead, they should focus on fostering continuous professional development (see also Collie & Martin, 2024). Pedagogical innovations are needed to enrich students' learning experiences and outcomes with GenAI. Therefore, teacher education programs should not only prepare future educators to use GenAI pedagogically, but also encourage them to cultivate a spirit of critical experimentation and active exploration with their students (Zhang & Zhang, 2024). There is a risk that institutional guidelines may restrict the innovative use of GenAI—particularly if they overlook its role as a learning assistant—which, in turn, may hinder the development of new forms of human–technology collaboration. HE should thus remain responsive to the evolving needs of both students and educators, ensuring that GenAI is fully leveraged for its educational potential. Through this comprehensive approach, HE is prepared to navigate the complex landscape of digital transformation.

References

- Abdelwahab, H. R., Rauf, A., & Chen, D. (2023). Business students' perceptions of Dutch higher educational institutions in preparing them for artificial intelligence work environments. *Industry and Higher Education*, 37(1), 22–34. <https://doi.org/10.1177/09504222221087614>
- An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, 22(1), 10. <https://doi.org/10.1186/s41239-025-00507-3>
- 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>
- Almaraz-López, C., Almaraz-Menéndez, F., & López-Esteban, C. (2023). Comparative study of the attitudes and perceptions of university students in business administration and management

- and in education toward artificial intelligence. *Education Sciences*, 13(6), 609.
<https://doi.org/10.3390/educsci13060609>
- Arene. (2024). Arene's recommendations on the use of artificial intelligence for universities of applied sciences. <https://bit.ly/43itRLz>
- Atchley, P., Pannell, H., Wofford, K. (2024). Human and AI collaboration in the higher education environment: opportunities and concerns. *Cogn. Research* 9, 20.
<https://doi.org/10.1186/s41235-024-00547-9>
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62.
- Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20(1), 59. <https://doi.org/10.1186/s41239-023-00427-0>
- Bates, T., Cobo, C., Mariño, O., & Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*, 17(1), 42, s41239-020-00218-x. <https://doi.org/10.1186/s41239-020-00218-x>
- Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*.
<https://doi.org/10.1111/bjet.13337>
- Bissessar, C. (2023). To use or not to use ChatGPT and assistive artificial intelligence tools in higher education institutions? The modern-day conundrum – students' and faculty's perspectives. *Equity in Education & Society*, 27526461231215083.
<https://doi.org/10.1177/27526461231215083>
- Cacho, R. (2024). Integrating generative AI in university teaching and learning: A model for balanced guidelines. *Online Learning*, 28(3), 55-81 <https://doi.org/10.24059/olj.v28i3.4508>
- Carvalho, C., Friman, M., & Mahlamäki-Kultanen, S. (2019). Pedagogy in Finnish higher education: a case example of Häme University of Applied Sciences. *Revista Praksis*, 3, 04.
<https://doi.org/10.25112/rpr.v3i0.1976>
- Cha, Y., Dai, Y., Lin, Z., Liu, A., & Lim, C. P. (2024). Empowering university educators to support generative AI-enabled learning: Proposing a competency framework. *Procedia CIRP*, 128, 256–261. <https://doi.org/10.1016/j.procir.2024.06.021>
- Chan, C. K. Y., & Tsi, L. H. Y. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, 83, 101395.
<https://doi.org/10.1016/j.stueduc.2024.101395>
- Chan, C.K.Y., & Lee, K.K.W. (2023) The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers?. *Smart Learn. Environ.* 10, 60
<https://doi.org/10.1186/s40561-023-00269-3>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A Review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118.
<https://doi.org/10.1016/j.caeai.2022.100118>
- Collie, R. J., & Martin, A. J. (2024). Teachers' motivation and engagement to harness generative AI for teaching and learning: The role of contextual, occupational, and background factors. *Computers and Education: Artificial Intelligence*, 6,
<https://doi.org/10.1016/j.caeai.2024.100224>
- Crawford, J., Vallis, C., Yang, J., Fitzgerald, R., O'Dea, C., & Cowling, M. (2024). Artificial Intelligence is Awesome, but Good Teaching Should Always Come First. *Journal of University Teaching and Learning Practice*, 20(7). <https://doi.org/10.53761/1.20.7.01>

- Crawford, J., Cowling, M., & Allen, K. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, 20(3). <https://doi.org/10.53761/1.20.3.02>
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Dabis, A., Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. *Humanit Soc Sci Commun* 11, 1006. <https://doi.org/10.1057/s41599-024-03526-z>
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Digital Education Council. (2024). AI or not AI: What students want. Digital Education Council. <https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-student-survey-2024>
- Duah, J. E., & McGivern, P. (2024). How generative artificial intelligence has blurred notions of authorial identity and academic norms in higher education, necessitating clear university usage policies. *The International Journal of Information and Learning Technology*, 41(2), 180–193. <https://doi.org/10.1108/IJILT-11-2023-0213>
- Eager, B., & Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *Journal of University Teaching and Learning Practice*, 20(5). <https://doi.org/10.53761/1.20.5.02>
- Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>
- Hamerman, E.J., Aggarwal, A. & Martins, C. (2025). An investigation of generative AI in the classroom and its implications for university policy. *Quality Assurance in Education*, 33(2) pp. 253–266. <https://doi.org/10.1108/QAE-08-2024-0149>
- Hu, S. (2024). The effect of artificial intelligence-assisted personalized learning on student learning outcomes: A meta-analysis based on 31 empirical research papers. *Science Insights Education Frontiers*, 24(1), 3873–3894. <https://doi.org/10.15354/sief.24.re395>
- Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, 9(6), 584. <https://doi.org/10.3390/math9060584>
- Ibrahim, H., Liu, F., Asim, R., Battu, B., Benabderrahmane, S., Alhafni, B., Adnan, W., Alhanai, T., AlShebli, B., Baghdadi, R., Bélanger, J. J., Beretta, E., Celik, K., Chaqfeh, M., Daqaq, M. F., Bernoussi, Z. E., Fougne, D., Garcia de Soto, B., Gandolfi, A., ... Zaki, Y. (2023). Perception, performance, and detectability of conversational artificial intelligence across 32 university courses. *Scientific Reports*, 13(1), 12187. <https://doi.org/10.1038/s41598-023-38964-3>
- Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, 20(1), 2. <https://doi.org/10.1007/s40979-024-00149-4>
- Kasnezi, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasnezi, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kizilcec, R. (2024). To Advance AI use in education, focus on understanding educators. *Int J Artif Intell Educ* 34, 12–19. <https://doi.org/10.1007/s40593-023-00351-4>
- Lemke, C., Kirchner, K., Anandarajah, L., & Herfurth, F. (2023). Exploring the student perspective: assessing technology readiness and acceptance for adopting large language models in higher

- education. *European Conference on E-Learning*, 22(1), 156–164.
<https://doi.org/10.34190/ecel.22.1.1828>
- Lodge, J. M., Howard, S., Bearman, M., Dawson, P., & Associates. (2023). Assessment reform for the age of Artificial Intelligence. Tertiary Education Quality and Standards Agency.
<https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf>
- Luo, J., (2024a). A critical review of GenAI policies in higher education assessment: A call to reconsider the “originality” of students’ work. *Assessment & Evaluation in Higher Education*, 49 (5) <https://doi.org/10.1080/02602938.2024.2309963>
- Luo, J. (2024b). How does GenAI affect trust in teacher-student relationships? Insights from students’ assessment experiences. *Teaching in High Education* 1–16
<https://doi.org/10.1080/13562517.2024.2341005>
- McDonald, P., Hay, S., Cathcart, A., & Feldman, A. (2024). Apostles, agnostics and atheists: Engagement with generative AI by Australian university staff. QUT Centre for Decent Work and Industry, Brisbane, Qld. <https://eprints.qut.edu.au/241763/>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. *Education Sciences*, 13(9), 856.
<https://doi.org/10.3390/educsci13090856>
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world’s top-ranking universities. *Computers and Education Open*, 5, 100151.
<https://doi.org/10.1016/j.caeo.2023.100151>
- Mulder, L. B., Jordan, J., & Rink, F. (2015). The effect of specific and general rules on ethical decisions. *Organizational Behavior and Human Decision Processes*, 126, 115–129.
<https://doi.org/10.1016/j.obhdp.2014.11.002>
- Onesi-Ozigagun, O., Olofade, Y.J., Eyo-Udo, N. L., & Ogundipe, D.O. (2024). Revolutionizing education through AI: A comprehensive review of enhancing learning experiences. *International Journal of Applied Research in Social Sciences*, 6(4), 589–607.
<https://doi.org/10.51594/ijarss.v6i4.1011>
- Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*, 40(2), 269–275.
<https://doi.org/10.1016/j.ijresmar.2023.03.001>
- Pratschke, B. M. (2024). *Generative AI and Education: Digital Pedagogies, Teaching Innovation and Learning Design*. Springer.
- Ravšelj, D., Keržič, D., Tomažević, N., Umek, L., Brezovar, N., Iahad, N. A., Abdulla, A. A., Akopyan, A., Segura, M. W. A., AlHumaid, J., Allam, M. F., Alló, M., Andoh, R. P. K., Andronic, O., Arthur, Y. D., Aydin, F., Badran, A., Balbontín-Alvarado, R., Saad, H. B., ... Aristovnik, A. (2025). Higher education students’ perceptions of ChatGPT: A global study of early reactions. *PLOS ONE*, 20(2), e0315011.
<https://doi.org/10.1371/journal.pone.0315011>
- Ruiz-Rojas, L. I., Acosta-Vargas, P., De-Moreta-Llovet, J., & Gonzalez-Rodriguez, M. (2023). Empowering Education with Generative Artificial Intelligence Tools: Approach with an Instructional Design Matrix. *Sustainability*, 15(15), 11524.
<https://doi.org/10.3390/su151511524>
- 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>
- Southworth, J., Migliaccio, K., Glover, J., Glover, J., Reed, D., McCarty, C., Brendemuhl, J., & Thomas, A. (2023). Developing a model for AI across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artificial Intelligence*, 4, 100127. <https://doi.org/10.1016/j.caeai.2023.100127>

- Suonpää, M., Heikkilä, J. & Dimkar, A. (2024) Students' perceptions of generative AI usage and risks in a Finnish higher education institution, *INTED2024 Proceedings*, pp. 3071-3077. <https://doi.org/10.21125/inted.2024.0825>
- Tekniikan akateemiset TEK. (2023). TEK student survey: AI in studies. <https://www.tek.fi/en/news-blogs/universities-and-workplaces-need-guidance-on-how-to-use-artificial-intelligence>
[Survey results in Finnish: <https://www.tek.fi/fi/tietoa-tekista/tutkimus/tek-tutkii-opiskelijat/opiskelijatutkimus-tekoaly-opinnoissa>]
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. <https://doi.org/10.1186/s40561-023-00237-x>
- UNESCO (2023). Guidance for generative AI in education and research (F. Miao & W. Holmes, Authors). UNESCO. <https://doi.org/10.54675/EWZM9535>
- Virolainen, M. H., Heikkinen, H. L. T., & Laitinen-Väänänen, S. (2024). The evolving role of Finnish universities of applied sciences in the regional innovation ecosystem. *Journal of Vocational Education & Training*, 1–21. <https://doi.org/10.1080/13636820.2024.2443905>
- Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, 42(9), 1324–1337. <https://doi.org/10.1080/0144929X.2022.2072768>
- Xiao, P., Chen, Y. and Bao, W. (2023). Waiting, banning, and embracing: An empirical analysis of adapting policies for generative AI in higher education. <http://dx.doi.org/10.2139/ssrn.4458269>
- Yilmaz, F. G. K., Yilmaz, R., & Ceylan, M. (2023). Generative artificial intelligence acceptance scale: A validity and reliability study. *International Journal of Human–Computer Interaction*, 1–13. <https://doi.org/10.1080/10447318.2023.2288730>
- Zhai, X. (2024). Transforming teachers' roles and agencies in the era of generative AI: Perceptions, acceptance, knowledge, and practices. *Journal of Science Education and Technology*, 1-11. <https://doi.org/10.48550/arXiv.2410.03018>
- Zhang, J., & Zhang, Z. (2024). AI in teacher education: Unlocking new dimensions in teaching support, inclusive learning, and digital literacy. *Journal of Computer Assisted Learning*, 40(4), 1871–1885. <https://doi.org/10.1111/jcal.12988>

LLM PIPELINE FOR MAPPING HETEROGENEOUS DATA: A CASE STUDY IN FOOD CLASSIFICATION

KEVIN NILS RÖHL,¹ RAINER ALT,² JAN WIRSAM¹

¹ HTW Berlin University of Applied Sciences, Berlin, Germany

roehl@htw-berlin.de, wirsam@htw-berlin.de

² Leipzig University, Leipzig, Germany

rainer.alt@uni-leipzig.de

Accurate food classification is essential for ensuring compliance with dietary regulations, nutritional standards, and sustainability guidelines, but it remains challenging due to fragmented data and semantic complexity. This study presents a pipeline leveraging large language model (LLM) embeddings, ontology mapping, and human-in-the-loop validation to enhance food classification in institutional food services. The pipeline achieves high accuracy in dietary-group mapping (precision 0.94, recall 0.91, F1-score 0.92), though precise FoodEx2 code matching remains challenging. A confidence-based validation strategy effectively balances automated processes with expert oversight to manage ambiguity. The proposed approach enables digital transformation of traditionally fragmented food service systems, enhancing transparency, operational efficiency, and alignment with dietary and public health guidelines. Future research should deploy this pipeline in operational canteen settings to refine embedding techniques, enhance accuracy, and support sustainable nutrition management.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.30](https://doi.org/10.18690/um.fov.4.2025.30)

ISBN

978-961-286-998-4

Keywords:

AI pipeline,
LLM embeddings,
classification,
food,
ontology



University of Maribor Press

1 Introduction

The rapid advancement of Emerging Digital Technologies (EDT), including Artificial Intelligence (AI), data pipelines, and automation, is reshaping industries by enabling data-driven decision-making, and process optimization (Oluwaseun Badmus *et al.*, 2024). From healthcare and finance to supply chain management and public services, these technologies offer increased efficiency (Bialas *et al.*, 2023; Elias *et al.*, 2024; Khedr and S, 2024; Kulal *et al.*, 2024). However, their widespread adoption presents significant challenges, particularly regarding data interoperability, ethical AI governance, and transparency (Danks and Trusilo, 2022).

In institutional food services, these challenges manifest in heterogeneous data formats, fragmented digital ecosystems, and the complexity of integrating AI-driven solutions into existing infrastructure (Wolfert *et al.*, 2023; Agrawal *et al.*, 2025; Dhal and Kar, 2025). Ensuring sustainable and nutritionally balanced food offerings in canteens requires a structured approach to data harmonization and regulatory alignment (Gaitán-Cremaschi and Valbuena, 2024). One approach to this is FoodEx2, which is a comprehensive food classification system developed by the European Food Safety Authority and is used to categorize and standardize food products, ingredients, and food-related data (EFSA, 2015). By utilizing food ontologies like FoodOn, food items can be systematically classified and connections between food, health, and the environment established (Dooley *et al.*, 2018). Unlike isolated ontologies, FoodOntoMap links multiple sources with semantic tags, enhancing interoperability (Popovski *et al.*, 2019). This ontology mapping addresses the lack of annotated food corpora and named-entity recognition systems, supporting research on food systems, human health, and sustainability (Popovski, Seljak and Eftimov, 2020). Food classification is essential for evaluating national and international dietary guidelines, as all food-based dietary guidelines (FBDG) rely on food groups. Across 90 countries, FBDG emphasizes core principles such as dietary variety, prioritizing fruits, vegetables, and legumes, and limiting sugar, fat, and salt. In detail, the recommendations on dairy, red meat, and fats can vary (Herforth *et al.*, 2019). A systematic review by Leme *et al.*, 2021 found that, on average, only 40% of individuals across both high-income and low- to middle-income countries meet their national dietary recommendations. Addressing this issue requires scalable tools to evaluate the nutritional quality of the food supply and support the development of effective public health interventions. Despite the challenges, existing research shows

promising AI approaches in extracting food data and estimating nutritional values (Hu, Ahmed and L'Abbé, 2023; Harris *et al.*, 2025). To further address these challenges, knowledge graphs, NER and machine learning integration have emerged as a promising solution (Cudré-Mauroux, 2020). Ontologies provide a structured framework for standardizing food-related data, allowing AI-driven systems to harmonize diverse nutritional databases, procurement records, and sustainability tracking tools (Popovski *et al.*, 2019; Shirai *et al.*, 2021; Min *et al.*, 2022).

This study applies the EDT framework to enable business integration and digital transformation within institutional food services, exemplified through company canteens (Serrano-Santoyo *et al.*, 2021). The developed pipeline, leveraging Large Language Models (LLM), addresses fragmented data, inconsistent food classification, and regulatory compliance, thereby enhancing transparency and operational efficiency. By aligning food offerings with dietary guidelines and public health goals, the pipeline supports the nutritional and operational transformation to an adaptive digital ecosystem.

2 Literature

2.1 Data Pipelines

Data pipelines play a critical role in integrating, processing, and analyzing data across multiple sources. They enable data flow between disparate systems, ensuring that information is structured and ready for analysis. In sectors where data originates from heterogeneous and unstructured sources, traditional data pipelines often struggle to maintain interoperability (Foidl *et al.*, 2024). Unlike conventional extract, transform and load systems, modern data pipelines leverage artificial intelligence and automation to handle real-time ingestion, data standardization, and transformation (Kolluri, 2024). In institutional food services, information is often scattered across different systems, including Enterprise Resource Planning (ERP) software, supplier databases, nutritional databases, and menu management tools. These datasets frequently exist in incompatible formats such as Excel spreadsheets, PDFs, CSV files, and proprietary ERP exports. The lack of a standardized data exchange format makes integration difficult, requiring complex preprocessing before meaningful analysis can take place (Zadeh *et al.*, 2018). Greater access to real-world data further complicates the evolution of data schemas across platforms (Zhang *et al.*, 2022).

Current data pipelines struggle with inconsistencies, quality issues, and inefficiencies in dynamic environments. While machine learning can automate schema adaptation and improve data quality, their use in scalable, cross-platform workflows is still limited. More research is needed to develop AI-driven automation for better interoperability, quality assurance, and regulatory compliance (Santhosh Bussa, 2024).

2.2 Ontology Mapping for Data Integration

Ontologies provide structured vocabularies that enable semantic interoperability between disparate datasets. In domains where data consistency is critical, ontology-based classification ensures that different terms referring to the same entity are aligned under a common framework (Gruber, 1995). While ontologies have traditionally been developed as rule-based systems, recent advances in artificial intelligence and natural language processing have enhanced their usability, enabling automated ontology mapping across multiple data sources (Wei and Li, 2025).

Machine learning models, particularly deep learning and transformer-based architectures have significantly improved the automation of entity recognition and classification. LLMs are now capable of extracting structured knowledge from unstructured text, making them valuable tools for ontology mapping in data pipelines (Ciatto *et al.*, 2025). With advanced LLMs from companies like OpenAI, ontology mapping with embedding models like text-embedding-3-large has significantly improved in accuracy. Embeddings represent concepts as high-dimensional vectors, capturing semantic similarities beyond lexical differences. This enhances alignment in complex domains like biomedicine and regulatory compliance, where terminologies evolve across institutions. By reducing reliance on rule-based mappings, embeddings improve scalability while maintaining high precision in entity linking (Sousa, Lima and Trojahn, 2025; Taboada *et al.*, 2025). In the healthcare sector, AI-driven Fast Healthcare Interoperability Resources mapping has enabled seamless integration of patient records across hospitals, insurance providers, and government agencies (Li *et al.*, 2023). Similarly, in supply management, AI-powered ontology mapping has improved logistics efficiency by standardizing product categories across global supplier networks (Regal and Pereira, 2018).

Despite the existence of standardized food classification systems like FoodEx2 and ontologies, their practical implementation remains limited, and current policies still fail to drive the necessary transformation toward a sustainable food system. To overcome fragmented approaches and address systemic crises, a new research and policy agenda is needed that strengthens cross-sectoral governance and effectively integrates food policies (Edwards, Sonnino and López Cifuentes, 2024). AI-based mapping offers a potential solution to those challenges by automatically linking raw ingredient data to a predefined classification system, reducing manual labor and improving data consistency (Hua *et al.*, no date; Goel and Bagler, 2022; Youn, Li and Tagkopoulos, 2023).

2.3 Digital Integration for Sustainable Food Systems

Digital tools like blockchain, AI, and digital twins can boost sustainability in the food value chain by improving efficiency and resource use. Yet, outdated infrastructure, regulatory hurdles, and fragmented data limit their integration and impact (Michel *et al.*, 2024). Achieving meaningful business integration in institutional food services requires overcoming siloed policy approaches by connecting policies across sectors and governance scales. Effective integration involves moving away from isolated processes toward interconnected governance frameworks, addressing interactions among health, environmental, and socio-economic sectors (Edwards, Sonnino and López Cifuentes, 2024). By systematically combining fragmented data from diverse stages of the food supply chain, AI and digitalization provide advanced predictive capabilities. Furthermore, integrating data-driven approaches fosters transparency, trust, and compliance with sustainability objectives throughout the food value chain (Marvin *et al.*, 2022).

3 Methodology

This study follows a Design Science Research approach to develop an AI-driven data pipeline for classification and mapping. The research focuses on structuring unstructured data, aligning it with predefined taxonomies, and integrating mechanisms for human oversight to ensure regulatory traceability. The study applies principles from the Exploratory Framework for Ethical and Regulatory Implications of EDT in Table 1, to examine how AI-driven classification can be validated and mapped to food guidelines. The methodology involves the development of a

crosswalk-based framework to establish interoperability between classification systems.

Table 1: Framework for Ethical and Regulatory Implications of EDT

Observed Conditions	Research Questions	Desired Conditions
Recipe data in PDFs is unstructured, making it difficult to analyze and categorize.	How can LLMs be leveraged to classify and map food data from unstructured text?	AI should support, rather than replace, expert decision-making in food classification and dietary assessments.
Current methods for classifying foods against nutritional guidelines are manual, time-consuming	How can regulatory elements be integrated into AI-driven food classification?	The pipeline should improve efficiency while ensuring transparency

3.1 Architecture

The LLM pipeline for heterogeneous data in Figure 1 supports transparent and regulatory-compliant AI decisions through ontology mapping, iterative learning, and structured expert oversight. By integrating human-in-the-loop validation, the pipeline ensures alignment with policy and ethical governance principles. Its architecture facilitates business integration by transforming fragmented food-service data into a unified digital ecosystem, shifting from manual, disconnected processes toward integrated compliance and sustainability management. The pipeline’s adaptable design further allows implementation across multiple domains, including public health, sustainability initiatives, and environmental monitoring, where regulatory tracking and accuracy are critical.

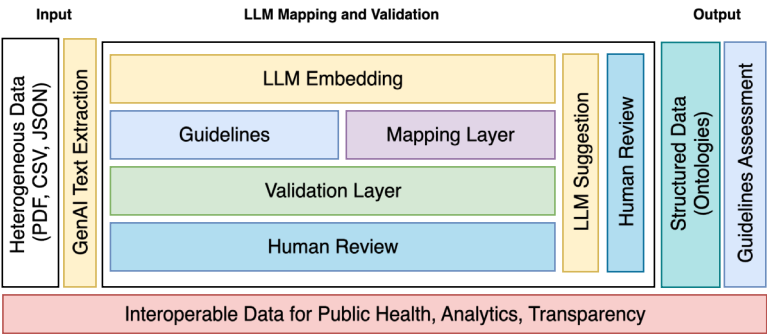


Figure 1: LLM Pipeline for Heterogeneous Data Mapping

Unstructured data from diverse sources (e.g., text documents, PDFs, databases) can be extracted and standardized using Natural Language Processing and entity recognition models. AI models generate semantic embeddings that allow for ontology-based classification, linking extracted data to structured taxonomies or regulatory databases. The use of crosswalk files ensures consistency and facilitates interoperability between national and international guidelines. This approach allows the pipeline to adapt to country-specific regulatory frameworks while maintaining cross-domain applicability. AI-generated classifications undergo a confidence-based validation process, where low-confidence mappings are flagged for human review. This ensures that automated classification aligns with national guidelines, balancing efficiency with expert oversight. Validated classifications feed into an iterative learning system, improving ontology mapping accuracy over time. This continuous refinement makes the system adaptive to policy changes.

3.2 Use-Case Canteen Recipe Ingredient Classification

OpenAI's LLM embedding “text-embedding-3-large” is used to match extracted entities to predefined taxonomies, selected for its semantic performance and ease of integration. Human validation processes refine mappings and address uncertainties. The pipeline architecture is implemented in the context of a company canteen setting, where ingredient names can be extracted from PDFs, and mapped to Planetary Health Diet (PHD) groups and FoodEx2 term codes.

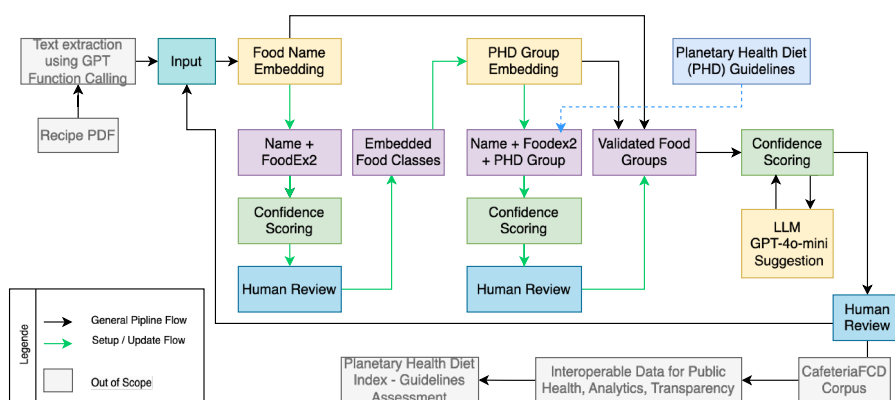


Figure 2: Company Canteen Pipeline for PHD Mapping

3.3 Validation

To validate the embeddings mapping, a test set of FoodEx2 names were manually labeled with PHD categories from the 430 subgroups of the FoodEx2 hierarchy code and applied to the underlying names. The first and second hierarchy has been removed to reduce misleading general food category matching, resulting in 4,416 entries. Flavors, vitamins, chemical elements, and composite dishes, which could belong to multiple food groups, were labeled as “unknown”. The final validation file consists of 2,750 entries, including FoodEx2 codes and PHD group classifications. To assess the effectiveness of the pipeline, F1-score and recall metrics were used to evaluate the PHD group embeddings generated by the “text-embedding-3-large” model against the validated file. A second test was conducted using zero-shot prompting with gpt4o-mini for mapping PHD groups to the FoodEx2 validation names. Another validation for ingredient data, was performed by mapping FoodEx2 name embeddings to given ingredients from FRIDA, the national food database from Denmark. FoodEx2 codes from FRIDA have been filtered based on the “unknown” codes list earlier, resulting in 935 entries with food names and their validated FoodEx2 codes. Embeddings have been created for all entries using “text-embedding-3-large” to compare them with the embedded FoodEx2 names. Cosine similarity was applied to determine the accuracy of this mapping.

4 Results

Figure 3 presents the classification results for a food group matching task using an embedding model (“text-embedding-3-large”) and GPT-4o-mini. The GPT-4o-mini model exhibits higher performance across all food groups. The “text-embedding-3-large” model achieves a precision of 0.84, recall of 0.78, and an F1-score of 0.79, while GPT-4o-mini attains a precision of 0.94, recall of 0.91, and an F1-score of 0.92. In the “text-embedding-3-large” classification, the food categories “eggs” (precision: 0.17, recall: 1.00) and “unsaturated fats” (precision: 1.00, recall: 0.06) represent the lowest-performing cases, with eggs displaying low precision but high recall, and unsaturated fats exhibiting high-precision but low recall. Additionally, starchy vegetables show low precision (0.51) and recall (0.53). In contrast, dairy (precision: 0.94, recall: 0.84) is among the highest-performing categories. The GPT-4o-mini model demonstrates its lowest classification performance in the “unsaturated fats” (precision: 0.79, recall: 0.62) and “saturated fats” (precision: 0.71,

recall: 0.78) categories, while achieving the highest precision and recall for “fish and seafood” (precision: 0.99, recall: 0.98) and “eggs” (precision: 0.97, recall: 0.97).

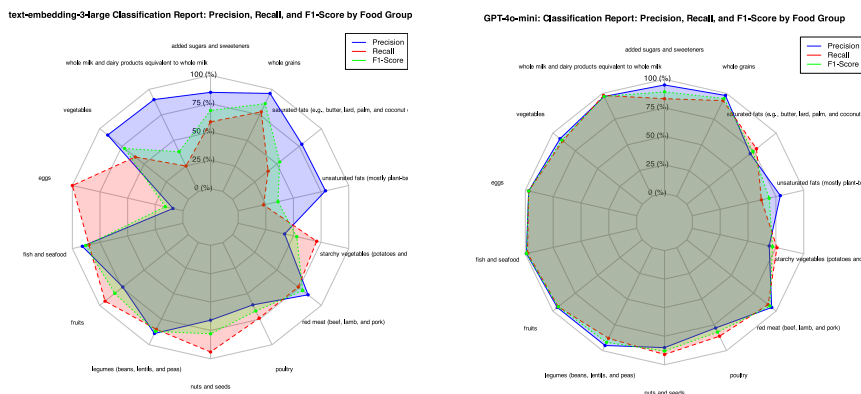


Figure 3: Food Group Classification Report

The similarity score distribution in Figure 4 shows a distinction between correct and incorrect classifications of FoodEx codes, with correct matches generally having higher similarity scores. Incorrect classifications tend to cluster at lower values, with some overlap between the distributions. The median similarity score for correct matches is higher than for incorrect ones.

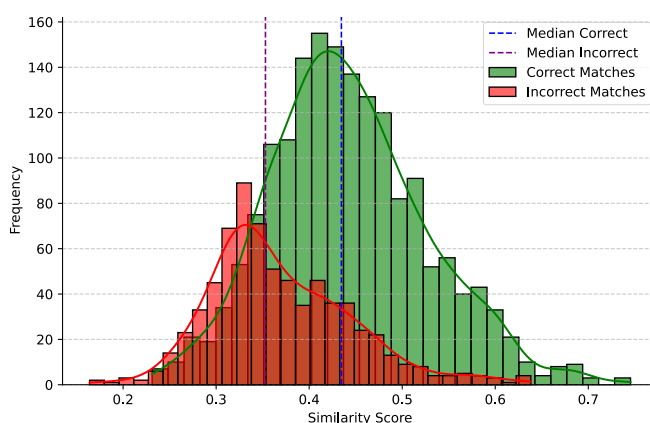


Figure 4: Distribution of Similarity Scores for PHD Category Matches

Table 2 provides two examples of false classifications in the text-embedding model. The model misclassified “emperors” as eggs (similarity score: 0.257) and “sea asters” as fish and seafood (similarity score: 0.372).

Table 2: False PHD Category Labels

Food Name	Correct Category	Incorrect Category	Cosine Similarity
Emperors	Fish and seafood	Eggs	0.257
Sea asters	Vegetables	Fish and seafood	0.372

4.1 FoodEx2 Code Matching Results

Figure 5 visualizes the distribution of similarity scores for correct and incorrect FoodEx2 code matches using OpenAI’s “text-embedding-3-large” model. Correct matches are shown in green (median 0.73), while incorrect matches are shown in red (median 0.66). The graph on the right displays the accuracy of FoodEx2 code matching based on cosine similarity across the closest three matches. The 1st match has an accuracy of 46.28%, the 2nd match has 13.19%, and the 3rd match has 5.74%.

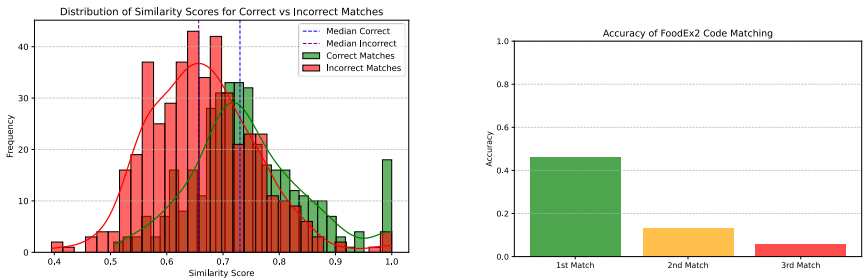


Figure 5: Embedding Matches of FoodEx2 and FRIDA Food Names

5 Discussion

The developed LLM-based classification pipeline demonstrates significant potential for aligning food categorization with national and international dietary guidelines, such as the PHD. By integrating crosswalk files, category assessment, and confidence-based validation, the system provides a structured approach to food classification. The ability to automate category assignment using GPT-4o-mini has

shown promising results, with the model successfully aligning food items to broader dietary groups with high accuracy. This suggests that LLM-assisted classification can support large-scale food policy assessments, improving transparency in food system monitoring. However, while category-level classification performs well, the exact matching of FoodEx2 codes remains a challenge due to the nuances in food naming conventions. The results showed that 46.28% of food items were correctly classified in the first match, while additional codes were identified in the second and third matches. A small part of the matching is shown in Table 3 in the appendix. This highlights that even when the top match is incorrect, valid classifications frequently appear in the nearest similarity ranks, demonstrating that embeddings capture meaningful relationships between food names even if perfect alignment is not always achieved.

5.1 Balancing Precision

Overlapping similarity scores between correct and incorrect classifications of FoodEx2 codes, complicate establishing a clear decision threshold. Especially when food names differ subtly due to preparation methods or regional variations. For example, categories like “eggs” and “unsaturated fats” illustrate how semantic ambiguity can significantly impact precision and recall.

This underscores the need for a strategic balance between automation accuracy and expert validation. Setting similarity thresholds too strictly reduces classification errors but increases manual effort, while lower thresholds boost automation but risk misclassification. Thus, confidence-based human oversight is essential for balancing precision effectively. Ensuring this balance is critical for reliable dietary assessment, effective regulatory compliance, and successful digital transformation in nutritional management and business integration.

5.2 Limitations and Future Research Directions

The results indicate specific challenges in food category classification, particularly in groups like “eggs” and “unsaturated fats”, where semantic ambiguity and overlapping similarity scores reduce classification accuracy. This highlights the need for further refinement in category-level matching, especially for food names overlapping multiple categories. A potential improvement would be to introduce

finer-grained subcategories within broader food groups and use crosswalk files to systematically link them to higher-level categories. This could improve classification accuracy by allowing AI models to recognize context-specific variations.

A key limitation of this study is the reliance on FRIDA FoodEx2 codes as the validation dataset. As demonstrated in Table 3, a direct name match for boiled potatoes does not correspond to the expected code due to FRIDA's assignment of a distinct FoodEx2 code for the boiling process, despite the existence of a code for boiled potatoes. Further cleaning by removing or separating cooking processes could improve the matching. This discrepancy constrains validation accuracy, underscoring the need for a more comprehensive and systematically curated validation dataset. More broadly, this issue reflects fundamental challenges within food systems ontology, where inconsistencies in classification, granularity, and contextual interpretation can limit interoperability and data harmonization across different databases.

For FoodEx2 classification, a more effective strategy could also involve ranking the top five similarity-based matches instead of relying solely on the highest-scoring option. Presenting multiple candidate classifications would allow an AI-driven decision layer to assess contextual relevance and select the most appropriate match. This approach could improve classification robustness, particularly in cases where food names have slight variations due to preparation methods, processing techniques, or regional terminology differences. Further research could also explore reasoning models, few-shot approaches, and smaller models for local usage. Reasoning models and Few-shot learning approaches could help resolve semantic ambiguities by leveraging contextual understanding, improving classification in complex categories like “eggs” and “unsaturated fats”. Smaller, locally deployable models could enhance data security and privacy by processing sensitive food classification data on-device. This approach minimizes data transmission risks, ensures compliance with privacy regulations, and allows organizations to retain control over proprietary datasets. Further research should implement the LLM pipeline in operational canteens, with active involvement of stakeholders, for evaluating the pipeline adoption in institutional food services.

6 Conclusion

This study explored an LLM food classification pipeline to address technical and regulatory challenges in institutional food services. The developed LLM pipeline, combining embeddings, crosswalk mapping, and human validation, improves dietary assessment, data standardization, and regulatory compliance. While GPT-4o-mini achieved high accuracy at the dietary group level, precise FoodEx2 code matching remains challenging due to semantic ambiguity and naming variations. A confidence-based validation ensures transparency and accountability by balancing automation with human oversight, supporting expert decision-making. It demonstrates AI's potential to transform fragmented processes into integrated digital ecosystems through practical and ethical interventions. Future research should focus on evaluating the pipeline within operational canteen environments, engaging stakeholders directly, investigating LLM embeddings, reasoning models, and locally deployable AI models to enhance data privacy, nutritional quality, and sustainability outcomes.

References

- Agrawal, K. *et al.* (2025) 'AI-driven transformation in food manufacturing: a pathway to sustainable efficiency and quality assurance', *Frontiers in Nutrition*, 12. Available at: <https://doi.org/10.3389/FNUT.2025.1553942>.
- Bialas, C. *et al.* (2023) 'A Holistic View on the Adoption and Cost-Effectiveness of Technology-Driven Supply Chain Management Practices in Healthcare', *Sustainability* 2023, Vol. 15, Page 5541, 15(6), p. 5541. Available at: <https://doi.org/10.3390/SU15065541>.
- Ciatto, G. *et al.* (2025) 'Large language models as oracles for instantiating ontologies with domain-specific knowledge', *Knowledge-Based Systems*, 310, p. 112940. Available at: <https://doi.org/10.1016/J.KNOSYS.2024.112940>.
- Cudré-Mauroux, P. (2020) 'Leveraging Knowledge Graphs for Big Data Integration: The XI Pipeline', *Semantic Web*, 11(1), pp. 13–17. Available at: <https://doi.org/10.3233/SW-190371>.
- Danks, D. and Trusilo, D. (2022) 'The Challenge of Ethical Interoperability', *Digital Society* 2022 1:2, 1(2), pp. 1–20. Available at: <https://doi.org/10.1007/S44206-022-00014-2>.
- Dhal, S.B. and Kar, D. (2025) 'Leveraging artificial intelligence and advanced food processing techniques for enhanced food safety, quality, and security: a comprehensive review', *Discover Applied Sciences*, 7(1). Available at: <https://doi.org/10.1007/S42452-025-06472-W>.
- Dooley, D.M. *et al.* (2018) 'FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration', *npj Science of Food* 2018 2:1, 2(1), pp. 1–10. Available at: <https://doi.org/10.1038/s41538-018-0032-6>.
- Edwards, F., Sonnino, R. and López Cifuentes, M. (2024) 'Connecting the dots: Integrating food policies towards food system transformation', *Environmental Science & Policy*, 156, p. 103735. Available at: <https://doi.org/10.1016/J.ENVSCI.2024.103735>.

- EFSA (2015) 'The food classification and description system FoodEx 2 (revision 2)', *EFSA Supporting Publications*, 12(5), p. 804E. Available at: <https://doi.org/https://doi.org/10.2903/sp.efsa.2015.EN-804>.
- Elias, O. *et al.* (2024) 'The evolution of green fintech: Leveraging AI and IoT for sustainable financial services and smart contract implementation', <https://wjarr.com/sites/default/files/WJ-ARR-2024-2272.pdf>, 23(1), pp. 2710–2723. Available at: <https://doi.org/10.30574/WJARR.2024.23.1.2272>.
- Foidl, H. *et al.* (2024) 'Data pipeline quality: Influencing factors, root causes of data-related issues, and processing problem areas for developers', *Journal of Systems and Software*, 207, p. 111855. Available at: <https://doi.org/10.1016/J.JSS.2023.111855>.
- Gaitán-Cremaschi, D. and Valbuena, D. (2024) 'Examining purchasing strategies in public food procurement: Integrating sustainability, nutrition, and health in Spanish school meals and social care centres', *Food Policy*, 129, p. 102742. Available at: <https://doi.org/10.1016/J.FOODPOL.2024.102742>.
- Goel, M. and Bagler, G. (2022) 'Computational gastronomy: A data science approach to food', *Journal of Biosciences*, 47(1). Available at: <https://doi.org/10.1007/S12038-021-00248-1>.
- Gruber, T.R. (1995) 'Toward principles for the design of ontologies used for knowledge sharing?', *International Journal of Human-Computer Studies*, 43(5–6), pp. 907–928. Available at: <https://doi.org/10.1006/IJHC.1995.1081>.
- Harris, J. *et al.* (2025) 'Evaluating Large Language Models for Public Health Classification and Extraction Tasks'.
- Herforth, A. *et al.* (2019) 'A Global Review of Food-Based Dietary Guidelines', *Advances in Nutrition*, 10(4), pp. 590–605. Available at: <https://doi.org/10.1093/ADVANCES/NMY130>.
- Hu, G., Ahmed, M. and L'Abbé, M.R. (2023) 'Natural language processing and machine learning approaches for food categorization and nutrition quality prediction compared with traditional methods', *American Journal of Clinical Nutrition*, 117(3), pp. 553–563. Available at: <https://doi.org/10.1016/J.AJCNUT.2022.11.022>.
- Hua, A. *et al.* (no date) 'NUTRIBENCH: A Dataset for Evaluating Large Language Models on Nutrition Estimation from Meal Descriptions'. Available at: <https://mehak126.github.io/nutribench.html> (Accessed: 6 February 2025).
- Khedr, A.M. and S, S.R. (2024) 'Enhancing supply chain management with deep learning and machine learning techniques: A review', *Journal of Open Innovation: Technology, Market, and Complexity*, 10(4), p. 100379. Available at: <https://doi.org/10.1016/J.JOITMC.2024.100379>.
- Kolluri, S. (2024) 'Automating Data Pipelines with AI for Scalable, Real-Time Process Optimization in the Cloud', *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(6), pp. 2070–2079. Available at: <https://doi.org/10.32628/CSEIT242612405>.
- Kulal, A. *et al.* (2024) 'Enhancing public service delivery efficiency: Exploring the impact of AI', *Journal of Open Innovation: Technology, Market, and Complexity*, 10(3), p. 100329. Available at: <https://doi.org/10.1016/J.JOITMC.2024.100329>.
- Leme, A.C.B. *et al.* (2021) 'Adherence to Food-Based Dietary Guidelines: A Systemic Review of High-Income and Low- and Middle-Income Countries', *Nutrients*, 13(3). Available at: <https://doi.org/10.3390/NU13031038>.
- Li, Y. *et al.* (2023) 'Enhancing Health Data Interoperability with Large Language Models: A FHIR Study'. Available at: <https://arxiv.org/abs/2310.12989v1> (Accessed: 21 March 2025).
- Marvin, H.J.P. *et al.* (2022) 'Digitalisation and Artificial Intelligence for sustainable food systems', *Trends in Food Science & Technology*, 120, pp. 344–348. Available at: <https://doi.org/10.1016/J.TIFS.2022.01.020>.
- Michel, M. *et al.* (2024) 'Benefits and challenges of food processing in the context of food systems, value chains and sustainable development goals', *Trends in Food Science & Technology*, 153, p. 104703. Available at: <https://doi.org/10.1016/J.TIFS.2024.104703>.
- Min, W. *et al.* (2022) 'Applications of knowledge graphs for food science and industry', *Patterns*, 3(5), p. 100484. Available at: <https://doi.org/10.1016/J.PATTER.2022.100484>.

- Oluwaseun Badmus *et al.* (2024) ‘AI-driven business analytics and decision making’, *World Journal of Advanced Research and Reviews*, 24(1), pp. 616–633. Available at: <https://doi.org/10.30574/WJARR.2024.24.1.3093>.
- Popovski, G. *et al.* (2019) ‘FoodOntoMap: Linking Food Concepts across Different Food Ontologies’. Available at: <https://doi.org/10.5220/0008353201950202>.
- Popovski, G., Seljak, B.K. and Eftimov, T. (2020) ‘FoodOntoMapV2: Food Concepts Normalization Across Food Ontologies’, *Communications in Computer and Information Science*, 1297, pp. 413–426. Available at: https://doi.org/10.1007/978-3-030-66196-0_19/TABLES/5.
- Regal, T. and Pereira, C.E. (2018) ‘Ontology for Conceptual Modelling of Intelligent Maintenance Systems and Spare Parts Supply Chain Integration’, *IFAC-PapersOnLine*, 51(11), pp. 1511–1516. Available at: <https://doi.org/10.1016/J.IFACOL.2018.08.285>.
- Santhosh Bussa (2024) ‘Evolution of Data Engineering in Modern Software Development’, *Journal of Sustainable Solutions*, 1(4), pp. 116–130. Available at: <https://doi.org/10.36676/J.SUST.SOL.V1.I4.43>.
- Serrano-Santoyo, A. *et al.* (2021) ‘Ethical implications regarding the adoption of emerging digital technologies: An exploratory framework’, *Progress in Ethical Practices of Businesses: A Focus on Behavioral Interactions*, pp. 219–239. Available at: https://doi.org/10.1007/978-3-030-60727-2_12/FIGURES/6.
- Shirai, S.S. *et al.* (2021) ‘Identifying Ingredient Substitutions Using a Knowledge Graph of Food’, *Frontiers in Artificial Intelligence*, 3. Available at: <https://doi.org/10.3389/FRAI.2020.621766>.
- Sousa, G., Lima, R. and Trojahn, C. (2025) ‘Complex Ontology Matching with Large Language Model Embeddings’. Available at: <http://arxiv.org/abs/2502.13619> (Accessed: 19 March 2025).
- Taboada, M. *et al.* (2025) ‘Ontology Matching with Large Language Models and Prioritized Depth-First Search’. Available at: <http://arxiv.org/abs/2501.11441> (Accessed: 19 March 2025).
- Wei, Y. and Li, X. (2025) ‘Knowledge-enhanced ontology-to-vector for automated ontology concept enrichment in BIM’, *Journal of Industrial Information Integration*, p. 100836. Available at: <https://doi.org/10.1016/J.JII.2025.100836>.
- Wolfert, S. *et al.* (2023) ‘Digital innovation ecosystems in agri-food: design principles and organizational framework’, *Agricultural Systems*, 204, p. 103558. Available at: <https://doi.org/10.1016/J.AGSY.2022.103558>.
- Youn, J., Li, F. and Tagkopoulos, I. (2023) ‘Semi-Automated Construction of Food Composition Knowledge Base’. Available at: <https://github.com/ibpa/SemiAutomatedFoodKBC>. (Accessed: 6 February 2025).
- Zadeh, A.H. *et al.* (2018) ‘Cloud ERP systems for smalland-medium enterprises: A case study in the food industry’, *Journal of Cases on Information Technology*, 20(4), pp. 53–70. Available at: <https://doi.org/10.4018/JCIT.2018100104>.
- Zhang, J. *et al.* (2022) ‘Best practices in the real-world data life cycle’, *PLOS Digital Health*, 1(1), p. e0000003. Available at: <https://doi.org/10.1371/JOURNAL.PDIG.0000003>.

Appendix

Table 3: FRIDA food name embedding cosine similarity FoodEx2 matches

FRIDA Name	FoodEx2 Name 1	cosSim 1	Match 1	FoodEx2 Name 2	cosSim 2	Match 2	FoodEx2 Name 3	cosSim 1	Match 3
goat milk	goat milk	1,00	True	goat milk fat	0,84	False	yoghurt goat milk	0,78	False
potato boiled	potato boiled	1,00	False	potato baked	0,73	False	potatoes	0,63	True
semolina	semolina	0,99	False	wheat semolina	0,82	True	maize semolina	0,74	False
rice flour	rice flour	1,00	True	rice grain	0,77	False	rice rolled grains	0,72	False

DEVELOPING A VALUE SENSITIVE VR LEARNING ENVIRONMENT FOSTERING PROFESSIONALS' COMMUNICATION SKILLS

OLIVIER VAN DER MOLEN,¹ EMANUEL VAN DONGEN,²
STAN VAN GINKEL¹

¹ University of Applied Sciences Utrecht, the Netherlands
olivier.vandermolen@hu.nl, stan.vanginkel@hu.nl

² Utrecht University, Utrecht, the Netherlands
E.G.D.vanDongen@uu.nl

E-learning using virtual reality (also called v-learning) made it possible to create a 'scalable v-learning platform', enabling anyone without programming experience to create v- learning content for soft skill training. In the current project, Large Language Model algorithms are developed that help make it easier to create content and to generate more realistic and interactive virtual trainings. In order to study which human values and design principles are relevant for developing a skill training virtual reality learning environment, we used a Value Sensitive Design (Friedman, 1996) method in order to gather the human values of varying stakeholders. First, we studied values to be taken into account from both student and teacher perspectives, including educational and didactical values. These human values based on literature studies were supplemented by the perceptions of stakeholders, including commercial ones, and their human values in mutual interactions (empirical perspective). For this, an ethical matrix (comp. Van der Stappen & Van Steenberghe, 2020) was used as an instrument (for empirical research) for (a) identifying the harms and benefits of the use of the VR learning environment and (b) discussing essential human values in the development and implementation process. Finally, these two perspectives – one conceptual the other empirical – are brought together and some recommendations are offered for designing ethical, transparent, and effective digital learning environments that support both students and educators in their professional development.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.31](https://doi.org/10.18690/um.fov.4.2025.31)

ISBN
978-961-286-998-4

Keywords:
artificial intelligence,
virtual reality,
communication skills,
design principles,
e-learning



University of Maribor Press

1 Introduction

Higher education and professional training programs face increasing challenges, including rising student populations, heightened workloads for educators, and persistent shortages of teaching staff (Adubra et al., 2019). These issues contribute to reduced student-teacher interactions, limited opportunities for skill development, and potential declines in learning outcomes (van Ginkel et al., 2015). To address these challenges, digital learning environments powered by Artificial Intelligence (AI) have emerged as a promising solution, providing scalable, efficient, and interactive platforms for education and professional training. AI-driven virtual reality (VR) environments, in particular, offer immersive, adaptive, and realistic simulations that support skill acquisition in ways traditional methods often cannot (Bond et al., 2024). AI-enhanced digital learning environments have been shown to benefit student development across various domains, including science education (Lamb et al., 2021), language acquisition (Wei, 2023), and soft skill training (van Ginkel et al., 2019). By enabling personalized learning experiences, automating feedback, and providing real-time insights into learners' progress, these technologies can help alleviate the burden on educators while enhancing the quality of instruction.

However, the widespread integration of AI and VR (or broader: Extended Reality (XR)) in education (in a broad sense) has raised ethical concerns regarding autonomy of users or students, data privacy, and equitable access to learning resources (Van der Stappen & Van Steenbergen, 2020). These concerns are particularly relevant as AI-based tools become more autonomous in guiding learning processes, potentially affecting students' cognitive abilities, such as decision-making, analytical reasoning, and critical thinking (Zhai et al., 2024). Additionally, the use of AI-driven learning analytics to collect and interpret student data introduces questions about consent, data ownership, and fairness in educational settings (Ferguson, 2019). Further, the adoption of VR environments for learning brings challenges related to user autonomy and the potential manipulation of learning experiences (Skulmowski, 2023). While immersive simulations can enhance engagement and facilitate experiential learning, they may also inadvertently shape perceptions and behaviours in ways that warrant ethical scrutiny. These concerns highlight the need for a structured approach to integrating AI-driven VR tools in education—one that ensures alignment with educational goals while respecting the values and needs of diverse learners (Nguyen et al., 2023).

E-learning using virtual reality (also called v-learning) made it possible to create a 'scalable v-learning platform', enabling anyone without programming experience to create v-learning content for soft skill training. In a project financed by Kansen voor West, Large Language Model algorithms are developed by commercial companies that help make it easier to create content and to generate more realistic and interactive virtual trainings (project called 'AI generated v-learning content'). This study explores the ethical considerations and design principles for creating such a platform, aimed at enhancing professional communication skills. By adopting a Value Sensitive Design (VSD) approach (Friedman & Hendry, 2019), we aim to identify the key concerns and expectations of students, teachers, and professional trainees regarding the use of these technologies, in particular a scalable v-learning platform. Through two phases of our VSD-approach, a conceptual and empirical phase, we assess the potential benefits and risks associated with AI-enhanced learning, gathering insights from i) a (educational and technological) literature study and ii) structured workshops. To structure our empirical analysis, we employed the ethical matrix methodology (Van der Stappen & Van Steenbergen, 2020), enabling us to systematically capture the perspectives of different stakeholders on the potential harms and benefits of AI-driven VR environments. This contribution will conclude by offering recommendations for designing ethical, transparent, and effective digital learning environments that support both students and educators in their professional development. In doing so, the findings from this study contribute to the broader discourse on responsible AI and Virtual Reality use in education.

The study addresses the following research questions:

- (1) Which human values are essential for stakeholders in AI-driven VR learning environments?
- (2) How can these identified values effectively inform the design of ethically-aligned VR learning platforms?

2 Theoretical framework

2.1 Conceptualizing this study

This study investigates how a v-learning platform can foster professional communication skills while accounting for ethical considerations and stakeholder perspectives. As mentioned before, we used the framework of Value Sensitive Design (VSD), which has been widely used in various domains, including information and communication technology, biotechnology, sustainability, and healthcare (Friedman et al., 2006). Despite its broad applications, VSD remains relatively underexplored in the context of professional post-higher education training and digital learning environments in higher education. Recently, the ethical matrix, a tool used to support VSD, has been used by the University of Applied Sciences Utrecht in digital innovation projects to assess stakeholder perceptions on ethical values in education (Van der Stappen & Van Steenberghe, 2020). These theoretical insights have been used as a guide for our current study and for the following section on the VSD framework and the ethical matrix methodology.

2.2 Technological Advancements in Learning

Experiential learning, a pedagogical approach emphasizing active engagement, reflection, and real-world application (Kolb, 2014), has been significantly transformed by advancements in artificial intelligence (AI) and virtual reality (VR). These technologies offer immersive, adaptive, and data-driven learning experiences that enhance skill development across educational and professional training contexts (Makransky & Petersen, 2021). By simulating complex environments, VR enables learners to practice communication, problem-solving, and decision-making in realistic yet controlled settings, bridging the gap between theoretical knowledge and practical application (Radianti et al., 2020).

AI-driven learning environments further augment experiential learning by providing personalized feedback, intelligent tutoring, and automated assessment (Luckin, 2017). Such enhancements support self-directed learning and continuous improvement, making AI-powered VR tools particularly effective for skill-intensive domains, including professional communication and interpersonal training (Gopalan et al., 2023). Practice and feedback are some of the foremost aspects, aided

by the VR technology, that aid in the education of soft-skills (Van Ginkel et al., 2019; Van Dongen 2024).

Despite these advantages, the ethical implications of AI and VR in experiential learning remain a growing concern. Issues such as data privacy, algorithmic bias, and the potential over-reliance on technology challenge the responsible deployment of these tools in educational settings (Zhai et al., 2024). Moreover, questions about the impact of immersive simulations on learners' cognitive load and emotional well-being necessitate further investigation (Parsons, 2021). A recent exploratory study in law showed significant increases in heart rate and electrodermal activity occurred from rest into the activity of presenting in VR (called 'pleading'; Van Dongen 2024). These described risks need to be managed with a safe and holistic approach, where an overview of information and priorities such as intrinsic stakeholder values regarding the technology play a prominent role. The design and implementation of AI-driven VR learning environments must therefore adopt a value-sensitive approach, ensuring that technological advancements align with educational goals while safeguarding ethical considerations (Nguyen et al., 2023).

2.3 The concept of Value Sensitive Design

Value Sensitive Design (VSD) is an approach that integrates human values throughout the design process, ensuring that ethical considerations are embedded in technological development (Friedman et al., 2006). Rather than focusing solely on usability and functionality, VSD examines the broader societal and ethical implications of new technologies (Flanagan et al., 2008). Within this framework, values are defined as 'what is important to people in their lives, with a focus on ethics and morality' (Friedman & Hendry, 2019, p. 24).

VSD identifies four categories of stakeholders whose values should be accounted for: (1) the sponsors or funders of the design, (2) the project team developing the technology, (3) direct stakeholders who actively use the technology, and (4) indirect stakeholders who may be affected by the technology without directly using it. Understanding the impact on indirect stakeholders requires forward-looking and adaptive design thinking, considering both the long-term effects and the potential for unintended consequences as AI-driven tools expand in use (Nathan et al., 2008). In this study, indirect stakeholders specifically include institutional IT support staff,

policymakers involved in educational technology regulation, and future students indirectly impacted by technology adoption.

The VSD approach is applied through an iterative process involving two perspectives: both conceptual and empirical. At the conceptual level, stakeholders and relevant values are identified based on existing literature and theoretical considerations. The empirical level focuses on gathering stakeholders' perspectives using methods such as focus groups, interviews, or surveys, allowing for the identification of tensions between different values. The combination of these perspectives translates these insights into design strategies and recommendations for further implementation, ensuring that identified values are reflected in the system's functionality and user experience on long-term. These stages interact dynamically, allowing for continuous refinement and adaptation as insights emerge (Friedman & Hendry, 2019). Over time, VSD has been successfully applied in diverse domains, including web security (Friedman et al., 2002), renewable energy technologies (Oosterlaken, 2015), AI system design (Umbrello & van de Poel, 2021), and music copyrights (Nerko et al., 2024).

3 Methodology

3.1 A method for Value Sensitive Design

To explore the human values and design principles relevant to the development of a virtual reality learning environment for fostering communication skills, the VSD approach was adopted (Friedman, 1996). The study was conducted through two experimental workshops involving students, teachers, and professional trainees, utilizing the ethical matrix (Van der Stappen & Van Steenberg, 2020) as a structuring tool for gathering and analyzing data (see figure 1). The ethical matrix is used to identify possible harms and benefits of the technical subject on its direct and indirect stakeholders, by creating a discussion between creators and users to foster transparent conversations for the development of the technology.

In the conceptual research, a systematic literature review was conducted to identify and categorize human values relevant to AI-driven VR learning environments based on other cases. The review followed the structured approach outlined by Fink

(2014), covering values applicable to educational, professional, and commercial contexts, ensuring a broad perspective. The selection criteria included peer-reviewed studies that (a) examined the role of digital learning environments in communication skill training, (b) discussed ethical and human value considerations within these training technology contexts, and (c) provided empirical insights into AI and VR implementation in education and professional training. The identified values from this literature review were not discussed during the empirical research to prevent participant biases.

Table 1: The ethical matrix template

	Value 1	Value 2	...
Stakeholder group 1	Harms and benefits	Harms and benefits	
Stakeholder group 2	Harms and benefits	Harms and benefits	
...			

Source: Van der Stappen & Van Steenberghe, 2020

The empirical phase involved two experimental VSD-workshops, each designed to elicit insights from three stakeholder groups: students, teachers, and professional trainees. The objective was to examine how these groups perceive the benefits and potential harms of the AI-driven VR training program to their specific roles and to identify values that should be considered in its development and implementation to adapt to the stakeholders' intrinsic values.

3.2 Research phases and data collection

Each VSD workshop, conducted by two independent researchers, consisted of three phases:

Introduction and Familiarization: Participants were introduced to the AI-driven VR training program and the study's objectives. The AI tool was shown, a conversation took place between the creator and the program and afterwards, questions were answered. They were briefed on the ethical matrix as a tool for structuring their reflections and the definitions used within these methods for specific intrinsic values.

Ethical Matrix Exercise: Participants independently identified potential harms and benefits of the VR learning environment based on their role, writing these aspects down on post-it notes. These post-its were either written on green notes identifying the benefits, where harms are written on red notes. Harms were defined by the risk or potential damage the software could have in its implementation context, where benefits represented the potential positive impact.

Focus Group Discussion: A facilitated discussion followed, where participants elaborated on their identified concerns and values, engaging in a structured dialogue to refine their perspectives, where open, non-steering questions were used such as: ‘Why does this seem like a harmful aspect to you?’ or ‘Why could this be a benefit and why is this important to you?’. Audio recordings were made to aid in the translation from the ethical matrix towards the intrinsic values. The combination of both audio recordings and the filled ethical matrix as outcome of the workshops were then used to summarize the most-noted values and quotes, highlighting the specific harms or values from which the values were derived. The conversations held in both groups were facilitated by researchers with deep understanding of VSD and the Ethical Matrix. One of these researchers conducted the decoding of the data of both conversations to the paper’s findings. Another researcher checked the paper findings.

3.3 Participants and privacy

A total of 18 participants were recruited across the three stakeholder groups, with six representatives from each category, ensuring balanced perspectives in the further discussions. Ethical approval was obtained from the University of Applied Sciences Utrecht Research Ethics review board, and all participants signed an informed consent before participation. The audio recordings were deleted after having been used as a check to the notes made during the workshops. Furthermore, in VSD terminology, participants were referred to as ‘stakeholder’, due to the nature of their relationship to the product, where for this study the term ‘participant’ is used.

4 Findings

4.1 Introduction

This chapter presents the findings derived from i) the literature study on values before the stakeholder meetings (“workshops”) were held; ii) the stakeholder analysis based on various key values as discussed during the beforementioned VSD-workshops. Each value is briefly defined based on our literature study, followed by an overview of both positive and negative perspectives from different stakeholder groups.

4.2 Conceptual

Based on the literature review, we identified fourteen key values: efficiency, autonomy, human connection, competency, quality, accessibility, flexibility, adaptability, authenticity, safety, inclusion, enjoyment, effectiveness, and ethics. As outlined in the methodology section, these values stem from a wide range of studies on digital learning environments in different contexts (Yulian et al., 2022, Kent, 2021, Lenca et al., 2023). Importantly, the process required translating reported harms and benefits in these studies into overarching values — a subjective exercise that involved interpreting the intended meaning from each source. Consequently, these values cannot be directly attributed to any one specific tool or platform. Nonetheless, they serve as a useful framework for comparing and contextualizing the focus group findings, offering an initial, exploratory sense of how broadly applicable those findings might be. These values illustrate significant trade-offs in deploying VR technologies: while they enhance efficiency and autonomy, they can simultaneously raise issues such as increased cognitive load or reduced interpersonal interactions.

Many of the values identified in the literature were associated with both advantages and drawbacks. For instance, didactic quality can be enhanced through technology’s ability to present material in innovative ways or support diverse teaching strategies. Yet, the same quality can be compromised if technological solutions are rigid or fail to account for individual learner differences or contextual nuances. Similarly, human connection may be strengthened by promoting peer interaction and collaborative learning but could also suffer from reduced teacher-student engagement. Autonomy

is another value that shows duality: while technology can empower learners to independently explore content, it may also foster dependency on these digital tools. Interestingly, four values were consistently described in the literature as offering only benefits: safety, enjoyment, effectiveness, and inclusion. Inclusion was noted in the context of voice-user interfaces, enabling better access for students with difficulties using keyboards or traditional search engines. Safety was highlighted as learners could practice skills without real-world repercussions, helping them feel more secure and confident before applying those skills in actual settings. Enjoyment was linked to increased student engagement, particularly when technology introduced novel ways of presenting content or offered greater diversity in learning activities. Lastly, multiple studies reported that educators observed clear learning gains among students using these digital learning tools (McLaren et al., 2017. Wang & Lieberoth, 2016).

4.3 Empirical

Multiple core values – relating to the virtual reality learning environment – were recurring within the two workshop groups. Although the values were labelled differently by each of the workshop groups, they were clustered based on their given definitions on the harms and benefits, resulting in the following outcomes:

As shown in the Ethical Matrix, the value Reputation is a crucial factor in how the tool is perceived in terms of credibility and professionalism. Trainers, teachers, and learning coordinators emphasized that the tool must maintain a serious and polished appearance, avoiding any design elements that could make it look cheap or unprofessional. This concern was mirrored by students, who expressed worries that using the tool might impact how seriously they are taken in their learning environments. Furthermore, the perception of professionalism ties directly into the effectiveness of the tool, as users are more likely to engage with a system they deem reliable.

Beyond its Reputation, the participants mention that the tool offers New Learning Experiences, which were recognized as a key advantage. Trainers, teachers, and learning coordinators highlighted its ability to introduce engaging and innovative learning concepts that go beyond traditional methods, using a broad scala of even the most recent information. Students found the experience novel and stimulating,

while trainees, professionals, and users appreciated how it allowed for further research into evolving learning experiences. However, it was also emphasized that the novelty of these experiences must be balanced with ensuring that the tool does not feel gimmicky or lack depth in its educational offerings.

Table 2: The combined ethical matrix outcome of the two workshop groups

	Reputation	New Learning Experiences	Safety	Human Connection	Scalability & Impact	Quality	Accuracy	Transparency	Ease of Use	Sustainability
Trainer / Teacher / Learning Coordinator	Should appear serious, should not look cheap	Concept is interesting and offers new learning experiences	Must be well secured, safe practice with new behaviour	Need for face-to-face moments	Ability to reach a broader target audience	Realistic view is important, constant support and updates are crucial	I'm afraid the tool might work against me	Black box: can it be explained?		Already uses a lot of energy
Student	Not taken seriously as a student	New experience	students feel vulnerable and can feel unsafe in practice	Reduction of human interaction	Students feedback: does it really work?	Should align with existing teaching methods	AI does not do exactly what I mean (loophole)			A lot of data — how do we store it?
Trainee / Professional / User		Research into learning experiences				Effectivity should be scientifically proven to make sure to receive consistent results over time			Why feedback only later? Prefer immediate feedback	Criteria: sensitivity and selection rules

Source: data collection

Participant: “The conversations look really well and in-depth, but I do see this uncanny valley form with these AI avatars looking really fake and creepy. I doubt people will take this seriously in the long term unless we move out of this uncanny valley.”

Safety, both in terms of data security and emotional comfort, was another major consideration of the participants. Distinct views can be distinguished. Trainers, teachers, and learning coordinators underscored the importance of ensuring a well-secured system that provides a supportive environment for practicing new behaviours. Students however reported feeling vulnerable while using the tool, with some expressing discomfort in how it handled their interactions. This highlights the tension between innovation and user comfort—while the tool provides novel learning opportunities, it must also ensure a sense of security to encourage full engagement.

A discussion between participants showed both positive and negative impact on the value of safety through this digital learning environment:

Participant 1: “Usually, when there are training actors involved, one student stands in front of the group to do the exercise. The rest of the group learns from the observation of this as well.”

Participant 2: “Yeah, so there is safety in being able to watch from the sidelines for the first time.”

Yet other participants also highlighted the secondary effect on safety:

Participant 3: “But there is also safety in being able to practice on your own at home, instead of in the presence of your peers.”

This discussion shows the ethical challenge of both the positive and the negative implication on Safety, throughout the creation of a time and place independent learning environment. This conversation fueled other values later described.

In the discussions, the balance between Personal Interaction and Digital Automation within the learning process was a recurring theme. While trainers, teachers, and learning coordinators stressed the necessity of maintaining face-to-face moments in education, students voiced concerns that the tool might diminish Human Interaction, making learning feel less personal. This contrast underscores a key challenge: while automation can enhance accessibility and scalability, it must not come at the cost of meaningful human engagement.

Participant: “There is value on being able to practice on your own. But on the other side, I do see that people do not act collectively anymore. Even learning is now individualized.”

Speaking of Scalability & Impact, trainers, teachers, and learning coordinators saw great potential in reaching a broader audience through the tool. However, students were more sceptical, questioning whether the tool’s approach would be as effective in real-world learning environments. This divide raises an important point about implementation—while scalability is advantageous, it must be paired with adaptability to ensure effectiveness across different user groups.

When it came to quality of the virtual environment, stakeholders placed high importance on the tool’s Effectiveness, Realism, and Reliability. Trainers, teachers, and learning coordinators insisted on a realistic interface with consistent support and updates. Meanwhile, students emphasized the need for alignment with existing teaching methods, and trainees, professionals, and users highlighted the necessity of scientific validation and long-term consistency. However, ensuring high quality may present challenges in terms of Accuracy, which was another significant concern which came up during the workshops. While a realistic interface is essential, trainers, teachers, and learning coordinators were wary of the potential for misleading or incorrect outputs:

Participant: “A lot of these AI-assistants don’t seem to provide consistent results over a longer conversation. They start dreaming and moving out of bounds of the prompted instructions. This could become an issue as the tool starts to respond differently.”

Students similarly noted inconsistencies in AI-driven functionalities, which sometimes failed to behave as expected or misinterpreted their inputs. Thus, while striving for quality, developers must also prioritize precision to avoid user frustration and misinformation.

Similarly, Transparency emerged as a crucial consideration. Trainers, teachers, and learning coordinators voiced concerns that the tool might function as a "black box," making it difficult for users to understand how it arrives at its conclusions. This concern ties closely with accuracy—if users cannot comprehend the logic behind the tool's outputs, trust in its accuracy may diminish. Therefore, providing explanations for how decisions are made could enhance both trust and engagement.

Another major factor was Ease of Use, particularly regarding accessibility and responsiveness. Trainees, professionals, and users expressed frustration with delayed feedback, stating a preference for immediate responses that could enhance their learning experience. This need for efficiency further reinforces the demand for accuracy and transparency. Furthermore, quick feedback must also be correct and clearly explained.

Finally, Sustainability was discussed in terms of both environmental impact and data management. Trainers, teachers, and learning coordinators noted that the tool already consumes significant energy, raising questions about its long-term viability:

Participant: "Every day more datacentres are being built for all this AI. I saw that they use a lot of energy and water, is this not something we need to account for when we want to broadly implement these tools?"

Students questioned how large datasets should be stored efficiently, while trainees, professionals, and users emphasized the importance of establishing clear criteria for data selection and storage to ensure responsible management. While sustainability may seem like a separate issue, it is closely tied to the tool's long-term effectiveness—without efficient data handling and energy-conscious design, its usability and credibility may be compromised over time.

5 Conclusion

This study enhances the transparency of discussions on ethical values regarding v-learning context of an human-AI communication technology to incorporate ethics by design on the development of this technology.

The outcomes of both the literature and the empirical perspectives, revealed that participant values regarding this v-learning context, such as safety and inclusion, are foundational to user acceptance. Safety, encompassing both emotional comfort and data security, was consistently highlighted as crucial, particularly given the vulnerability participants felt when interacting with AI avatars. Inclusion, exemplified through accessibility features like voice-user interfaces, is important as well as it ensures that learners of all backgrounds and abilities can participate meaningfully. Enjoyment and effectiveness were also prominent: stakeholders noted that when learning tools are engaging and demonstrably effective, they foster deeper learning and motivation.

However, the current study also highlighted the inherent duality of certain values. For instance, while autonomy and flexibility are beneficial in promoting independent learning, they risk reducing human connection and collaborative learning experiences. Similarly, scalability and automation, though valuable for reaching larger audiences, may inadvertently diminish the personal touch and lead to perceptions of detachment or reduced educational quality.

Caution is therefore needed when balancing these values. The tension between realism and accuracy was particularly evident. While stakeholders appreciated realistic interfaces, they cautioned against inconsistencies in AI outputs that could undermine credibility. Transparency is another area requiring careful design; without clear explanations of system decisions, users may lose trust in the technology. Moreover, concerns about environmental sustainability and data storage were raised, highlighting that technological innovation must not come at the expense of long-term viability.

This study, while rich in qualitative data, was limited in scale, with only 18 participants. Broader quantitative research would provide stronger validation of these findings. Additionally, the study focused on perceptions prior to full-scale

implementation; longitudinal studies are needed to assess how these values evolve with prolonged use. Finally, cultural factors and disciplinary differences were not deeply explored and may influence value prioritization.

Future research could address these limitations by expanding participant diversity and employing longitudinal study designs to observe how perceptions on harms and benefits, and even values, can change during the process. Furthermore, more research can be conducted to explore adaptive interfaces that dynamically balance competing values such as autonomy versus guidance, or scalability versus personalization. Practical implications of this study include incorporating adaptive feedback mechanisms, embedding transparency features to clarify AI decision-making, and developing emotionally responsive design elements to address user comfort.

Furthermore, given the identified risks such as manipulation and cognitive impact, a comparison with other methods such as Value Based Engineering (VBE) could provide more insight into the risks of specific values in these contexts.

The VSD methodology using the ethical matrix showcases the overview that can be created based on stakeholder values, providing information on priorities towards the development of v-learning technology. Using these values as a roadmap for further development and policy development regarding aspects as ethics, autonomy and more could provide a significantly more suitable practical v-learning environment than without the VSD approach.

References

- Adubra, A. L., George, R., & Duveskog, M. (2019). Addressing teacher shortages in Sub-Saharan Africa: Policies, practices, and lessons learned. *International Review of Education*, 65(1), 147–166.
- Bond, M., Zawacki-Richter, O., & Nichols, M. (2024). Trends in digital education: A global systematic review. *Educational Technology Research and Development*, 72(1), 12–28.
- Ferguson, R. (2019). Ethical issues in learning analytics: Implications for the future. *British Journal of Educational Technology*, 50(6), 2784–2794. <https://doi.org/10.1111/bjet.12883>
- Flanagan, M., Howe, D. C., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. In *Information technology and moral philosophy* (pp. 322–353). Cambridge University Press.
- Friedman, B. (1996). Value-sensitive design. *Interactions*, 3(6), 16–23.
- Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press.

- Friedman, B., Kahn, P. H., Jr., & Borning, A. (2006). Value sensitive design and information systems. In *Human-computer interaction and management information systems: Foundations* (pp. 348–372). M.E. Sharpe.
- Gopalan, N., Puthussery, A., & Claiborne, T. (2023). AI-enhanced experiential learning: Transforming soft skill training in digital environments. *Educational Technology Research and Development*, 71(2), 445–463. <https://doi.org/10.1007/s11423-023-10123-8>
- Kent, M. (2021). Voice-user interfaces for TESOL: Potential and barriers. *TESOL Journal*, 12(1), e00516. <https://doi.org/10.1002/tesj.516>
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development* (2nd ed.). Pearson Education.
- Lamb, R., Akmal, T., & Petrie, K. (2021). The role of AI in science education: A systematic review. *Journal of Science Education and Technology*, 30(2), 123–139.
- Makransky, G., & Petersen, G. B. (2021). The cognitive and affective benefits of immersive virtual reality in education. *Educational Psychology Review*, 33(4), 1535–1561. <https://doi.org/10.1007/s10648-020-09557-3>
- McLaren, B. M., Adams, D. M., Mayer, R. E., & Forlizzi, J. (2017). A computer-based game that promotes mathematics learning more than a conventional approach. *International Journal of Game-Based Learning*, 7(1), 36–56.
- 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>
- Oosterlaken, I. (2015). Applying value sensitive design (VSD) to wind turbines and wind parks: An exploration. *Science and Engineering Ethics*, 21(2), 359–379.
- Parsons, S. (2021). Immersive virtual reality and learning: A critical reflection on the potential of VR for education. *Frontiers in Education*, 6, 653234. <https://doi.org/10.3389/feduc.2021.653234>
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147, 103778. <https://doi.org/10.1016/j.compedu.2019.103778>
- Skulmowski, A. (2023). Ethical implications of immersive technology in education. *Educational Technology & Society*, 26(2), 1–12.
- Umbrello, S., & van de Poel, I. (2021). Mapping value sensitive design onto AI for social good principles. *AI and Ethics*, 1(3), 283–296.
- Van der Stappen, E., & Van Steenbergen, N. (2020). Ethics by design: Practical applications in educational technology. *Journal of Educational Innovation*, 2(1), 45–62.
- E.G.D. van Dongen, 'Pleading in the Virtual Courtroom: Exploring Experiential Learning in Law through Virtual-Reality-Based Exercises and Student Feedback', *European Journal of Legal Education* 5 (2024), 1, p. 157-190.
- Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2015). Towards a set of design principles for developing oral communication skills in higher education. *Learning Environments Research*, 18(4), 489–516.
- Van Ginkel, S., et al. (2019). The impact of VR on oral communication training in higher education. *Computers & Education*, 134, 78–97.
- Wei, L. (2023). AI applications in language learning: A review of recent advancements. *Language Learning & Technology*, 27(1), 1–20.
- Wang, A. I., & Lieberoth, A. (2016). The effect of points and audio on concentration, engagement, enjoyment, learning, motivation, and classroom dynamics using Kahoot! In *Proceedings of the European Conference on Games Based Learning* (pp. 738–746).
- Yulian, F., Ginting, R., & Zulfikar, Z. (2022). EFL slow learners' perception in speaking with avatars. *International Journal of Emerging Technologies in Learning*, 17(4), 147–160. <https://doi.org/10.3991/ijet.v17i04.27369>

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>

ENERGY BEHAVIOR AND SMART HOME SYSTEMS: A BIBLIOMETRIC ANALYSIS

XIAO PENG,¹ XIUZHU ZANG,^{2,3} MENGQIU DENG⁴

¹ University of Applied Sciences Utrecht, Utrecht, the Netherlands

xiao.peng@hu.nl

² China University of Mining and Technology, School of Economics and Management, Jiangsu, China

ts21070071a31ld@cumt.edu.cn

³ China University of Mining and Technology, The Carbon Neutrality and Energy Strategy Think Tank, Jiangsu, China

ts21070071a31ld@cumt.edu.cn

⁴ Zhejiang University, School of Energy Engineering, Zhejiang, China
mengqiu_deng@zju.edu.cn

This study explores how households interact with smart systems for energy usage, providing insights into the field's trends, themes and evolution through a bibliometric analysis of 547 relevant literature from 2015 to 2025. Our findings discover: (1) Research activity has grown over the past decade, with leading journals recognizing several productive authors. Increased collaboration and interdisciplinary work are expected to expand; (2) Key research hotspots, identified through keyword co-occurrence, with two (exploration and development) stages, highlighting the interplay between technological, economic, environmental, and behavioral factors within the field; (3) Future research should place greater emphasis on understanding how emerging technologies interact with human, with a deeper understanding of users. Beyond the individual perspective, social dimensions also demand investigation. Finally, research should also aim to support policy development. To conclude, this study contributes to a broader perspective of this topic and highlights directions for future research development.

DOI

<https://doi.org/10.18690/um.fov.4.2025.32>

ISBN

978-961-286-998-4

Keywords:

energy behavior,
smart home systems,
bibliometric analysis,
interdisciplinary work,
social dimensions



University of Maribor Press

1 Introduction

The global energy transition is accelerating as societies shift from fossil fuels to more sustainable energy sources, and it is aided by rapid digital and technological advancements (K. Khan & Su, 2023; Dekeyrel & Fessler, 2024). Among these technology innovations, smart home systems are practical tools that help optimize energy efficiency and reduce overall consumption for energy end users (Wilson et al., 2015). However, despite their potential, the social dimensions of smart energy systems have been largely neglected in mainstream energy transition agendas. Robison et al. (2023) argue that social sciences and humanities (SSH) should take a leading role in addressing this gap by examining how individuals interact with and benefit from technologies to fasten energy transition.

Understanding individual behavior is crucial for societal transitions such as the energy transition, where individuals serve as drivers, beneficiaries, and adaptors of change (Whitmarsh, 2011). We believe that digital technologies, particularly smart home systems, play an important role in shaping behavioral shifts by not only enhancing energy efficiency but also fostering pro-environmental behavior (PEB) (Froehlich, 2009; Nilsson et al., 2018; Stieglitz et al., 2023). Smart systems enable real-time behavior tracking, allowing for immediate adjustments, such as running a washer when energy prices are low either by households themselves or by automated system controls (Costanza et al., 2014; M. Khan et al., 2016). Smart systems also can enable the implementation of Just-in-Time Adaptive Interventions (JITAI), such as sending real-time notifications about low energy availability or sharing social norms on neighborhood energy consumption to encourage adaptive user behavior (Nilsson et al., 2018). Finally, smart systems facilitate connection with others, fostering awareness that collective efforts are being made toward the energy transition – you are not alone (Froehlich, 2009). An often-heard problem in sustainability is the persistent belief that others may not feel as deeply about the environment as you yourself, which may diminish the likelihood for efficient or sustainable energy consumption (Gadenne et al., 2011). By making PEB visible, smart systems can reinforce social norms and encourage individuals to adopt energy-efficient practices. Additionally, users who share their energy usage online may feel social pressure to adopt energy-efficient behaviors (Froehlich, 2009).

Smart systems have also as benefit that they can be designed with more general (behavioral) rules, with the opportunity for its user to personalize the environment to meet his or her own preferences or needs. The latter is particularly important, as some users may greatly prefer high granulation of available data and options, whereas others want to keep it as simple as just seeing a price (Risteska Stojkoska & Trivodaliev, 2017). Moreover, visual design also influences household energy behavior to varying degrees, depending on individual preferences (Froehlich, 2009). Further, a great advantage is the ability to track behavior for a longer period. This does not only allow for measurement of lasting behavior change but also allows for follow-up interventions when effects seem to dissipate. Smart systems can incorporate framing and nudging strategies – such as demonstrating the specific impact of energy consumption reduction on CO₂ emissions or gamifying energy-saving goals with challenges and rewards – to educate end users, increase awareness, encourage desirable behaviors and serve as long-term interventions for behavior change (Kroll et al., 2019; Stieglitz et al., 2023). With the question in mind, “*What is the role of smart home systems in shaping household energy behavior?*” we conduct a literature analysis to explore the interaction between individuals and technology within the context of energy conservation at home. The following chapters outline our methodology, present the results, and propose directions for future research.

2 Data Acquisition and Bibliometric methods

We systematically retrieved relevant publications from the Web of Science (WOS) database, which is highly aligned with our research field through its SCI/SSCI indexing system, and conducted comprehensive bibliometric analyses, including descriptive statistics, word frequency analysis with cloud visualization, and keyword co-occurrence network mapping, to investigate the dynamic evolution of research hotspots in this field.

2.1 Data Acquisition and Screening

To find all relevant articles, we defined the following search strategy. The keywords consist of a wide range of synonyms to ensure the inclusion of all records that might be relevant to the topic. Below is the complete search term:

(digit OR smart* OR ICT* OR tech* OR "Internet of Things") AND*

(energ OR resource* OR gas OR water OR electric* OR solar OR wind OR heat) AND*

(efficien OR conserv* OR consum* OR use OR usage OR using OR reduc* OR sav* OR curtail* or behavio?r* OR prod* OR invest*) AND*

(consumer OR house OR commun* OR collect* OR group OR citi? OR pers* OR individual)*

We filtered only peer-reviewed articles to maintain an academic perspective, and they are written in English. The publication years were set between 2015 and 2025 to provide an overview of the last decade. Additionally, we focused on studies from the following fields, which we considered the most relevant: "environmental sciences ecology" or "water resources" or "business economics" or "psychology" or "urban studies" or "sociology" or "behavioral sciences" or "communication" or "public administration" or "social sciences other topics" or "social issues" or "demography". In total, 11,712 articles were retrieved from the Web of Science as of February 10, 2025.

Systematically screening 11,712 articles to identify publications relevant to the main subject of this review is a time-consuming and challenging task. Large language models, an emerging technology capable of automating language-related tasks, have been applied in literature reviews in some fields to screen titles and abstracts (Dennstädt et al., 2024). This paper adopted this approach, developing a prompt for screening titles and abstracts in line with the theme of this literature review. The process utilized Python to call the OpenAI ChatGPT-4o model via an API, helping to reduce labor costs. To refine the screening process, we conducted a self-scanning of 700 articles, iteratively improving the prompt to enhance accuracy and evaluation. The prompt incorporated inclusion criteria such as *"studies that explore smart devices, systems, meters, automation, or feedback mechanisms in home settings influencing energy-related human behavior"*. It also specified exclusion criteria such as *"if a study is mainly technical (>90% focused on engineering, algorithms, sensors, or optimization), and its discussion of human behavior is minimal, secondary, or incidental, it should be excluded"*. Additionally, we defined *human behavior refers to household energy consumption or production actions, behavioral interventions, lifestyle adjustments, or PEB influencing home energy use. It does not include technology adoption studies, policy evaluations, or studies focused solely on technical efficiency*. We compared the results of GPT-based screening with manual screening results and

verified the feasibility of this approach. Using this approach, 547 papers related to energy behavior and smart home systems were selected for further analysis.

2.2 Bibliometric Methods

Bibliometrics is a method for evaluating literature data, serving as a quantitative scientific assessment approach. It systematically collects, organizes, and analyzes academic literature data to reveal the development trajectory, research hotspots, and future trends of a discipline.

In this study, we conducted both external feature analysis and content analysis of the literature. External feature analysis forms the descriptive basics of bibliometric research, encompassing statistical analysis of data such as authors, institutions, countries, publication years, and journals. Content analysis, the core of bibliometric research, includes methods such as keyword co-occurrence analysis, and topic identification and evolution analysis. These approaches uncover research hotspots and provide critical insights into the evolving focus areas of energy behavior and smart home systems. We selected several analyses for this paper, and the details are outlined below:

(1) Descriptive Statistical Analysis:

We first conducted a systematic statistical analysis of the years, journals of the publications to gain insights into their trend and their academic impact and relevance to our research topic. This involved a quantitative assessment of number of publications per year, journal, measuring the productivity. Additionally, we examine the distribution across different journals and check the top journals quality, considering factors such as impact factor and journal quartile rankings (Q1–Q4) to evaluate their academic influence. This approach allowed us to identify the most influential journals in the field and understand their role in shaping the discourse on the topic.

(2) Word Frequency and Word Cloud

We used word frequency analysis to examine titles and abstracts, identifying the most frequently occurring terms in the research articles. This method helps highlight key themes that researchers focus on within the topic. By analyzing word frequency, we can determine the most researched concepts, emerging trends, and commonly discussed keywords, providing insights into the academic landscape.

To visually represent these findings, we generated word clouds, a widely used data visualization tool that displays words in different sizes based on their frequency of occurrence. The larger a word appears in the cloud, the more frequently it is mentioned in the dataset. This approach allows for an intuitive and immediate understanding of dominant research topics and their relative importance.

(3) Keyword Co-occurrence Analysis:

Using Python, we preprocessed the literature text data, including tokenization, stop-word removal, and lemmatization. The TF-IDF (Term Frequency-Inverse Document Frequency) algorithm was then applied to quantify keyword importance and filter core keywords for each article. By constructing a keyword co-occurrence matrix, we calculated the co-occurrence frequency and association strength between keywords, revealing the intrinsic connections among research themes. Moreover, we extracted latent topics from the titles and abstracts of publications across two distinct stages: exploration stage (2015-2020) and development stage (2021-2025). The topic identification for each stage highlights the evolution of the research agenda over the past decade.

Further, Gephi network analysis tools were used to transform the co-occurrence matrix into a keyword co-occurrence network graph. Through node size, edge weight, and community clustering visualization, we intuitively displayed the relationships and clustering characteristics of keywords.

3 Results

The research findings provide a detailed presentation of multiple analytical dimensions, including the timeline of publications, author productivity and collaboration patterns, journal statistics, word frequency and word cloud visualizations, as well as co-word and co-cited network analyses. Particularly noteworthy is the temporal segmentation into two distinct research phases: the Exploration Stage (2015-2020) and the Development Stage (2021-2025), through which the study systematically examines the shifting trajectories of research hotspots within the target domain. This phased analytical approach enables a nuanced understanding of the field's evolving research priorities and intellectual structure over time.

3.1 Timeline of Publications

We began by analyzing the number of publications per year. The following **Figure 1** illustrates the publication trend over the past decade.

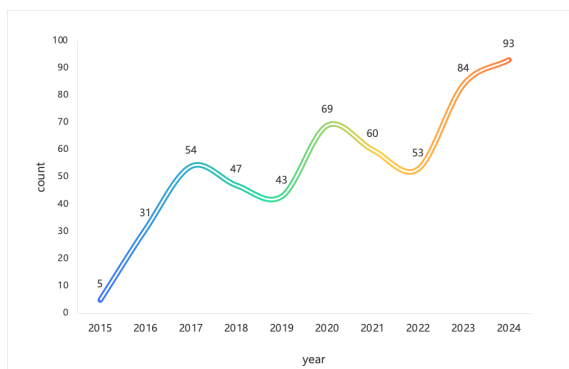


Figure 1: Publication Trend

The year 2025 (with a total of 8 publications) is excluded, as it is not a complete year and does not accurately reflect the total output. Overall, the research shows a steadily growing trend in interest and activity, increasing from just 5 publications in 2015 to 93 in 2024. The growth pattern also reflects some fluctuations. A slight decline in publication numbers in 2018 and 2019 may be linked to rising concerns over data privacy following the growing popularity of smart home systems, which could have

temporarily slowed research momentum. Another dip in 2021 and 2022 may be attributed to the lingering impacts of the COVID-19 pandemic, which disrupted research activities and priorities across many fields.

3.2 Authors Productivity and Collaboration

We analyzed author statistics, focusing on the number of publications in the field and patterns of collaboration (Figure 2). The most productive author is Gram-Hanssen, K, with eight publications, followed by Sovacool, B.K. with six publications, and Wang, B. and Skjolsvold, T.M. with five publications each. The co-authorship network diagram illustrates collaborative relationships among authors; however, the results indicate a scattered and loosely connected structure. This suggests that the field is currently at a stage where researchers are initiating their work independently. In the future, we expect to see increased collaboration, including more interdisciplinary partnerships and cross-institutional research efforts.

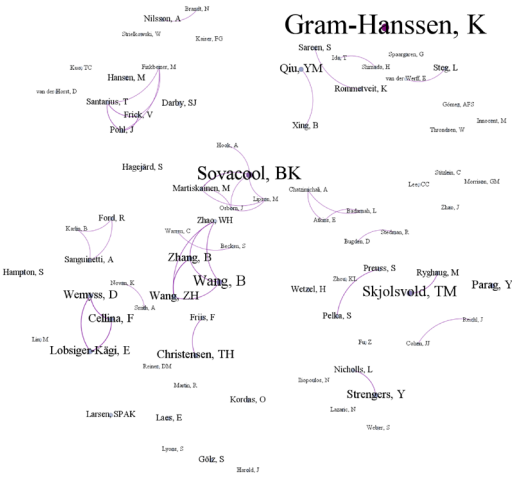


Figure 2: Co-authorship network

3.3 Journal Statistics

We conducted a descriptive analysis to identify the journals in which research on this topic has been published. The popularity of the top 10 journals is presented in the following Table 1 and Figure 3. Energy Research & Social Science (88

publications), Sustainability (68 publications), and Energy Policy (62 publications) have the highest number of publications, and all three are Q1 journals. According to the latest available data as of 2025, the impact factors for the journals Energy Research & Social Science and Energy Policy are both notably high, at 6.9 and 9.3, respectively.

Table 1: Table of Journal Statistics

Source Title	Account	Percentage
Energy Research & Social Science	88	16.1%
Sustainability	68	12.4%
Energy Policy	62	11.3%
Journal Of Cleaner Production	37	6.8%
Energy Efficiency	31	5.7%
Energy Economics	15	2.7%
Environmental Research Letters	8	1.5%
Resources Conservation and Recycling	7	1.3%
Transportation Research Part D-Transport and Environment	7	1.3%
Environmental Innovation and Societal Transitions	6	1.1%

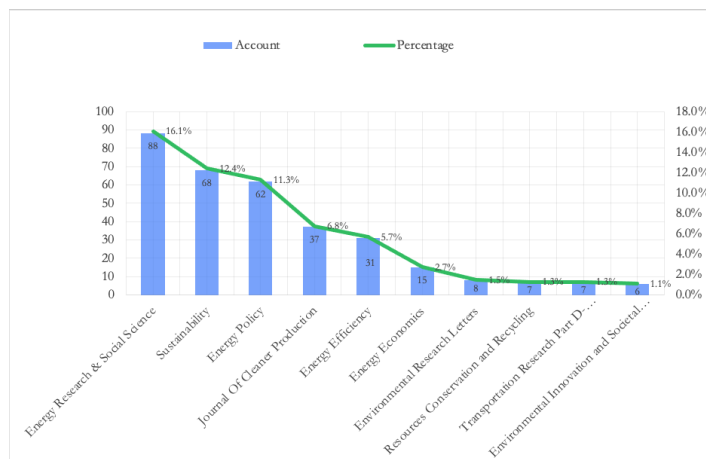


Figure 3: Journal Popularity

3.4 Word Frequency and Word Cloud

We analyzed word frequency in both titles and abstracts to identify the most researched terms. We generated word clouds for both, which appear similar. **Figure 4** is the word cloud for titles, which visually represents the most frequently occurring

terms. As expected, “energy” is the most dominant term, reflecting the central focus of the research in this domain. Additionally, “smart” appears frequently, often in conjunction with “meter”, “system”, and “consumption”. These keywords indicate that much of the research emphasizes smart energy systems, metering technologies, and consumption patterns. This distribution of terms confirms that our search strategy successfully captured relevant studies within the topic, as the retrieved articles focus on expected key areas. Furthermore, the absence of surprising or unrelated keywords suggests that the dataset is well-refined and accurately represents the research landscape in this field.

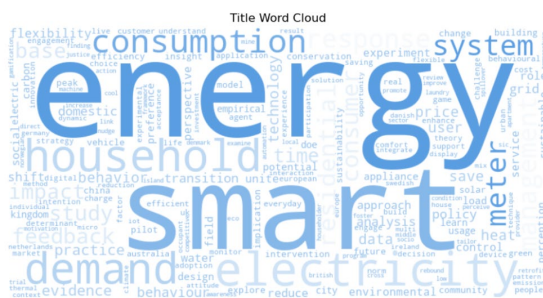


Figure 4: Title word frequency cloud

3.4 Co-Word and Co-Cited Analysis

From the 574 articles published between 2015 and 2025, we divided the research development into two stages: the exploration stage (2015–2020) and the development stage (2021–2025). The following figures and analyses present the co-word and co-citation patterns specific to each stage, highlighting key thematic clusters centered around the topic of energy. These clusters reflect interconnected research areas and reveal the dominant themes that have shaped the literature in each phase. In general, several distinct clusters emerge in the last decade clearly around Energy (illustrated from combining Figure 5 and 6):

Energy Sources: This cluster includes terms such as electricity, gas, water, and fuel, representing the types of energy sources discussed in the literature.

Home Technology: This cluster includes terms such as smart technology, meter, device, system, apps, and model etc., which are closely related to advancements in home energy management and digital innovations in energy efficiency.

Human-Centric: Another prominent group of terms focuses on people and their roles in energy consumption, including household, consumer, producer, and user. This indicates a strong research interest in how individuals and communities interact with energy systems.

Economic Factors: A separate cluster emphasizes monetary aspects, featuring terms like price, incentive, cost, bill, and tariff. This reflects the role of economic mechanisms in shaping energy consumption behaviors, such as pricing strategies and financial incentives.

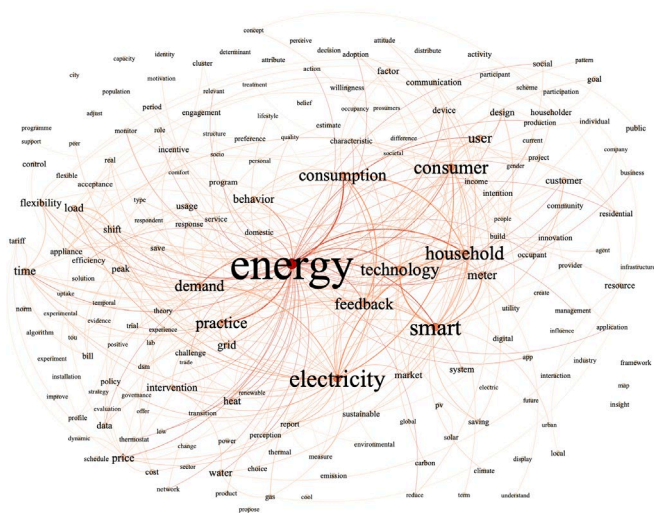
Environmental Concerns: The importance of sustainability and climate impact is evident in terms like carbon and green, highlighting the connection between energy use and environmental policies.

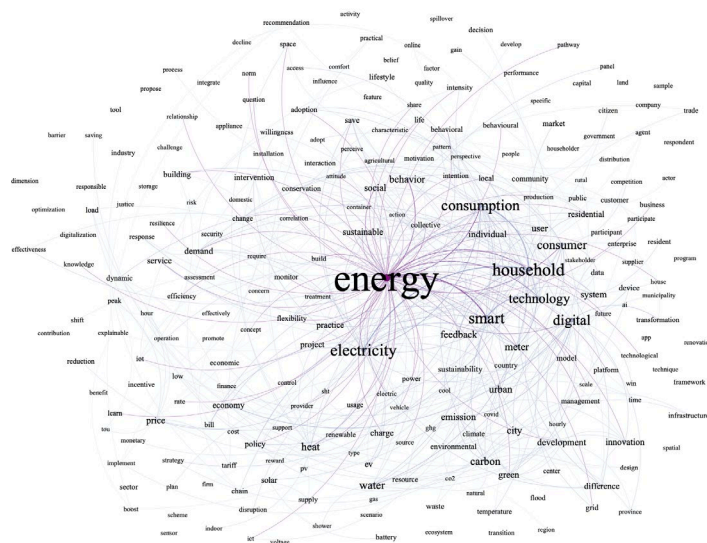
Behavioral Aspects: Lastly, the behavioral dimension is captured through terms such as behavior (behaviour), usage, effectiveness, conservation, reduce, save, and practice. This cluster emphasizes research on energy-saving behaviors, decision-making processes, and the effectiveness of various interventions.

When comparing the two stages, several differences emerge, which are discussed in the following paragraphs.

Exploration stage (2015-2020): the thematic clusters from the first stage are illustrated in **Figure 5**. The clusters and keywords that distinguish this stage from stage two are:

1. Practice: practice, experience, lab, trial, theory, challenge.
2. Demand: peak, load, flexible, solution, control, efficiency.
3. Behavior: lifestyle, belief, characteristic, willingness.





emphasis on user-centric solutions, and behavior change approaches in the energy sector.

There are several avenues for future research. Based on thematic analysis, we examined multiple studies in detail and identified key directions for further exploration. Future studies could investigate the role of Artificial Intelligence (AI), blockchain, and Internet of Things (IoT) not only in optimizing energy efficiency and demand-side management but also in understanding how these technologies interact with human behavior to enhance practical applications (Froehlich, 2009; Nilsson et al., 2018; Stieglitz et al., 2023). Additionally, a deeper understanding of users is lacking. It is important to understand who the users of smart home technologies are and how they interact with and engage with these systems. Gram-Hanssen & Darby (2018) argued that the home represents a place of safety and control, raising questions about the boundaries of control and cooperation in smart home systems. Moreover, improving energy efficiency could lead to a rebound effect, where consumers use more energy than necessary (Kroll et al., 2019). Future research should examine this phenomenon through the lens of digital nudging in smart home applications to better motivate and foster deeply ingrained, intrinsic sustainable behavior that persists. However, it is also crucial to consider the trade-offs between comfort, value, and cost, as these vary among users. People could prioritize warmth and comfort at home over others (Sovacool et al., 2020). Moreover, the impact of social influence and peer pressure on encouraging PEB warrants further investigation, as these factors might be rather effective than focusing solely on individual motivations (Froehlich, 2009; Gadenne et al., 2011). Furthermore, research should also aim at providing policy recommendations, such as strategies to support vulnerable groups (Sovacool et al., 2017) and approaches to promote the adoption of smart home systems (Froehlich, 2009; M. Khan et al., 2016; Robison et al., 2023).

By addressing these research gaps, future studies can further enhance the understanding of energy users' behavior in relation with energy systems, improve policy effectiveness, and drive innovation in sustainable energy solutions.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 52161135202) and the Merian Fund by the Dutch Research Council (482.20.608).

References

- Costanza, E., Fischer, J. E., Colley, J. A., Rodden, T., Ramchurn, S. D., & Jennings, N. R. (2014). Doing the laundry with agents: A field trial of a future smart energy system in the home. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 813–822. <https://doi.org/10.1145/2556288.2557167>
- Dekeyrel, S., & Fessler, M. (2024). Digitalisation: An Enabler for the Clean Energy Transition. *Journal of Energy and Natural Resources Law*, 42, 185.
- Dennstädt, F., Zink, J., Putora, P. M., Hastings, J., & Cihoric, N. (2024). Title and abstract screening for literature reviews using large language models: An exploratory study in the biomedical domain. *Systematic Reviews*, 13(1), 158. <https://doi.org/10.1186/s13643-024-02575-4>
- Froehlich, J. (2009). Promoting Energy Efficient Behaviors in the Home through Feedback: The Role of Human-Computer Interaction. *HCIC Winter Workshop Boaster Paper*.
- Gadenne, D., Sharma, B., Kerr, D., & Smith, T. (2011). The influence of consumers' environmental beliefs and attitudes on energy saving behaviours. *Energy Policy*, 39(12), 7684–7694. <https://doi.org/10.1016/j.enpol.2011.09.002>
- Gram-Hanssen, K., & Darby, S. J. (2018). “Home is where the smart is”? Evaluating smart home research and approaches against the concept of home. *Energy Research & Social Science*, 37, 94–101. <https://doi.org/10.1016/j.erss.2017.09.037>
- Khan, K., & Su, C. wei. (2023). Does technology innovation complement the renewable energy transition? *Environmental Science and Pollution Research*, 30(11), 30144–30154. <https://doi.org/10.1007/s11356-022-24336-3>
- Khan, M., Silva, B. N., & Han, K. (2016). Internet of Things Based Energy Aware Smart Home Control System. *IEEE Access*, 4, 7556–7566. IEEE Access. <https://doi.org/10.1109/ACCESS.2016.2621752>
- Kroll, T., Paukstadt, U., Kreidermann, K., & Mirbabaie, M. (2019). Nudging People to Save Energy in Smart Homes with Social Norms and Self-Commitment. *Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019*. <https://fis.uni-bamberg.de/handle/uniba/95858>
- Nilsson, A., Wester, M., Lazarevic, D., & Brandt, N. (2018). Smart homes, home energy management systems and real-time feedback: Lessons for influencing household energy consumption from a Swedish field study. *Energy and Buildings*, 179, 15–25. <https://doi.org/10.1016/j.enbuild.2018.08.026>
- Risteska Stojkoska, B. L., & Trivodaliev, K. V. (2017). A review of Internet of Things for smart home: Challenges and solutions. *Journal of Cleaner Production*, 140, 1454–1464. <https://doi.org/10.1016/j.jclepro.2016.10.006>
- Robison, R., Skjølsvold, T. M., Hargreaves, T., Renström, S., Wolsink, M., Judson, E., Pechancová, V., Demirbağ-Kaplan, M., March, H., Lehne, J., Foulds, C., Bharucha, Z., Bilous, L., Büscher, C., Carrus, G., Darby, S., Douzou, S., Drevenšek, M., Frantál, B., ... Wyckmans, A. (2023). Shifts in the smart research agenda? 100 priority questions to accelerate sustainable energy futures. *Journal of Cleaner Production*, 419, 137946. <https://doi.org/10.1016/j.jclepro.2023.137946>
- Sovacool, B. K., Kivimaa, P., Hielscher, S., & Jenkins, K. (2017). Vulnerability and resistance in the United Kingdom's smart meter transition. *Energy Policy*, 109, 767–781. <https://doi.org/10.1016/j.enpol.2017.07.037>
- Sovacool, B. K., Osborn, J., Martiskainen, M., & Lipson, M. (2020). Testing smarter control and feedback with users: Time, temperature and space in household heating preferences and

- practices in a Living Laboratory. *Global Environmental Change Part A: Human & Policy Dimensions*, 65, N.PAG-N.PAG. 8gh. <https://doi.org/10.1016/j.gloenvcha.2020.102185>
- Stieglitz, S., Mirbabaie, M., Deubel, A., Braun, L.-M., & Kissmer, T. (2023). The potential of digital nudging to bridge the gap between environmental attitude and behavior in the usage of smart home applications. *International Journal of Information Management*, 72, 102665. <https://doi.org/10.1016/j.ijinfomgt.2023.102665>
- Whitmarsh, L. (2011). Social and Psychological Drivers of Energy Consumption Behaviour and Energy Transitions. In *Political Economy of the Environment*. Routledge.
- Wilson, C., Hargreaves, T., & Hauxwell-Baldwin, R. (2015). Smart homes and their users: A systematic analysis and key challenges. *Personal and Ubiquitous Computing*, 19(2), 463–476. <https://doi.org/10.1007/s00779-014-0813-0>

BREAKING THE ICE? EXAMINING THE COGNITIVE, SOCIAL, AND PSYCHOLOGICAL DETERMINANTS OF USER RESISTANCE TO ELECTRIC NON-ROAD WORK VEHICLE ADOPTION

HENRI JALO, MARKO SEPPÄNEN

Tampere University, Faculty of Management and Business, Tampere, Finland
henri.jalo@tuni.fi, marko.seppanen@tuni.fi

The success of organizationally mandated technology adoption often hinges on whether employees are receptive or resistant to the new technology. This study examines non-road work vehicle operators' attitudes toward switching to using electric work vehicles. A research model investigating the impact of cognitive evaluations of vehicle attributes, organizational social context, and psychological antecedents on resistance was tested using survey data from 1460 respondents collected via an online panel. Results from hierarchical multiple regression analysis show that beneficial vehicle attributes (perceived sustainability and quietness of electric work vehicles) and positive social context (colleague opinions and organizational adoption intentions) lower resistance. In contrast, psychological attitudes related to technostress and disrupting the status quo (techno-overload and inertia) significantly increase resistance. Theoretical contributions and practical implications of these findings are discussed.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.33](https://doi.org/10.18690/um.fov.4.2025.33)

ISBN
978-961-286-998-4

Keywords:
adoption,
resistance,
end user,
organization,
sustainability,
electric vehicle,
regression



University of Maribor Press

1 Introduction

Employees often face situations where they are mandated to adopt technologies or information systems by their organization (Heath et al., 2022; Ilie & Turel, 2020; Klaus et al., 2010). Although new technologies often bring benefits to employees, resistance stemming from various sources, such as inertia (Polites & Karahanna, 2012), is a common reaction because people typically prefer to maintain the status quo (Kim & Kankanhalli, 2009). However, the dynamics and relative strength of these drivers and inhibitors of resistance are still not well understood. This is a critical concern, as user resistance may lead to underutilization of the new technology (Ilie & Turel, 2020) or even employee turnover (Califf et al., 2020; Lapointe & Rivard, 2005), resulting in the loss of institutional knowledge and necessitating additional training investments for new hires. With the pace of technological development and adoption increasing, a better understanding of how these factors influence user resistance is needed.

Electric vehicles (EVs) are being increasingly adopted by consumers, and organizations are now turning their attention to electrifying work fleets—especially non-road work vehicles—due to potential operational and employee benefits (Dehkordi et al., 2024; McKinsey & Company, 2023). These benefits include improved air quality in indoor and urban work environments and reduced greenhouse emissions (Lajunen et al., 2016, 2018). For instance, in 2019 non-road work vehicles accounted for an estimated 2% of total greenhouse emissions and up to 11% in certain EU industry sectors (Lončarević et al., 2022). However, adopting such vehicles can be challenging for companies because employees often vary in their receptiveness. The added complexity of operating the more digitalized control systems of the vehicles (Strayer et al., 2019) may, for instance, require employees to learn new operating procedures and disrupt established routines. Identifying the factors that contribute to resistance is crucial for organizations to proactively mitigate its emergence. Otherwise, resistance behaviors may spread and become entrenched (Jalo & Pirkkalainen, 2024; Lapointe & Rivard, 2005), making them particularly difficult to overcome (Selander & Henfridsson, 2012).

This paper examines how beneficial vehicle characteristics, an organization's social context, and employees' psychological attitudes toward change impact their resistance to adopting battery electric non-road vehicles (BENVs). The research

model is tested using hierarchical multiple-regression analysis based on a sample of 1460 current non-road work vehicle operators collected via the Prolific online survey panel. The results show that although the perceived benefits of BENVs, namely their higher perceived environmental sustainability and lower noise, and an adoption-supportive social environment can dampen resistance attitudes, users' psychological predispositions toward switching to BENVs, specifically inertia and fear of techno-overload, are on balance more strongly related to resistance. We contribute to the innovation adoption and resistance literatures by empirically validating the relevance of the proposed resistance antecedents through a large survey study.

The rest of the paper is structured as follows. In Section 2, the theoretical background and the research model are presented. Section 3 describes the survey development and data collection. Section 4 presents the hierarchical multiple-regression analysis results. The results are discussed in Section 5, along with the study's limitations and suggestions for future research.

2 Theoretical background

EV adoption has received substantial attention in the innovation adoption literature (Kumar & Alok, 2020); however, the enterprise adoption context remains less studied. Switching entire vehicle fleets from internal combustion engine (ICE) vehicles to electric motor vehicles represents a significant investment for companies. Beyond the business case and larger ecosystem concerns (Dehkordi et al., 2024), it is also essential for companies' employees to be receptive toward the use BENVs, as resistance is one of the leading causes for failed technology adoptions (Cieslak & Valor, 2025). We ground our examination of employee resistance attitudes in the perceived beneficial *characteristics* of BENVs, the organization's *social context*, and employees' *psychological attitudes* related to disrupting the status quo (Kim & Kankanhalli, 2009; Tarafdar et al., 2007).

When employees evaluate a technology, they often consider its technological or functional benefits, including a cognitive assessment of both switching benefits and switching costs (Kim & Kankanhalli, 2009). In the context of EVs, their perceived superior environmental impact compared to ICE vehicles has been identified as a key determinant of adoption in the consumer context (Kumar & Alok, 2020). Although employees might not conduct formal financial cost-benefit analyses of

company investments, they may still evaluate the sustainability of BENVs similarly, potentially reducing resistance due to a personal preference for more environmentally friendly vehicles. EV users have also been shown to appreciate the lower noise levels produced by electric vehicles (Schmalfuß et al., 2017). The relevance of reduced noise may be particularly salient in work settings, where vehicles are typically operated for several hours a day. We therefore operationalize the *sustainability of BENVs* and employees' *low-noise preference* as key vehicle characteristics in our research model.

In the context of mandated technology adoption, the user's social context plays a central role (Khechine et al., 2023), as it shapes user attitudes through social pressure and the need to align with organizational objectives. When users are initially resistant to using the new technology, pressure from their immediate social environment (i.e., co-workers) can reduce resistance (Kim & Kankanhalli, 2009). However, social pressure can also increase resistance if employees discourage others from using the new solution (Lapointe & Rivard, 2005). However, the impact of employees' perceptions of their organization's adoption intentions on resistance remains less well understood, although an inverse relationship has been established from the organizational perspective—namely, that expected employee resistance can lower organizational adoption intentions (Jalo & Pirkkalainen, 2024). Accordingly, we operationalize the social context to include *colleague opinion* and *perceived organizational adoption intention*.

Employees typically favor the status quo (Kim & Kankanhalli, 2009). New technologies introduce uncertainty and require changes to established routines, which can engender resistance (Polites & Karahanna, 2012). Moreover, complex technologies often cause technostress, as employees may fear increased workloads—especially during the learning phase (Tarafdar et al., 2007). In addition to their power generation, BENVs differ from ICE vehicles through their heavier reliance on digital control systems, display interfaces, and integration with various information systems (Strayer et al., 2019). This adds a new layer of complexity for end users, who must adapt to digital interfaces that may differ substantially from the more analog controls of ICE work vehicles. Such shifts can trigger perceptions of added workload and feelings of techno-overload (Tarafdar et al., 2007), driven by anticipated complexity of BENV operation. These psychological responses have been linked to negative employee outcomes, such as lowered employee satisfaction, as well as increased

attrition and turnover intentions (Califf et al., 2020). As a psychological coping mechanism, employees may prefer to continue using their current vehicles, which are more familiar and aligned with their existing expertise. Routine-seeking and cognitive rigidity also contribute to inertia, both of which have been linked to resistance attitudes (Laumer et al., 2016). Organizations must therefore account for employee inertia when implementing new technologies (Polites & Karahanna, 2012). Accordingly, the psychological dimension of our research model is operationalized through *techno-overload* and *inertia*.

As control variables, we included age, gender, and prior experience with EVs in non-work contexts. These variables are commonly included in technology adoption models (Venkatesh et al., 2003). Controlling for prior experience is particularly important, as users with limited exposure may hold inaccurate perceptions about EV performance (Burgess et al., 2013).

3 Methodology

The survey data were collected between November and December 2024 using the Prolific online survey panel and the Alchemer survey tool. The target group comprised operational employees and lower- and middle-level managers expected to be closely involved in vehicle operation. A total of 1573 responses were gathered. After removing 111 responses due to incorrect answers to one of the two attention trap questions, 1462 responses remained. Further examination identified six responses with a standard deviation of less than 0.7. Two of these were completed in under five minutes, suggesting potential inattentiveness, given that the estimated completion time for the survey was 15 minutes. These two responses were removed, resulting in a final sample size of 1460. Demographic and background information of the respondents is reported in Table 1. The respondents' companies were located in the United States (554), United Kingdom (458), Canada (94), Poland (74), Germany (50), Spain (25), Australia (25), Netherlands, (25), and 155 in other countries. The primary industries of the respondents were Manufacturing or Industrial (331), Warehousing and Distribution (275), Construction and Architecture (234), Consumer Goods (133), Agriculture (76), Aerospace or Aviation (65), and 346 worked in other industries.

Table 1: Sample demographics and background information

	Frequency	%
Gender		
Male	1073	73.5
Female	387	26.5
Age (years)		
18 to 29	530	36.3
30 to 39	504	34.5
40 to 49	237	16.2
50 to 59	148	10.1
60 to 76	41	2.8
Organizational position		
Operational level employee (e.g., forklift driver)	591	40.5
Lower-level manager (e.g., team supervisor)	384	26.3
Middle manager (e.g., worksite manager)	485	33.2
Non-work EV experience		
Not familiar with electric vehicles outside of work	214	14.7
Familiar with electric vehicles, but haven't driven one	382	26.2
Ridden as a passenger, but haven't driven one	273	18.7
Have driven an electric vehicle before	485	33.2
I own an electric vehicle	106	7.3

The constructs, item wordings, means, standard deviations, Cronbach's alpha values, and standardized item loadings are reported in Table 2. All scales were measured using 7-point Likert items ranging from Strongly disagree to Strongly agree. The BENV sustainability scale was adapted from Fang and Li's (2022) sustainability scale, which was originally based on Möhlmann's (2015) environmental impact scale. The low-noise enjoyability scale (Schmalfuß et al., 2017) was adapted to focus on preference for low vehicle noise rather than driving style adaptations. The colleague opinion scale (Kim & Kankanhalli, 2009) was slightly reworded to focus specifically on BENVs. The behavioral intention to use scale (Venkatesh et al., 2003) was adapted to assess employees' perceptions of organizational-level adoption intention, reflecting the mandatory adoption context. The inertia scale (Polites & Karahanna, 2012) was adapted to examine employee preferences for their current vehicles rather than their use continuance intention. The techno-overload scale (Tarafdar et al., 2007) was modified to assess perceptions of whether the complexity of BENV adoption would increase workload. Finally, the dependent variable of resistance was adapted from Kim and Kankanhalli (2009) to capture attitudinal rather than active resistance.

Table 2: Survey constructs and items (*) $p < 0.001$**

Construct	Item wording	Loading
Sustainability of BENVs (Fang & Li, 2022; Möhlmann, 2015) $\alpha = 0.872$ Mean = 5.22 SD = 1.19	SUST1: Using electric work vehicles helps conserve more resources than traditional vehicles	0.823***
	SUST2: Using electric work vehicles is more sustainable than traditional vehicles	0.788***
	SUST3: Using electric work vehicles is more efficient in terms of resource utilization than traditional vehicles	0.781***
	SUST4: Using electric work vehicles is more environmentally friendly than using traditional vehicles	0.786***
Low-noise preference (Schmalfuß et al., 2017) $\alpha = 0.842$ Mean = 5.47 SD = 1.17	LONO1: The low noise level of electric work vehicles would make driving more enjoyable	0.775***
	LONO2: I would like the quietness of electric work vehicles	0.815***
	LONO3: I would perceive the low noise level of electric work vehicles as pleasant	0.815***
Colleague opinion (Kim & Kankanhalli, 2009) $\alpha = 0.897$ Mean = 4.59 SD = 1.34	COOP1: Most of my colleagues think switching to electric work vehicles is a good idea	0.879***
	COOP2: My peers support switching to electric work vehicles	0.854***
	COOP3: Most people I work with encourage switching to electric work vehicles	0.857***
Organizational future use intention (Venkatesh et al., 2003) $\alpha = 0.913$ Mean = 5.20 SD = 1.35	OFUS1: I predict that our organization will use electric work vehicles in the future	0.826***
	OFUS2: Our organization plans to use electric work vehicles in the future	0.909***
	OFUS3: Our organization intends to use electric work vehicles in the future	0.916***
Inertia (Polites & Karahanna, 2012; Shi et al., 2018) $\alpha = 0.907$ Mean = 3.62 SD = 1.47	INER1: I prefer traditional work vehicles because they are stress-free	0.802***
	INER2: I prefer traditional work vehicles because using them is more comfortable for me	0.862***
	INER3: I prefer traditional work vehicles because I have been working with them for so long	0.859***
	INER4: I prefer traditional work vehicles because I have used them regularly in the past	0.850***
Technology-overload (Tarafdar et al., 2007) $\alpha = 0.843$ Mean = 2.96 SD = 1.38	OVLO1: The complexity of using electric vehicles would force me to do more work than I can handle	0.787***
	OVLO1: The complexity of using electric vehicles would leave me less time to focus on my actual work	0.821***
	OVLO1: I would have a higher workload because of the complexity of electric vehicles	0.793***
Resistance (Kim & Kankanhalli, 2009) $\alpha = 0.861$ Mean = 2.68 SD = 1.44	RESI1: I am reluctant to switch to using electric work vehicles	0.800***
	RESI2: I am unwilling to switch to using electric work vehicles	0.839***
	RESI3: I oppose switching to using electric work vehicles	0.828***

4 Results

4.1 Reliability and discriminant and convergent validity

IBM SPSS version 29 was used to conduct an exploratory factor analysis (EFA) to examine the factor structure by assessing item cross-loadings. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.935, exceeding the recommended threshold of 0.8, and Bartlett's test of sphericity was significant ($\chi^2(276) = 23390.377$; $p < 0.001$), indicating sufficient inter-item correlations for factor analysis (Hair et al., 2014). EFA was performed using principal axis factoring with promax rotation (Costello & Osborne, 2019; Matsunaga, 2010). The seven-factor solution explained 68.5% of the total variance. All items loaded onto their expected constructs with adequate to strong loadings ranging from 0.627 to 0.950, and no strong cross-loadings (> 0.32) were identified (Costello & Osborne, 2019).

Next, we conducted confirmatory factor analysis (CFA) using IBM Amos version 28. We began by examining the standardized factor loadings. One resistance item had a loading of 0.64, below the commonly accepted threshold of 0.707 (Hair et al., 2014), indicating it explained less than 50% of the variance in the latent construct. This item was therefore dropped from subsequent analyses, resulting in three items measuring resistance.

We then assessed convergent and discriminant validity using average variance extracted (AVE), maximum shared variance (MSV), and composite reliability (CR) (Fornell & Larcker, 1981), as well as the heterotrait-monotrait (HTMT) ratio of correlations (Henseler et al., 2015; Voorhees et al., 2016). The Fornell-Larcker and HTMT results, obtained using the Master Validity tool by Gaskin et al. (2019), are shown in Table 3. Following the guidelines of Hair et al. (2014) and Henseler et al. (2015) ($CR > 0.7$; $AVE > 0.5$; $MSV < AVE$; square root of AVE greater than inter-construct correlations; $HTMT < 0.85$), the measurement model demonstrated adequate convergent and discriminant validity.

Table 3: Fornell-Larcker (lower triangle) and HTMT (upper triangle) results

	CR	AVE	MSV	INER	SUST	OFUS	OVLO	COOP	RESI	LONO
INER	0.908	0.712	0.535	0.844	0.435	0.373	0.583	0.360	0.653	0.380
SUST	0.873	0.632	0.371	-0.485	0.795	0.441	0.231	0.540	0.518	0.523
OFUS	0.915	0.782	0.458	-0.403	0.478	0.885	0.193	0.618	0.444	0.342
OVLO	0.843	0.641	0.472	0.660	-0.268	-0.207	0.801	0.058	0.585	0.267
COOP	0.898	0.745	0.458	-0.404	0.608	0.677	-0.072	0.863	0.406	0.353
RESI	0.863	0.677	0.535	0.732	-0.595	-0.481	0.687	-0.463	0.823	0.443
LONO	0.844	0.643	0.371	-0.430	0.609	0.379	-0.321	0.407	-0.524	0.802

4.2 Common method variance

We assessed common method variance (CMV) using Harman's single factor test with principal axis factoring and no rotation (Fuller et al., 2016). The single factor explained only 36.997% of the variance, well below the suggested 50% threshold. We also examined the correlations between a marker variable (MV, measuring attitude toward the color blue) and the theoretical variables (Miller & Simmering, 2023; Williams et al., 2010) during CFA. The absolute mean correlation between the MV and the theoretical constructs was 0.09, indicating minimal CMV (Malhotra et al., 2006). Having established the suitability of the measurement model, we calculated mean scores for each construct in SPSS for use in hierarchical regression analysis. Lastly, we checked variance inflation factors (VIF) during the regression analysis, with the highest VIF value being 2.008, well below the more stringent cutoff of 3 (Hair et al., 2014). Thus, CMV is not a threat to the validity of the results.

4.3 Hierarchical multiple-regression analysis results

Table 4 presents the four hierarchical multiple-regression models predicting resistance. Model 1 (control variables) explains 1.4% of the variance ($R^2 = 0.014$). Age ($\beta = -0.108^{***}$) and prior EV experience ($\beta = -0.066^*$) are both negatively associated with resistance.

Model 2 adds vehicle characteristics, increasing R^2 to 0.313. Perceived BENV sustainability ($\beta = -0.395^{***}$) and low-noise preference ($\beta = -0.230^{***}$) both reduce resistance; prior EV experience is no longer significant.

Model 3 adds social-context variables, slightly increasing R^2 to 0.355. Organizational adoption intention ($\beta = -0.214^{***}$) is negatively associated with resistance, whereas colleague opinion is not. Vehicle characteristic effects remain statistically significant ($p < 0.001$), although their coefficients are slightly smaller.

Model 4 introduces psychological predictors, boosting R^2 to 0.589. Inertia ($\beta = 0.274^{***}$) and techno-overload ($\beta = 0.336^{***}$) show the strongest positive relations with resistance. All vehicle- and social-context predictors remain negative and significant, albeit with substantially lower coefficients. Age is no longer significant, while colleague opinion attains significance ($\beta = -0.091^{***}$).

Table 4: Hierarchical regression analysis results ($p < 0.001$, ** $p < 0.01$, * $p < 0.05$)**

	Model 1	Model 2	Model 3	Model 4
	Beta (sig.)	Beta (sig.)	Beta (sig.)	Beta (sig.)
Demographic and background characteristics				
Age	-0.108***	-0.78***	-0.066**	-0.025
Gender	-0.011	0.031	0.020	-0.021
Non-work EV experience	-0.066*	-0.018	0.011	-0.017
Vehicle characteristics				
Sustainability of BENVs		-0.395***	-0.299***	-0.173***
Low-noise preference		-0.230***	-0.194***	-0.107***
Organizational social context				
Colleague opinion			-0.042	-0.091***
Organizational adoption intention			-0.214***	-0.109***
Psychological attitudes				
Inertia				0.274***
Techno-overload				0.336***
Adjusted R^2	0.014	0.313	0.355	0.589
ΔR^2	-	0.299	0.042	0.234
F change	F(3, 1456)= 7.945***	F(2, 1454)= 317.542***	F(2, 1452)= 48.374***	F(2, 1450)= 415.604***

5 Discussion

Companies need to consider employee attitudes when mandating the use of new technologies to ensure smooth roll-outs (Heath et al., 2022; Klaus et al., 2010). Our study examined how employees' perceptions of BENVs shape resistance to replacing ICE vehicles. The hierarchical regression results reveal a clear pattern. In the first step, age and prior non-work EV experience showed small negative links to

resistance. However, once perceived BENV sustainability and low-noise preference were added, prior experience lost its influence. This shift indicates that employees are influenced less by mere familiarity with EVs than by conviction that BENVs deliver concrete environmental and noise-related advantages.

Adding social-context variables produced only a modest gain in explained variance, yet perceptions of the organization's intention to adopt BENVs still lowered resistance. Colleague opinion was insignificant until the final step, when techno-overload and inertia were included. Evidently, employees who feel overwhelmed by the technology or are strongly attached to their current vehicles discount peer enthusiasm for BENVs. Once those psychological barriers are accounted for, positive peer sentiment emerges as a distinct force that dampens resistance.

Psychological attitudes clearly dominate the story. Inertia and techno-overload not only have the largest coefficients but also suppress the age effect observed in earlier models. Although older employees appeared less resistant at first, their stance is better explained by lower perceived overload and weaker inertia rather than by age itself. Overall, potential losses in terms of extra workload and disruption of routines loom larger than potential gains, a finding consistent with the status-quo-bias perspective (Kim & Kankanhalli, 2009). Yet when employees clearly perceive BENVs' sustainability and noise advantages, those benefits can overcome default reluctance, provided that overload and inertia are addressed.

5.1 Theoretical contributions

This study contributes to theory in several ways. First, by demonstrating that techno-overload is a direct antecedent of resistance, it bridges the technostress and user-resistance literatures and extends prior work that linked overload mainly to employee satisfaction and turnover intentions (Califf et al., 2020; Tarafdar et al., 2007). Second, the findings refine status-quo-bias theory by indicating that demographic attributes such as age influence resistance mainly through their association with psychological factors—specifically inertia and techno-overload—rather than exerting an independent effect. Third, the study reveals a multi-level social-influence dynamic, indicating that an organization's declared adoption intention exerts a more consistent downward pressure on resistance than informal colleague opinion. This is noteworthy because most employee-level resistance models include only peer

opinion as a social antecedent (e.g., Kim & Kankanhalli, 2009), whereas organizational-level adoption research has shown a negative link between anticipated employee resistance and organizational adoption intention (Jalo & Pirkkalainen, 2024). Collectively, these insights deepen our understanding of why employees resist or embrace mandatory sustainable technologies.

5.2 Practical implications

Our results have several implications for companies. First, companies should prioritize alleviating employees' fears about the extra work they may associate with BENV adoption. Before full deployment, small-scale pilot projects could be conducted to provide hands-on experience and allow employees to explore the vehicles' interfaces in a low-stakes environment. Such experiential learning can build competence and may help reduce techno-overload. To ease the switch from the status quo, professional inertia related to employees' preference for ICE vehicles must be addressed. Highlighting the benefits of BENVs can help mitigate these issues, but companies should also leverage peer influence to foster more positive attitudes toward adoption. Finally, communicating the organization's adoption plans well in advance may help prevent the emergence of resistance.

5.3 Limitations and future research

Our study has a few limitations. First, the data were collected using a cross-sectional design, which restricts our ability to infer causality between the antecedents and resistance (Maier et al., 2023). Longitudinal research that tracks pre- and post-adoption attitudes could clarify how resistance evolves over time in mandatory adoption settings. Second, the hierarchical regression model examined only direct relationships between the predictor and dependent variables. Future research might utilize structural equation modeling to examine whether some of the predictors mediate the influence of the antecedents or employ fuzzy-set qualitative comparative analysis to explore causal asymmetry and identify distinct resistance profiles (Pappas & Woodside, 2021). Third, our sample is mainly drawn from industrial, construction, and logistics contexts in which employees operate non-road work vehicles. Although the underlying mechanisms—status-quo bias, techno-overload, and multi-level social influence—should be broadly generalizable, sector-specific factors such as regulatory intensity or task complexity may moderate their effects. Replicating the

model in other sectors, such as healthcare or public services, would help clarify boundary conditions and enhance external validity.

Acknowledgements

This research was supported by the Business Finland project Decarbonization of Mobile Machine Systems (Decarbo) (7942/31/2022).

References

- Burgess, M., King, N., Harris, M., & Lewis, E. (2013). Electric vehicle drivers' reported interactions with the public: Driving stereotype change? *Transportation Research Part F: Traffic Psychology and Behaviour*, 17, 33–44. <https://doi.org/10.1016/j.trf.2012.09.003>
- Califf, C. B., Sarker, S., & Sarker, S. (2020). The Bright and Dark Sides of Technostress: A Mixed-Methods Study Involving Healthcare IT. *MIS Quarterly*, 44(2), 809–856. <https://doi.org/10.25300/MISQ/2020/14818>
- Cieslak, V., & Valor, C. (2025). Moving beyond conventional resistance and resistors: An integrative review of employee resistance to digital transformation. *Cogent Business & Management*, 12(1), 2442550. <https://doi.org/10.1080/23311975.2024.2442550>
- Costello, A. B., & Osborne, J. W. (2019). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research, and Evaluation*, 10(1), 7. <https://doi.org/10.7275/jyj1-4868>
- Dehkordi, R., Ahokangas, P., Evers, N., & Sorvisto, M. (2024). Business model design for Electric Commercial Vehicles (ECVs): An ecosystemic perspective. *Energy Policy*, 186, 113971. <https://doi.org/10.1016/j.enpol.2023.113971>
- Fang, Y.-H., & Li, C.-Y. (2022). Does the sharing economy change conventional consumption modes? *International Journal of Information Management*, 67, 102552. <https://doi.org/10.1016/j.ijinfomgt.2022.102552>
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>
- Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research*, 69(8), 3192–3198. <https://doi.org/10.1016/j.jbusres.2015.12.008>
- Gaskin, J., James, M., & Lim, J. (2019). Master Validity Tool, AMOS Plugin. Gaskination's StatWiki.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis: Pearson New International Edition* (7th edition). Pearson Education Limited.
- Heath, M., Appan, R., & Henry, R. (2022). Value alignment's role in mitigating resistance to IT use: The case of physicians' resistance to electronic health record systems. *Information and Management*, 59(8), 103702. <https://doi.org/10.1016/j.im.2022.103702>
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007%2Fs11747-014-0403-8>
- Ilie, V., & Turel, O. (2020). Manipulating user resistance to large-scale information systems through influence tactics. *Information and Management*, 57(3). <https://doi.org/10.1016/j.im.2019.103178>
- Jalo, H., & Pirkkalainen, H. (2024). Effect of user resistance on the organizational adoption of extended reality technologies: A mixed methods study. *International Journal of Information Management*, 75, 102731. <https://doi.org/10.1016/j.ijinfomgt.2023.102731>
- Khechine, H., Raymond, B., & Lakhal, S. (2023). Often asserted, never confirmed: The role of attitude in the acceptance of mandatory technology use, let's settle this question statistically

- for LMS use in the educational context. *Behaviour & Information Technology*, 42(11), 1801–1817. <https://doi.org/10.1080/0144929X.2022.2099297>
- Kim, H.-W., & Kankanhalli, A. (2009). Investigating user resistance to information systems implementation: A status quo bias perspective. *MIS Quarterly*, 33(3), 567–582. <https://doi.org/10.2307/20650309>
- Klaus, T., Wingreen, S. C., & Blanton, J. E. (2010). Resistant groups in enterprise system implementations: A Q-methodology examination. *Journal of Information Technology*, 25(1), 91–106. <https://doi.org/10.1057/jit.2009.7>
- Kumar, R. R., & Alok, K. (2020). Adoption of electric vehicle: A literature review and prospects for sustainability. *Journal of Cleaner Production*, 253, 119911. <https://doi.org/10.1016/j.jclepro.2019.119911>
- Lajunen, A., Sainio, P., Laurila, L., Pippuri-Mäkeläinen, J., & Tammi, K. (2018). Overview of Powertrain Electrification and Future Scenarios for Non-Road Mobile Machinery. *Energies*, 11(5), 1184. <https://doi.org/10.3390/en11051184>
- Lajunen, A., Suomela, J., Pippuri, J., Tammi, K., Lehmuspelto, T., & Sainio, P. (2016). Electric and Hybrid Electric Non-Road Mobile Machinery – Present Situation and Future Trends. *World Electric Vehicle Journal*, 8(1), 172–183. <https://doi.org/10.3390/wevj8010172>
- 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>
- Laumer, S., Maier, C., Eckhardt, A., & Weitzel, T. (2016). User personality and resistance to mandatory information systems in organizations: A theoretical model and empirical test of dispositional resistance to change. *Journal of Information Technology*, 31(1), 67–82. <https://doi.org/10.1057/jit.2015.17>
- Lončarević, Š., Ilinčić, P., Šagi, G., & Lulić, Z. (2022). Problems and Directions in Creating a National Non-Road Mobile Machinery Emission Inventory: A Critical Review. *Sustainability*, 14(6), 3471. <https://doi.org/10.3390/su14063471>
- Maier, C., Thatcher, J. B., Grover, V., & Dwivedi, Y. K. (2023). Cross-sectional research: A critical perspective, use cases, and recommendations for IS research. *International Journal of Information Management*, 70, 102625. <https://doi.org/10.1016/j.ijinfomgt.2023.102625>
- Malhotra, N. K., Kim, S. S., & Patil, A. (2006). Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research. *Management Science*, 52(12), 1865–1883. <https://doi.org/10.1287/mnsc.1060.0597>
- Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, 3(1), 97–110. <https://doi.org/10.21500/20112084.854>
- McKinsey & Company (2023). Built for purpose: EV adoption in light commercial vehicles. Available at <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/built-for-purpose-ev-adoption-in-light-commercial-vehicles>
- Miller, B. K., & Simmering, M. J. (2023). Attitude Toward the Color Blue: An Ideal Marker Variable. *Organizational Research Methods*, 26(3), 409–440. <https://doi.org/10.1177/10944281221075361>
- Möhlmann, M. (2015). Collaborative consumption: Determinants of satisfaction and the likelihood of using a sharing economy option again. *Journal of Consumer Behaviour*, 14(3), 193–207. <https://doi.org/10.1002/cb.1512>
- 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>
- Polites, G. L., & Karahanna, E. (2012). Shackled to the Status Quo: The Inhibiting Effects of Incumbent System Habit, Switching Costs, and Inertia on New System Acceptance. *MIS Quarterly*, 36(1), 21–42. <https://doi.org/10.2307/41410404>
- Schmalfuß, F., Mühl, K., & Krems, J. F. (2017). Direct experience with battery electric vehicles (BEVs) matters when evaluating vehicle attributes, attitude and purchase intention.

- Transportation Research Part F: Traffic Psychology and Behaviour, 46, 47–69.
<https://doi.org/10.1016/j.trf.2017.01.004>
- Selander, L., & Henfridsson, O. (2012). Cynicism as user resistance in IT implementation. *Information Systems Journal*, 22(4), 289–312. <https://doi.org/10.1111/j.1365-2575.2011.00386.x>
- Shi, X., Lin, Z., Liu, J., & Hui, Y. K. (2018). Consumer loyalty toward smartphone brands: The determining roles of deliberate inertia and cognitive lock-in. *Information & Management*, 55(7), 866–876. <https://doi.org/10.1016/j.im.2018.03.013>
- Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D. J., & Biondi, F. (2019). Assessing the visual and cognitive demands of in-vehicle information systems. *Cognitive Research: Principles and Implications*, 4(1), 18. <https://doi.org/10.1186/s41235-019-0166-3>
- Tarafdar, M., Tu, Q., Ragu-Nathan, B. S., & Ragu-Nathan, T. S. (2007). The Impact of Technostress on Role Stress and Productivity. *Journal of Management Information Systems*, 24(1), 301–328. <https://doi.org/10.2753/MIS0742-1222240109>
- 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.
<https://doi.org/10.2307/30036540>
- Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, 44(1), 119–134. <https://doi.org/10.1007/s11747-015-0455-4>
- Williams, L. J., Hartman, N., & Cavazotte, F. (2010). Method variance and marker variables: A review and comprehensive cfa marker technique. *Organizational Research Methods*, 13(3), 477–514. <https://doi.org/10.1177/1094428110366036>

GENERATIONAL RESEARCH ON TECHNOLOGY ACCEPTANCE AND USE: WHAT'S POPULAR ISN'T ALWAYS RIGHT

ANNA SELL,¹ MARKUS MAKKONEN,² PIIRKKO WALDEN³

¹ Linnaeus University, Faculty of Technology, Växjö, Sweden
anna.sell@lnu.se

² Tampere University, Faculty of Management and Business, Tampere, Finland
markus.makkonen@tuni.fi

³ Åbo Akademi University and Institute for Advanced Management Systems Research,
Åbo, Finland
pirkko.walden@abo.fi

Generations are a popular tool used by practitioners and researchers to divide consumers into groups that purportedly share values and characteristics due to shared life experiences. In the technology use realm, baby boomers' attitudes and behaviors are assumed to differ from millennials due to their formative years having been markedly different regarding technology access and exposure. While the idea is intuitively compelling, in this paper, we discuss problems with the concept of generations and the proposed mechanisms behind generational differences. We present findings from a systematic literature review investigating generational assumptions and inferences in technology acceptance and use research, revealing a need for a stronger theoretical grounding of generational assumptions and to evaluate other possible causes for differences.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.34](https://doi.org/10.18690/um.fov.4.2025.34)

ISBN

978-961-286-998-4

Keywords:

generations,
digital natives,
stereotypes,
generational cohorts,
technology acceptance



University of Maribor Press

1 Introduction

Generational differences fascinate academics and practitioners alike, as generational monikers and their attached implications are easy to understand and identify with. Generational cohorts group individuals by birth years and are associated with assumed descriptive characteristics. E.g., millennials are born between 1981 and 1996 (Fry, 2020) and described as “*ambitious, entitled and technologically savvy*” in the workplace (Gabriel et al., 2020). Digital natives and digital immigrants describe a generational dichotomy between the ‘natives’ who grew up with technology and ‘immigrants’ who learned to use it later in life (Prensky, 2001). Digital natives are claimed to be inherently comfortable with technology use, which also purportedly makes them think and behave differently than digital immigrants (Judd, 2018; Mertala et al., 2024). A central premise in most generational research is that core events during an individual’s formative years, such as wars and new innovations, have a profound impact on their subsequent attitudes and behavior.

While popular, generational approaches have been criticized for ascribing every individual within a broad cohort the same values, attitudes, and behavior, for failing to recognize that important life events can impact individuals differently, and for ignoring other than historical influences on human development and the fact that experiences during the entire lifespan have an impact on the individual (Beier et al., 2022). In empirical research, digital natives exhibit diverse attitudes, preferences, use patterns, and access to technologies, they are not uniformly skillful and innovative technology users, and they do not natively speak “the language of computers” (Jones, 2011; Mertala et al., 2024). Over-emphasizing the role of generational membership can be described as causal over-simplification. Purhonen (2016) coins the term *generationalism* for the practice of using generations as the primary explanatory factor for phenomena in the social world. Incorrect categorizations or an over-reliance on generational cohorts can lead to negative implications such as internalized stereotypes (Birkland, 2024), age discrimination (Cox and Coulton, 2015), and poor research designs, for example choosing samples that arbitrarily exclude certain age groups (Goodwin et al., 2023).

Generational research has been extensively scrutinized within workplace research by, e.g., Costanza et al. (2020, 2023) and Rudolph et al. (2018), while a wide and critical debate on the differences between digital natives and digital immigrants can

be found in educational research, as exemplified by Bennett and Maton (2010), Brown and Czerniewicz (2010), and Jones (2011). This dialogue is less prevalent in information systems (IS) research. As many characteristics assigned to various generations describe use of technology or attitudes towards technology, we believe it is important to discuss the possibilities and drawbacks of generational lenses in IS and especially in technology acceptance and use research. In this paper, we report on a systematic literature review with the aims to (1) examine the prevalence of generational research in premier IS outlets, (2) identify problems, fallacies, and good practices within our sample, and (3) provide guidelines for IS researchers interested in applying generational concepts in their research. In the following, we outline the main problems and fallacies that previous research has recognized in generational research and construct a basis for analyzing the prevalence of these problems in our sample.

2 Previous Research

Researchers purport that the empirical evidence for the existence of generational differences is lacking or undistinguishable from alternative explanations (Rudolph et al., 2019; Costanza et al., 2012). Costanza et al. (2020) identify four main categories of issues in generational research: (1) theoretical, (2) methodological, (3) practical, and (4) legal issues. We focus on the first three, as legal issues are beyond the practical scope of our study. Theoretical issues concern whether the conducted research and the drawn inferences are grounded on robust theoretical reasoning in terms of (a) specifying the historical event assumed to produce the differences in a particular birth cohort, (b) specifying what these differences are, and (c) specifying what mechanisms produce these differences (Costanza et al., 2020). Is it reasonable to argue that all members of a particular birth cohort have experienced a historical event in the same way and, thus, can all be assumed to exhibit the differences in the same quantities and qualities? E.g., digital natives are assumed to be more tech-savvy than digital immigrants due to their greater exposure to digital technology during childhood. However, this exposure is not uniform and can vary based on factors such as socioeconomic status in adolescence, leading to differences in technological proficiency among individuals within the same generation. The risk of this fallacy can be considered particularly great when dealing with heterogenous populations or when making comparisons between different countries or cultures (Costanza et al., 2020).

Rudolph et al. (2018) and Rauvola et al. (2019) question the assumption that events stop influencing the development of individuals after adolescence. They suggest that the consequences are dynamic rather than static and might grow, shrink, or even vanish over time. Rauvola et al. (2019) point out that current events are equally likely to influence how people think and behave as core events during adolescence. E.g., most digital immigrants may have been less technologically savvy than most digital natives when entering work life due to lesser exposure to digital technologies in their youth. But the exposure to digital technologies at work may promote their technological skills beyond those of digital natives less exposed to digital technologies at work. In other words, the impact of early life events is exaggerated, and alternative explanations and influences are routinely ignored (Rudolph et al., 2018; Costanza et al., 2023). The risk of falling victim to this fallacy is particularly great when studying older generations because there has been more time for potential development after the historical event.

The methodological issues concern data collection and data analysis. Generational cohorts are commonly operationalized inconsistently, meaning that instead of using the same range of birth years to differentiate between generational cohorts across studies, some studies use arbitrary ranges (Costanza et al., 2012; Rudolph et al., 2018). Ranges may even be entirely driven by the data, for example, to divide the sample into equally sized generational cohorts. Sampling issues are also prevalent, such as whether samples are representative of the intended generational cohorts and sizeable enough for statistical analyses. E.g., Brown and Czerniewicz (2010) criticize the use of student samples to represent a generation, as students are an 'elite' subgroup of the population. Costanza et al. (2020, 2023) criticize the practice of studying generations in isolation. Rauvola et al. (2019) draw attention to the use of subgroup analyses as an error-prone strategy to tease out differences between artificially created groups, especially in the absence of strong theoretical support for the groupings.

A fundamental issue in data analysis is that no statistical analysis method is able to disentangle potential cohort effects (i.e., the differences attributable to the birth year of individuals) from potential age and period effects (i.e., the differences attributable to the age of the individuals or the time period when the data was collected). However, Costanza et al. (2023) suggest that some methods still perform better than others. E.g., instead of using cross-sectional study designs with statistical analysis

methods like t-tests and the analysis of variance (ANOVA), they recommend time-lagged study designs with statistical approaches like cross-temporal meta-analysis (CTMA) and cross-classified hierarchical linear modeling (CCHLM). We summarize the main theoretical and methodological problems in Table 1.

Table 1: Theoretical and methodological problems identified in previous research

Identified problems	Sources
Fuzzy categorization of generations	Costanza et al., 2012; Rudolph et al., 2018; Jones, 2011
Assumption that major events are experienced and impact individuals in the same way	Costanza et al., 2020
Cross-sectional study designs, unsuitable statistical methods	Costanza et al., 2020; Costanza et al., 2023
Assumption that generational differences remain static	Rudolph et al., 2018; Rauvola et al., 2019
Generations studied in isolation	Costanza et al., 2020; Costanza et al., 2023
Student samples representing a generation	Brown and Czerniewicz, 2010
Geographical or cultural context not adapted to	Costanza et al., 2020
Citing unsubstantiated claims (certainty-complacency spiral)	Bennett and Maton, 2010
Only history-related influences considered	Rudolph et al., 2018; Rauvola et al., 2019; Costanza et al., 2023
Lack of acknowledgement and effort to separate age, period, and cohort effects	Costanza et al., 2023
Digital native fallacy; describing young people as inherently tech-savvy	Jones, 2011; Mertala et al., 2024; Bennett and Maton, 2010

Finally, the practical issues concern whether the implications drawn from the conducted research have any practical value (Costanza et al., 2020). In the IS context, one can consider whether results can be used as a basis for managerial actions, e.g., to improve a system. We will next present our study scrutinizing theoretical, methodological, and practical issues in IS technology acceptance and use studies making use of generational concepts.

3 Methodology

Searches were conducted in April 2024 in the Web of Science (WoS) database, limited to research published during the last five years, and only to journals on the AIS Senior Scholars' List of Premier Journals, including those recommended by AIS Special Interest Groups (Association for Information Systems, 2023). Search terms used were "[*generation name**]" AND "*technology acceptance*" OR "[*generation name**]" AND "*technology use*" in all fields. Generations and concepts included in the searches

were: millennials (1 article found), Generation Z (5), Generation Y (0), Generation X (4), boomer (2), generational (10), cohort (129), digital native (6) and digital immigrant (1), amounting to a total of 158 peer-reviewed articles. Subsequently, the websites for the outlets were searched to identify 13 additional articles meeting the selection criteria. After screening the abstracts and removing duplicates, 28 articles from 8 outlets were selected for final analysis. Table 2 shows the distribution of articles over time and outlets. A complete list of selected articles is presented in Appendix A.

Table 2: Overview of selected articles

Outlet	2020	2021	2022	2023	2024	Total
Computers in Human Behavior	2	4	2	3	2	13
European Journal of Information Systems			1	1		2
Information & Management	1					1
Information Systems Frontiers			1			1
Information Technology & People			1	1		2
Journal of Medical Internet Research	2			2	1	5
Organizational Behavior and Human Decision Processes		1				1
Technological Forecasting and Social Change	1		1	1		3
Total	6	5	6	8	3	28

The selected articles were read in full by three researchers over several iterative rounds. During the first round, core research choices were extracted from the articles: (a) generations studied in the research and their indicated start/end years, (b) technolog(ies) examined in the study, (c) geographical/cultural context, (d) underlying theoretical framework, (e) sample description, and (f) methodology. During the second round, the articles were scrutinized for generational assumptions or descriptions underpinning the research and generational inferences drawn from the results. The third round was carried out to discover the presence of any generational research problems or fallacies found in previous research and to identify problems not described in previous studies. Care was taken also to recognize any insightful research choices avoiding the above-described problems and fallacies.

4 Results

Table 3: Generations, countries, and technological contexts of the selected articles

Study	Generations	Country	Context
[1]	Baby boomers, Gen X, Digital natives	Germany	Internet, social media
[2]	Millennials, Gen Z	South Korea	Metaverse platforms
[3]	Digital immigrants, Digital natives	China	Data analytics
[4]	Millennials	N/A	Digital assistants
[5]	Digital immigrants, Digital natives	United States	Online health information seeking
[6]	Millennials	South Africa	Mobile banking
[7]	Baby boomers, Gen X, Millennials	Hong Kong	Gaming
[8]	Gen X, Gen Y, Gen Z	Ghana	E-learning
[9]	Any	N/A	Virtual reality
[10]	Digital natives, other generations	United States	Online learning tool use
[11]	Gen Z, older generations	Switzerland	Health information seeking, digital health literacy
[12]	Digital natives, Digital immigrants	India	Learning management system use
[13]	Gen Z, Millennials, Gen X, Baby boomers	United States	Smartphone use
[14]	Silent generation, Baby boomers	Israel	Digital search engine use
[15]	Junior researchers, tenure-track faculty, tenured faculty	N/A	Pre-registration in open science
[16]	Gen Z	Spain	Mobile learning application use
[17]	Digital natives, Digital immigrants, Silent generation, Baby boomers, Gen X, Millennials	United States	Mobile technology adoption
[18]	Digital natives, Digital immigrants	Europe	Routine and innovative IS use
[19]	Gen Z	Italy	Smartphone addiction, online compulsive buying
[20]	Gen Z	Malaysia, Turkey	AI products
[21]	Gen X, Gen Y	Oman	IoT-enabled healthcare applications
[22]	All	N/A	'Cutting-edge' technologies
[23]	Digital immigrants, Digital natives	United States	Technology interruptions
[24]	Digital natives	France	Smart home technologies
[25]	Baby boomers, Gen Z	United States	Social media
[26]	Baby boomers, Gen X, Millennials	United States	Voice-activated smart home devices
[27]	Digital natives	Myanmar	Consumption patterns, digital flexibility
[28]	Baby boomers, Gen X, Gen Y, Gen Z	Thailand	Mobile health app

As shown in Table 3, the studies have been conducted in a variety of country contexts and with a diverse range of target technologies. Most of the articles study two or more generations, but seven focused on only one (e.g., articles [4] and [6] on millennials). Methodologically, most articles utilize surveys and quantitative

methods, but, e.g., articles [1], [13, study 1], and [23] conducted interviews and qualitative analysis. In turn, articles [9] and [22] are literature reviews. In the following, we discuss theoretical, methodological, and practical issues identified in our sample.

4.1 Theoretical issues

We identified three kinds of theoretical issues in the reviewed articles: (1) missing or ambiguous theoretical reasoning, (2) using generations as a proxy for other theoretical constructs, and (3) conceptual confusion regarding generations and age groups. In the case of missing or ambiguous theoretical reasoning, the reviewed article either lacked all argumentation or provided ambiguous argumentation for how and why the studied generations are assumed to differ, as well as how and why these differences are assumed to impact the phenomenon under investigation. An example of this is article [10], which hypothesizes digital nativity to act as a control variable for academic enjoyment, academic anxiety, satisfaction with learning process, and personal performance but does not present any theoretical argumentation for the proposed effect. Another example is article [16], which evaluates the efficacy and use of a new mobile learning platform among Gen Z, motivated by the statement that this generation “*has unique characteristics that require new teaching strategies*” (p. 2). However, the article does not provide details on what these characteristics are, and how and why they result in the requirement of new teaching strategies.

Many of the reviewed articles use generational cohorts as a proxy for other characteristics without actually measuring those characteristics. For example, some studies assume that younger generations, such as Gen Z, are more innovative than older generations, like baby boomers. While they might provide reasoning for why this difference is expected, they do not test whether the younger participants in their study really are more innovative than the older participants. Instead, they take the assumed generational difference as given and use it as the foundation of their analysis. This approach is problematic because it relies on stereotypical assumptions rather than empirical evidence. If innovation levels vary within generations – meaning that not all young people are more innovative than all older people – then using generations as a proxy for innovation can lead to misleading conclusions. An example of this is article [18], which studies the differences between

digital natives (DNs) and digital immigrants (DIs) and states: *"to communicate with colleagues, DNs use social networks while DIs use traditional communication modes"* (p. 2804) and *"while DI usually resist new technology, DN are more receptive and open to them"* (p. 2806). Although these characterizations may hold true for some digital natives and digital immigrants, they may also vary considerably within the group. Measuring the actual use of communication tools and level of technology resistance is a more reliable approach than assuming it from the fact that they belong to a specific generation.

Some of the reviewed articles did not consistently use generations (individuals born within a specific range of years) and age groups (individuals of a specific age) as two separate concepts but mixed them together, resulting in conceptual ambiguity. An example of this is article [13], which studies the interplay between smartphone use, flow, and well-being, focusing on two age groups: individuals under 24 years and individuals aged 24 years or over. Despite defining these as age groups, the article frequently refers to the younger group as Gen Z. This creates confusion on what is actually being measured: the differences between age groups (which would apply to all individuals of a certain age, regardless of birth cohort) or between generations (which would apply specifically to people born within a particular time range, independent of their current age)? The lack of clarity in distinguishing these concepts weakens the study's theoretical foundation and makes it difficult to interpret the findings correctly.

4.2 Methodological issues

We identified four kinds of methodological issues in the reviewed papers: (1) inappropriate operationalization of generations, (2) inconsistent operationalization of generations, (3) inappropriate statistical analyses, and (4) non-representative or small generational samples. In the first case, the reviewed articles referred to 'generations' that were not really generations, meaning that their operationalization was not based on the range of years when its members were born but on some other characteristic of its members. For example, in article [12], digital natives and digital immigrants are operationalized based on individual computer engagement, measured with a scale by Charlton and Danforth (2007), whereas in article [10], being a digital native is measured with a scale based on Bennett et al. (2008). In turn, article [18] uses an even more complex operationalization of digital natives and digital immigrants based on cluster analysis with age and task experience as the clustering

variables. Finally, in article [15], generations are operationalized based on the academic career stage of individuals. On the positive side, this practice partly addresses the issue of using generations as a proxy for other theoretical constructs because the operationalizations of generations are now based on these other theoretical constructs instead of the birth years of their members. On the negative side, these unusual operationalizations of generations introduce even more ambiguity to generational research. To avoid conceptual confusion, it would be preferable to use, e.g., the terms computer engagement, age, and task experience directly, instead of calling them ‘digital nativity’.

Most of the articles operationalized generations based on birth years, but the range of birth years used for a generation varied between the articles, leading to inconsistent operationalizations. E.g., Gen Z is operationalized as individuals born in 1995 or later in article [19], as individuals born in 1997 and later in article [13], as individuals born between 1995 and 2010 in article [11], as individuals born between 1996 and 2003 in article [8], and as individuals born between the mid-90s and the 2000s in article [16].

In the case of inappropriate statistical analyses, none of the reviewed articles use the time-lagged study designs and advanced statistical analyses like cross-temporal meta-analysis (CTMA) recommended by Costanza et al. (2023) but instead use cross-sectional study designs and simpler analyses like t-tests (e.g., [11] and [14]) and ANOVA (e.g., [17] and [13]). Also regression analyses (e.g., [5], [10], [13], and [14]) and structural equation modelling with multiple group analysis (e.g., [3], [12], and [18]) were frequent. Most of the reviewed articles base their statistical analyses on samples not representative of the population or too small for making reliable statistical inferences. E.g., in articles [20], [24], [16], [19], [13], and [18], the sampled digital natives or members of Gen Z are almost exclusively students, meaning that the sample cannot be representative of the whole population of digital natives or Gen Z. This is particularly true in many developing countries where educational opportunities can be closely tied to socioeconomic status (e.g., article [27] conducted in Myanmar). Calling young samples by a generational moniker may partly stem from a desire to avoid criticism for relying on student samples, i.e., camouflaging a student sample by relabeling it as digital natives or Gen Z. In turn, although articles [17] and [13] study the silent generation and baby boomers, respectively, their samples consist

of only 53 members of the silent generation and 25 baby boomers, which can be considered very small samples for making any kind of statistical inferences.

Most of the articles employed quantitative methods, but article [1] carried out focus group interviews with baby boomers, Gen X, and digital natives. The focus groups were conducted in generation-specific groups instead of mixing generations. This study design means that any topics introduced by respondents in one group can appear unique for that generation if not discussed in any of the other groups, creating the illusion of generational differences. A mixed group design might have revealed that members of different generations, in fact, share thoughts and attitudes.

4.3 Practical issues

We identified two practical issues in the reviewed articles: (1) practically non-significant generational differences and (2) suggesting generational differences when only studying one generation. In the case of (1), the reviewed articles may have found statistically significant generational differences, but the size of these differences is practically non-significant. E.g., in article [17], very few statistically significant and even fewer practically significant differences were found between baby boomers, Gen X, and millennials in terms of their perceptions of innovation attributes and their innovation adoption intentions.

In the case of suggesting generational differences by studying only one generation, the articles included only one generation but either explicitly or implicitly suggested differences between this generation and one or more other generations without any actual evidence for the existence of such differences. An example of this is article [19] that studies Gen Z technology addiction. Based on the high level of smartphone addiction and frequent online compulsive buying found among Gen Z in the article, the authors suggest that this generation is especially plagued by addiction and needs special attention to address these problems. This suggestion may be misleading without finding out how frequent these problems are among other generations.

5 Discussion and conclusions

In our review of IS articles published in premier outlets, we found evidence of similar generational fallacies as presented in previous research from the workplace and educational contexts. The theoretical, methodological, and practical issues identified are important to heed in future research in order to advance our knowledge of age and generations in technology acceptance and use. While technology acceptance and use are mainly a secondary interest in other fields, they are central in IS. Researchers in our field should, thus, be diligent regarding conceptual soundness and clarity in any core concepts describing technology use. We propose that IS researchers:

- a. Discuss any theoretical constructs (e.g., computer engagement and computer experience) directly instead of using generations as a proxy.
- b. Measure any theoretical constructs and characteristics of interest instead of assuming their levels in the studied population (e.g., measure the level of innovativeness in the sample).
- c. Be mindful of the difference between generations (a range of birth years) and age.
- d. Avoid sweeping statements about generations (e.g., “members of Gen X are independent”) and take care to explain the theoretical background, meaning, and relevance of any generational descriptions in relation to the study at hand.

We urge IS researchers to not fall into the certainty-complacency spiral when theorizing about generational differences, referring to the practice of supporting arguments by uncritically citing unsubstantiated claims made in other studies (Bennett & Maton, 2010). This is when researchers cite vague statements, such as “younger generations show a dramatic shift in behavior and attitudes”, but the behavior and attitudes are not described and elaborated further either in the cited or the citing research. This leads to a rise in volume and visibility of generational research without an accompanying rise in actual evidence on the existence and meaning of generations.

Theoretically, it would, in most cases, be advisable to forego generational descriptors altogether and avoid ascribing a wide set of characteristics to a large group of people, as the generational approach ignores the high variety within a group and the

possibility of changes over time. Assuming static generational differences is a lost opportunity for researchers to investigate how technology use evolves over the lifespan. Furthermore, researchers designing their studies on the assumption of generational differences risk limiting the depth and relevance of their findings. E.g., researchers might exclude certain generational groups from their sample altogether (typically older generations) or assume that their findings are not applicable to other generations than the one studied. When groupings are made based on wide generational cohorts, researchers forego the opportunity to test for other, more theoretically sound explanations for the potential differences. Also, researchers rarely test whether the generational assumptions forming their study designs are, in fact, correct.

In our review, we see a high prevalence of studies with student samples that refer to these samples as digital natives or Gen Z and generalize the results to the entire population. We caution against assuming that student samples are representative of the entire generation. Students are typically different regarding many characteristics known to influence technology use, such as computer experience and socio-demographic standing. In a similar vein, many studies take place in geographical and cultural contexts highly dissimilar from the countries where the generational descriptions were created. When using generational descriptors from a different context, care must be taken to ensure their applicability regarding, e.g., core events and levels of technological diffusion. In some studies, generational monikers were used more as a catchphrase than a theoretical construct. In these cases, the generations were mentioned only briefly and mostly in the title and introduction of the article. The study designs did not make use of generational cohorts and generational lenses were not used to interpret the results. This practice is essentially harmless, as researchers then steer clear of most generational fallacies mentioned.

In conclusion, we advise researchers to be critical and ambitious when choosing the core constructs for their studies. Many generational studies are explicitly or between the lines grounded on the digital native discourse. While popular, the myth of digital nativeness has not been theoretically proven. Furthermore, we can better advance our theories and understanding when we study specific constructs of interest rather than blanket descriptors of assumed differences. Finally, through striving for a more nuanced understanding of individuals' technology acceptance and use we can avoid

inadvertently strengthening the all too prevalent age discrimination and stereotyping present in technology use research and practice.

References

- Association for Information Systems (AIS) (February 2023). *Senior Scholars' List of Premier Journals*. <https://aisnet.org/page/SeniorScholarListofPremierJournals>. Accessed in April, 2024.
- Beier, M. E., Kanfer, R., Kooij, D. T., & Truxillo, D. M. (2022). What's age got to do with it? A primer and review of the workplace aging literature. *Personnel Psychology*, 75(4), 779–804.
- Bennett, S., & Maton, K. (2010). Beyond the 'digital natives' debate: Towards a more nuanced understanding of students' technology experiences. *Journal of Computer Assisted Learning*, 26(5), 321–331.
- Bennett, S., Maton, K., & Kervin, L. (2008). The 'digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology*, 39(5), 775–786.
- Birkland, J. L. H. (2024). How older adult information and communication technology users are impacted by aging stereotypes: A multigenerational perspective. *New Media & Society*, 26(7), 3967–3988.
- Brown, C., & Czerniewicz, L. (2010). Debunking the 'digital native': beyond digital apartheid, towards digital democracy. *Journal of Computer Assisted Learning*, 26(5), 357–369.
- Charlton, J. P., & Danforth, I. D. (2007). Distinguishing addiction and high engagement in the context of online game playing. *Computers in Human Behavior*, 23(3), 1531–1548.
- Costanza, D. P., Badger, J. M., Fraser, R. L., Severt, J. B., & Gade, P. A. (2012). Generational differences in work-related attitudes: A meta-analysis. *Journal of Business and Psychology*, 27(4), 375–394.
- Costanza, D. P., Finkelstein, L. M., Imose, R. A., & Ravid, D. M. (2020). Inappropriate inferences from generational research. In B. J. Hoffman, M. K. Shoss, & L. A. Wegman (Eds.), *The Cambridge handbook of the changing nature of work* (pp. 20–41). Cambridge University Press.
- Costanza, D. P., Rudolph, C. W., & Zacher, H. (2023). Are generations a useful concept? *Acta Psychologica*, 241, 104059.
- Cox, C. B., & Coulton, G. (2015). Fire all the boomers: How generational labeling legitimizes age discrimination. *Industrial and Organizational Psychology*, 8(3), 372–376.
- Gabriel, A. G., Alcantara, G. M., & Alvarez, J. D. (2020). How do millennial managers lead older employees? The Philippine workplace experience. *Sage Open*, 10(1).
- Goodwin, V. A., Low, M. S., Quinn, T. J., Cockcroft, E. J., Shepherd, V., Evans, P. H., ... & Witham, M. D. (2023). Including older people in health and social care research: best practice recommendations based on the INCLUDE framework. *Age and Ageing*, 52(6), afad082.
- Jones, C. (2011). Students, the net generation, and digital natives: Accounting for educational change. In M. Thomas (Ed.), *Deconstructing digital natives: Young people, technology, and the new literacies* (pp. 30–46). Routledge.
- Judd, T. (2018). The rise and fall (?) of the digital natives. *Australasian Journal of Educational Technology*, 34(5).
- Mertala, P., López-Pernas, S., Vartiainen, H., Saqr, M., & Tedre, M. (2024). Digital natives in the scientific literature: A topic modeling approach. *Computers in Human Behavior*, 152, 108076.
- Fry, R. (2020). *Millennials overtake Baby Boomers as America's largest generation*. <https://www.pewresearch.org/short-reads/2020/04/28/millennials-overtake-baby-boomers-as-americas-largest-generation/>. Accessed in June, 2024.
- Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, 9(5), 1–6.
- Purhonen, S. (2016). Generations on paper: Bourdieu and the critique of 'generationalism'. *Social Science Information*, 55(1), 94–114.

- Rauvola, R. S., Rudolph, C. W., & Zacher, H. (2019). Generationalism: Problems and implications. *Organizational Dynamics*, 48(4), 100664.
- Rudolph, C. W., Rauvola, R. S., Costanza, D. P., & Zacher, H. (2020). Answers to 10 questions about generations and generational differences in the workplace. *Public Policy & Aging Report*, 30(3), 82–88.
- Rudolph, C. W., Rauvola, R. S., & Zacher, H. (2018). Leadership and generations at work: A critical review. *Leadership Quarterly*, 29(1), 44–57.

Appendix A: Articles selected for review

1. Mayer, G., Alvarez, S., Gronewold, N., & Schultz, J. H. (2020). Expressions of individualization on the internet and social media: Multigenerational focus group study. *Journal of Medical Internet Research*, 22(11), e20528.
2. Oh, H. J., Kim, J., Chang, J. J., Park, N., & Lee, S. (2023). Social benefits of living in the metaverse: The relationships among social presence, supportive interaction, social self-efficacy, and feelings of loneliness. *Computers in Human Behavior*, 139, 107498.
3. Shao, Z., Benitez, J., Zhang, J., Zheng, H., & Ajamieh, A. (2023). Antecedents and performance outcomes of employees' data analytics skills: An adaptation structuration theory-based empirical investigation. *European Journal of Information Systems*, 32(6), 921–940.
4. Sharma, M., Joshi, S., Luthra, S., & Kumar, A. (2022). Impact of digital assistant attributes on millennials' purchasing intentions: A multi-group analysis using PLS-SEM, artificial neural network and fsQCA. *Information Systems Frontiers*, 26(3), 943–966.
5. Sinha, J., & Serin, N. (2024). Online health information seeking and preventative health actions: Cross-generational online survey study. *Journal of Medical Internet Research*, 26, e48977.
6. Thusi, P., & Maduku, D. K. (2020). South African millennials' acceptance and use of retail mobile banking apps: An integrated perspective. *Computers in Human Behavior*, 111, 106405.
7. Wang, H. Y., & Cheng, C. (2021). New perspectives on the prevalence and associated factors of gaming disorder in Hong Kong community adults: A generational approach. *Computers in Human Behavior*, 114, 106574.
8. Yawson, D. E., & Yamoah, F. A. (2021). Gender variability in e-learning utility essentials: Evidence from a multi-generational higher education cohort. *Computers in Human Behavior*, 114, 106558.
9. YiFei, L., & Othman, M. K. (2024). Investigating the behavioural intentions of museum visitors towards VR: A systematic literature review. *Computers in Human Behavior*, 155, 108167.
10. James, T. L., Zhang, J., Li, H., Ziegelmayer, J. L., & Villacis-Calderon, E. D. (2022). The moderating effect of technology overload on the ability of online learning to meet students' basic psychological needs. *Information Technology & People*, 35(4), 1364–1382.
11. Jiao, W., Chang, A., Ho, M., Lu, Q., Liu, M. T., & Schulz, P. J. (2023). Predicting and empowering health for Generation Z by comparing health information seeking and digital health literacy: Cross-sectional questionnaire study. *Journal of Medical Internet Research*, 25, e47595.
12. Kesharwani, A. (2020). Do (how) digital natives adopt a new technology differently than digital immigrants? A longitudinal study. *Information & Management*, 57(2), 103170.
13. Lavoie, R., & Zheng, Y. (2023). Smartphone use, flow and wellbeing: A case of Jekyll and Hyde. *Computers in Human Behavior*, 138, 107442.
14. Lissitsa, S., Zychlinski, E., & Kagan, M. (2022). The Silent Generation vs Baby Boomers: Socio-demographic and psychological predictors of the “gray” digital inequalities. *Computers in Human Behavior*, 128, 107098.
15. Logg, J. M., & Dorison, C. A. (2021). Pre-registration: Weighing costs and benefits for researchers. *Organizational Behavior and Human Decision Processes*, 167, 18–27.
16. Lozano-Lozano, M., Galiano-Castillo, N., Fernández-Lao, C., Postigo-Martin, P., Álvarez-Salvago, F., Arroyo-Morales, M., & Cantarero-Villanueva, I. (2020). The Ecofisio mobile app for assessment and diagnosis using ultrasound imaging for undergraduate health science students: Multicenter randomized controlled trial. *Journal of Medical Internet Research*, 22(3), e16258.
17. Magsamen-Conrad, K., & Dillon, J. M. (2020). Mobile technology adoption across the lifespan: A mixed methods investigation to clarify adoption stages, and the influence of diffusion attributes. *Computers in Human Behavior*, 112, 106456.
18. Martins, P., Picoto, W. N., & Bélanger, F. (2023). The effects of digital nativity on nonvolitional routine and innovative usage. *Information Technology & People*, 36(7), 2804–2825.

19. Mason, M. C., Zamparo, G., Marini, A., & Ameen, N. (2022). Glued to your phone? Generation Z's smartphone addiction and online compulsive buying. *Computers in Human Behavior*, 136, 107404.
20. Al-Sharaf, M. A., Al-Emran, M., Arpac, I., Iahad, N. A., AlQudah, A. A., Iranmanesh, M., & Al-Qaysi, N. (2023). Generation Z use of artificial intelligence products and its impact on environmental sustainability: A cross-cultural comparison. *Computers in Human Behavior*, 143, 107708.
21. Alraja, M. (2022). Frontline healthcare providers' behavioural intention to Internet of Things (IoT)-enabled healthcare applications: A gender-based, cross-generational study. *Technological Forecasting and Social Change*, 174, 121256.
22. Ameen, N., Hosany, S., & Tarhini, A. (2021). Consumer interaction with cutting-edge technologies: Implications for future research. *Computers in Human Behavior*, 120, 106761.
23. Baham, C., Kalgotra, P., Nasirpouri Shadbad, F., & Sharda, R. (2023). Generational differences in handling technology interruptions: A qualitative study. *European Journal of Information Systems*, 32(5), 858–878.
24. Baudier, P., Ammi, C., & Deboeuf-Rouchon, M. (2020). Smart home: Highly-educated students' acceptance. *Technological Forecasting and Social Change*, 153, 119355.
25. Bunker, C. J., & Kwan, V. S. (2024). Similarity between perceived selves on social media and offline and its relationship with psychological well-being in early and late adulthood. *Computers in Human Behavior*, 152, 108025.
26. Canziani, B., & MacSween, S. (2021). Consumer acceptance of voice-activated smart home devices for product information seeking and online ordering. *Computers in Human Behavior*, 119, 106714.
27. Han, J. H. (2023). Does 'born digital' mean 'being global' in characterizing millennial consumers in a less developed country context? – An empirical study in Myanmar. *Technological Forecasting and Social Change*, 195, 122801.
28. Inhusri T., Surangsirat D., Kwanmuang P., Poomivanichakij P., Apiwatgaroon P., Ongprakobkul S., Kongchu A., Klinpikul A., Taneeheng A., Pruphetkaew N., Thongseiratch T., Ngamchaliew P., Vichitkunakorn P. (2023). Association of generation and group size with the usage of a mobile health app in Thailand: Secondary analysis of the ThaiSook cohort study. *Journal of Medical Internet Research*, 25, e45374.

TOWARDS THE AUTOMATIC EXTRACTION OF DECISION MODEL & NOTATION FROM DUTCH LEGAL TEXT

ANNEMAE VAN DE HOEF, SAM LEEWIS, KOEN SMIT

HU University of Applied Sciences Utrecht, Digital Ethics, Utrecht, the Netherlands
annemae.vandehoef@hu.nl, sam.leewis@hu.nl, koen.smit@hu.nl

The translation from laws and regulations into actionable business rules remains challenging due to the complexity of Dutch legal text. In addition, the (semi-)manual translation of law into business rules is both time-consuming and error-prone. To address these issues, this research explores the use of Natural Language Processing (NLP) to automatically extract legal decisions and represent them in a Decision Model and Notation (DMN) model. For this purpose, existing research was reviewed to define requirements, which formed the basis for the NLP prototype. The current prototype evaluates an existing approach and aims to process unstructured Dutch legal text. However, a theoretical extension is proposed to address the structural complexity of extracting a DMN model from structured Dutch legal texts. Therefore, future research should focus on implementing the proposed approach and validating it in collaboration with legal analysts to extract a DMN model from structured Dutch legal texts.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.35](https://doi.org/10.18690/um.fov.4.2025.35)

ISBN

978-961-286-998-4

Keywords:

BRM,
DMN,
NLP,
Legislation,
legal text



University of Maribor Press

1 Introduction

Operational decisions play a crucial role in highly regulated industries (e.g., government) as they determine how laws, regulations, and policies are implemented in practice. These decisions are defined as: *“the act of determining an output value from a number of input values, using decision logic defining how the output is determined from the inputs”* (Object Management Group, 2019a). Organizations use business rules to structure and automate such operational decisions. Business rules are defined as: *“a statement that defines or constrains some aspect of the business intending to assert business structure or to control the behaviour of the business”* (Hay & Healy, 2000). Implementing well-defined business rules within an organization, such as in IT systems, can help prevent errors, save time, and reduce costs in software projects by, e.g., ensuring compliance with regulatory requirements (Morgan, 2002). However, for business rules to deliver these benefits, organizations must design, execute, and manage them. This could be achieved through Business Rule Management (BRM) (Graham, 2007). BRM is defined as: *“a systematic and controlled approach that supports the elicitation, design, specification, verification, validation, deployment, execution, evaluation, and governance of business decisions and business logic”* (Boyer & Mili, 2011; Morgan, 2002; Schlosser et al., 2014). Within BRM, the *elicitation* capability, in particular, remains time-consuming, e.g., as addressed by (Nelson et al. 2008). In addition, formalizing Dutch legal texts remains challenging due to their complex structure, with long sentences, passive voice, and extensive relative clauses (Bakker, de Boer, et al., 2022). For example, in the Netherlands, the translation process has caused years of delays in IT projects within the Dutch Tax Administration (Nederlandse Omroep Stichting, 2024). Moreover, mistranslation can have serious consequences, such as the Dutch childcare benefits scandal (Amnesty International, 2021) or fraud with personal budgets (Nederlandse Omroep Stichting, 2019).

Supporting governmental institutions in translating laws and regulations into business rules through the application of Natural Language Processing (NLP) could help law and rule analysts. NLP is used to extract semantics from text (Nadkarni et al., 2011). For example, research applied NLP in the legal domain to help save time (Barale et al., 2023; Hendrycks et al., 2021; Mokashi et al., 2024; Vayadande et al., 2024) and reduce errors (Barale et al., 2023; Vayadande et al., 2024). To facilitate the translation to business rules, several modeling languages have been developed that support the specification of business rules for subsequent implementation in IT

systems. These include the Semantics of Business Vocabulary and Business Rules (SBVR) (Object Management Group, 2019b), The Decision Model framework (TDM) (Von Halle & Goldberg, 2009), RuleSpeak (Ross, 2009), and Decision Model and Notation (DMN) (Object Management Group, 2019a). Among these, DMN is considered the industry standard for modeling decisions that define how business rules are implemented (Hasić et al., 2017; Leewis et al., 2020). In recent years, research has been conducted on NLP-based approaches for extracting (elements of) DMN models from natural language text (Arco et al., 2021; Etikala, 2021; Etikala et al., 2020; Goossens et al., 2021, 2022, 2023; Quishpi et al., 2021). For example, Goossens et al. (2023) use NLP to extract decision dependencies and decision logic for the creation of DMN models, while Quishpi et al. (2021) combine NLP with patterns to create DMN models. However, to the best of the authors' knowledge, no research has been conducted on an NLP-based approach for extracting DMN models from (Dutch) legal texts, nor on the requirements for such an approach. Therefore, the following research question is raised: *"In what way can a DMN model be automatically discovered from Dutch legal texts using NLP?"*

2 Background and Related Work

Before discussing DMN and the existing literature on the (semi-)automatic extraction of (elements of) DMN models from natural language text, the concept of BRM is further introduced. BRM consists of nine capabilities, which are performed (semi-)manually, as shown in Figure 1.

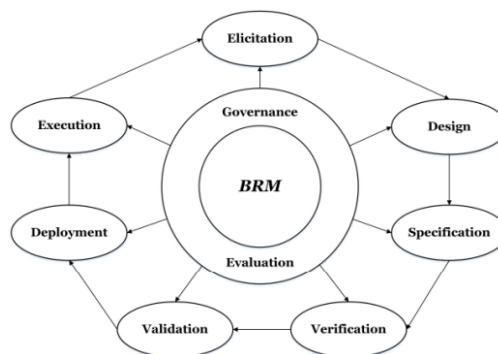


Figure 1: The nine BRM Capabilities (Smit, 2018)

First, an operational decision can be identified or modified through *elicitation*, based on, e.g., relevant legal knowledge. Second, *design* involves structuring the relevant information into a business rules architecture using a modeling language. Third, *specification* concerns defining the business logic. Fourth and fifth, *verification* focuses on ensuring logical consistency, while *validation* concerns confirming intended behavior. Sixth, *deployment* involves translating decisions into, e.g., implementation-dependent systems. Seventh, *execution* focuses on implementing business rules. Eighth and ninth, in parallel with the other capabilities, *governance* concerns ensuring traceability and version management, while *evaluation* involves monitoring implementation and performance (Boyer & Mili, 2011; Morgan, 2002; Smit, 2018). The business rules architecture defined in the *design* capability can, e.g., be modeled using SVBR (Object Management Group, 2019b), TDM (Von Halle & Goldberg, 2009), RuleSpeak (Ross, 2009), or DMN (Object Management Group, 2019a). Since DMN is considered an industry standard, it is used in this research.

DMN consists of two levels: the decision requirements level and the decision logic level (Hasić et al., 2017). The decision requirements level is represented by a Decision Requirement Diagram (DRD), which illustrates the requirements and dependencies of each decision. The decision logic level consists of a decision table for each decision in a DRD (Hasić et al., 2017). An example of both is shown in Figure 2. On the left, a decision is represented by (1), a dependency by (2), input data by (3), a knowledge source by (4), and business knowledge by (5) (Object Management Group, 2019a). On the right, a fact type is represented by (6) and a fact value by (7). An example of a decision rule would be: if the *Purpose of Use* is 'Education' and the *Type of Work* is 'Video,' then the *Permission Required* is 'false.'

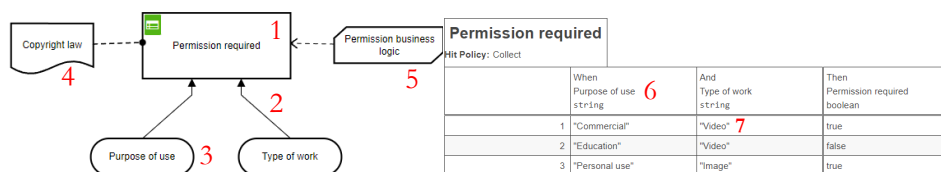


Figure 2: DRD (Left) and Decision Table (Right)

Research has been conducted on the NLP-based (semi-)automatic extraction of (elements of) DMN from natural language text. For example, a more recent approach is described by Goossens et al. (2023), which builds on previous work as

described in (Goossens et al., 2021, 2022). Goossens et al. (2023) present a three-step approach to extract a DRD and a single decision table, as also indicated in Figure 3:

1. The unstructured text undergoes coreference resolution (identifying words referring to the same concept (Ng & Cardie, 2001), followed by preprocessing, including lowercasing, sentence tokenization (e.g., specifying words or commas), and adding tokens to the glossary of the deep learning model.
2. A deep learning classifier then identifies relevant sentences containing decision-related information.
3. Next, Named Entity Recognition (NER) extracts decision logic and dependencies from these sentences. NER is used to identify entities of interest in text documents (Nadeau & Sekine, 2007).

The extracted information, i.e., decision dependencies and decision logic, is used to create a DMN model (Goossens et al., 2023). Goossens et al. (2023) concluded that the Bidirectional Encoder Representations from Transformers (BERT) outperformed the Bi-LSTM-CRF deep learning model in extracting dependencies. For text classification, BERT also achieved the best performance compared to traditional machine learning methods, including logistic regression, naive bayes, and Support Vector Machines (SVM). Since the authors did not focus on decision logic extraction, BERT was applied for that task without comparison to other approaches.

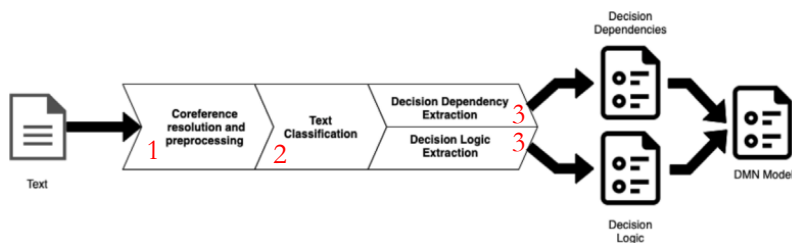


Figure 3: Implementation by Goossens et al. (2023)

One advantage of their approach is its language independence; for example, BERT_{je} (De Vries et al., 2019) can be applied to Dutch and CamemBERT (Martin et al., 2020) to French. The use of BERT makes the approach more resilient to emerging patterns. However, it has a limitation, as it uses only one decision table for all decisions, meaning that each decision does not have its dedicated table. In addition,

BERT has been trained only to consider a maximum of two decision dependency levels. Quishpi et al. (2021) also present an NLP-based semi-automatic approach for extracting a DMN model, similar to Goossens et al. (2023) in utilizing NLP to preprocess or parse natural language text. However, while Goossens et al. (2023) use NER for extraction, Quishpi et al. (2021) rely on English pattern-based rules (e.g., specific word structures) to identify decision logic and dependencies. This pattern-based method requires custom patterns for each case, making it less flexible compared to the approach by Goossens et al. (2023). In addition, not all decision expressions are covered by these patterns. However, the approach has the advantages of being extensible (by defining more patterns) and precise due to the clear pattern definitions. Furthermore, the authors extract both a DRD and corresponding decision tables. Two other NLP-based, framework-oriented automatic DMN extraction approaches are proposed by Etikala (2021) and Arco et al. (2021). Both address linguistic challenges, such as multiple interpretations and/or references in text. Etikala (2021) follows a pattern-based approach, similar to Quishpi et al. (2021), by identifying English decision patterns in sentences. The author also extracts a DRD and corresponding decision tables. In contrast, Arco et al. (2021) focus exclusively on extracting decision tables. Moreover, their approach does not support Dutch, as it relies on the Stanford NLP library, which does not support the Dutch language.

Although Goossens et al. (2023), Quishpi et al. (2021), Etikala (2021), and Arco et al. (2021) present approaches for the (semi-)automatic NLP-based extraction of (elements of) a DMN model from natural language text, there is currently no known approach for extracting a DMN model from Dutch legal text. In this regard, research has been conducted on automatic knowledge discovery from legal texts (Bartolini et al., 2004; Biagioli et al., 2005; Boella et al., 2013; Dragoni et al., 2018; Guissé et al., 2012; Kacfeh Emani, 2014; Michel et al., 2022; Palmirani et al., 2011; Sleimi et al., 2018, 2019; Spinosa et al., 2009; Wyner & Peters, 2011). While some of these extract DMN-related elements, e.g., decision rules (Michel et al., 2022), none of these approaches directly address DMN.

3 Research Method

This research adopted a Design Science Research (DSR) (Hevner et al., 2004) approach, supported by a narrative literature review (Paré et al., 2015), as illustrated in Figure 4. First, the relevant literature was reviewed. Second, requirements were defined based on the literature. Third and last, an artifact (prototype) was developed, assessed, and refined to support the extraction of DMN models from Dutch legal texts. Following DSR, the prototype was designed to address a practical business need, i.e., assisting with the manual translation of legal texts into executable business rules, and was informed by applicable frameworks and methods identified in the literature.

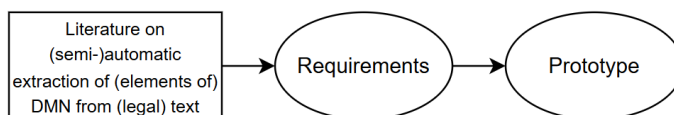


Figure 4: Steps used to implement the prototype

A narrative review was selected for this research, as it allows for a flexible examination of the most relevant literature. While this approach has the limitation of reduced transparency and reproducibility, it was appropriate for this research given the exploratory nature of the research and the goal of identifying relevant methods, frameworks, and requirements that can guide prototype development. Specifically, the objective was to identify existing solutions that extract structure using NLP from (legal) texts and generate (elements of) DMN models as output. To conduct the review, Google Scholar was used as the primary search engine as it is considered to be the most comprehensive academic databank (Gusenbauer, 2019). Based on the findings from this review, one or more of the following actions were taken:

- If an existing solution met the identified requirements, it was described and evaluated.
- If modifications or extensions were needed, the solution was refined based on insights from the literature.
- If an adequate solution was not found, a new approach was developed, integrating best practices and requirements extracted from prior research.

4 Data Collection and Analysis

The data collection for this research occurred between February and March 2025. Literature focused on extracting (elements of) DMN from natural language text was identified. It became evident that the body of literature on the (semi-)automatic extraction of DMN elements from (legal) texts is limited. To ensure completeness, the reference lists of the selected studies were examined (backward snowballing), confirming that no additional relevant sources were overlooked. This resulted in four relevant studies, as shown in Table 1. It is important to note that Etikala et al. (2020) and Goossens et al. (2021, 2022) were excluded as newer research provides a more comprehensive DMN extraction approach. The advantages and disadvantages relevant to the purpose of this research, specifically considering the extraction of DMN elements, robustness to new patterns (i.e., managing business rules can aid in maintaining systems in changing environments (Bajec et al., 2004)), especially with many Dutch laws having been in effect (Klein Haarhuis & Niemeijer, 2008), applicability to Dutch, and the consideration of edge cases, are outlined as follows:

Table 1: Advantages and disadvantages of relevant literature

Source	Advantages	Disadvantages
Goossens et al. (2023)	- Extracts DRD and one general decision table. - Natural language-independent. - More robust to changes.	- No decision table per decision. - Tested only on shallow (2-level) decision dependencies.
Quishpi et al. (2021)	- Extracts DRD and decision tables.	- Natural language-dependent; need to define Dutch patterns. - Does not cover all decision expressions. - Less robust to change; requires new patterns.
Etikala (2021)	- Extracts DRD and decision tables. - Addresses some natural language and decision modeling challenges.	- Natural language-dependent; need to define Dutch patterns. - Does not cover all decision expressions. - Less robust to change; requires new patterns.
Arco et al. (2021)	- Extracts decision tables. - More robust to changes. - Addresses some natural language characteristics.	- Does not extract DRD. - Single sentence analysis limits aggregating decision tables. - Does not support Dutch.

Based on the four studies focused on extracting DMN elements from natural language text, requirements for extracting a DMN model from Dutch legal texts were defined, as outlined in Table 2. Based on commonalities between the studies,

these requirements consist of having a text as input, being able to extract decision dependencies and/or logic, and modeling a DRD and/or logic. In addition, the requirement to address edge cases (i.e., ensuring all potential decision scenarios are accounted for) is added for completeness, such as potential errors in a DMN model or legal text, which is crucial for validating the DMN model or legal text. In this table, ‘X’ indicates full support, ‘x’ indicates partial support, and ‘-’ indicates no support. As shown in Table 2, Quishpi et al. (2021) do not fully address decision dependency extraction, decision logic extraction, or edge case coverage, as they acknowledge their approach’s limitations. A similar limitation applies to Etikala (2021), whose pattern-based approach suggests incomplete coverage of decision expressions. Arco et al. (2021) focus on decision tables rather than the DRD, while Goossens et al. (2023) only partially address edge cases due to a maximum of two-level decision dependencies.

Table 2: Identified requirements

High-Level Requirement / Source	Text as input	Decision dependency extraction	Decision logic extraction	Modeling DRD	Modeling decision logic	Edge case coverage
Goossens et al. (2023)	X	X	X	X	X	x
Quishpi et al. (2021)	X	x	x	X	X	x
Etikala (2021)	X	x	x	X	X	x
Arco et al. (2021)	X	x	X	-	X	-

The prototype must meet the requirements outlined in Table 2. Therefore, to develop a solution through a prototype, the approach by Goossens et al. (2023) was selected to evaluate whether it could work with (unstructured) Dutch legal text or if an extension would be necessary, as their approach best meets the requirements. Moreover, Goossens et al. (2023) was chosen because it is language-independent, does not require manual pattern definition, is more resistant to changes in legal texts, and facilitates the extraction of an (almost) complete DMN model, including a DRD and a general decision table with executable business rules. In contrast, Quishpi et al. (2021) and Etikala (2021) require predefined Dutch patterns, while Arco et al. (2021) focus only on decision logic and do not support Dutch. As a result, the approach by Goossens et al. (2023) is the most suitable for this research.

5 Prototype Development

The prototype follows the implementation by Goossens et al. (2023); for further details, we refer the to their paper. Three BERTje models were fine-tuned in this study, which is a BERT model trained in the Dutch language (De Vries et al., 2019). This required three labeled datasets, including Dutch unstructured text: one for text classification to categorize sentences and two for decision dependency and logic extraction using NER. The datasets were labeled using IOB tagging, as described by Goossens et al. (2023), and included unique sentences collected via ChatGPT: 400 decision dependency sentences, 400 decision logic sentences, and 200 irrelevant sentences. Furthermore, some of these sentences represented one of the four special cases described by Goossens et al. (2023), such as two if-else statements or three dependency levels. The text classification dataset contained all types of sentences, while the NER datasets included only decision dependency or decision logic sentences. The results of the BERTje models are presented in Table 3.

Table 3: Fine-tuned BERTje metrics

Text Classification	Precision	Recall	F1-Score
<i>Micro Average</i>	0.99 ± 0	0.99 ± 0	0.99 ± 0
<i>Macro Average</i>	0.99 ± 0	0.99 ± 0	0.99 ± 0
<i>Weighted Average</i>	0.99 ± 0	0.99 ± 0	0.99 ± 0
NER Dependency	Precision	Recall	F1-Score
<i>Micro Average</i>	0.82 ± 0.0078	0.89 ± 0.0084	0.86 ± 0.0082
<i>Macro Average</i>	0.83 ± 0.0085	0.89 ± 0.0079	0.86 ± 0.0082
<i>Weighted Average</i>	0.83 ± 0.0106	0.89 ± 0.0084	0.86 ± 0.0063
NER Logic	Precision	Recall	F1-Score
<i>Micro Average</i>	0.93 ± 0.014	0.94 ± 0.0094	0.93 ± 0.0097
<i>Macro Average</i>	0.88 ± 0.0329	0.88 ± 0.0227	0.88 ± 0.0274
<i>Weighted Average</i>	0.93 ± 0.0133	0.94 ± 0.0094	0.93 ± 0.0097

After fine-tuning three pre-trained BERTje models, the steps defined by Goossens et al. (2023) are followed. It is important to note that coreference resolution was not included in the prototype, as it might address errors that should instead be reflected in the DMN model. Therefore, as indicated by Goossens et al. (2023), the resulting prototype takes unstructured text as input, extracts decision dependencies and logic after text classification, and creates a DMN model using a single general decision table and DRD based on the extracted data. The source code, pseudocode, and HuggingFace links to the BERTje models can be found in an Open Science Framework repository (OSF) (Anonymous, 2025).

6 Validation

After the prototype was implemented, an unstructured Dutch legal text excerpt was used, as the prototype was inspired by the implementation of Goossens et al. (2023). For example, the following sentence from Article 15e of the Dutch copyright law illustrates a decision dependency used to model a DRD: “*Disputes concerning **Article 15c, Section 1**, referred **compensation** will be decided in the **first instance** by **exclusion** by the **District Court of The Hague**.*” Here, both **Article 15c, Section 1**, **compensation first instance** and **District Court of the Hague** are input to the **exclusion** decision. After this was given as input to the prototype, the DRD in Figure 5 was given as output. Besides the fact that the input data on the right-hand side should be modeled as a knowledge source, the DRD aligns well with the excerpt of the Dutch legal text.

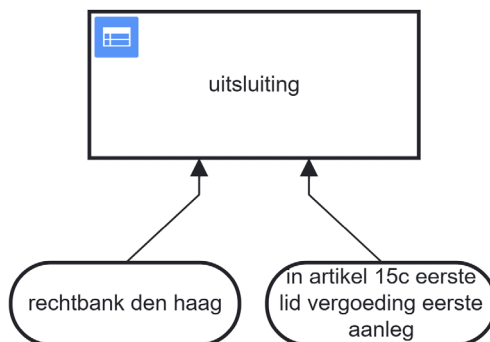


Figure 5: Results of using unstructured legal texts.

However, one apparent issue is that it is unclear which article the excerpt is from and how it refers to Article 15c, due to the structure of legal texts. In addition, as noted by Arco et al. (2021) and Etikala (2021), natural language challenges include bullet lists, a challenge also present in Dutch legal texts (De Maat & Winkels, 2010). Furthermore, Dutch legal texts contain long, complex sentences (Bakker, De Boer, et al., 2022; Sunkle et al., 2020) and, as shown in the DRD in Figure 5, include references to other legal sources (De Maat & Winkels, 2010). Therefore, the current prototype based on Goossens et al. (2023) needs to be extended to consider the structural complexity of Dutch legal texts. This results in the following additional requirements: R1) Support references to other legal sources, R2) Split text (e.g., lists)

into individual sentences, R3) Handle special cases involving long and complex sentences, R4) Support the full DMN syntax. These additional requirements led to a theoretical extension, shown in Figure 5, which extends the prototype implemented based on Goossens et al. (2023). In this regard, references (R1) and lists are now accounted for (R2). Two additional features are the need to handle long, complex sentences (R3), which could be addressed by labeling a dataset containing Dutch legal text, and the ability to model the full DMN syntax within the prototype, e.g., knowledge sources and decision tables per decision (R4).

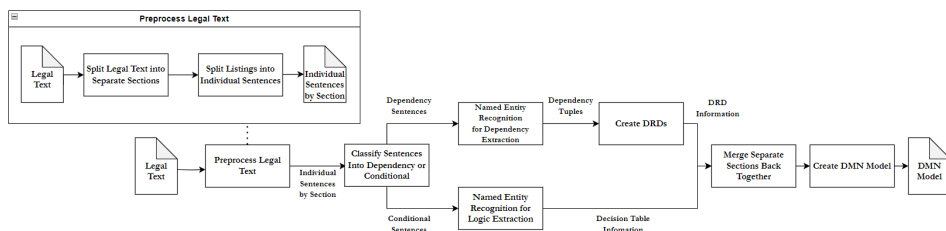


Figure 5: Proposed prototype extension.

7 Discussion and Future Research

This research has several limitations:

- *Limited data collection:* The data may have overlooked techniques such as context-aware deep learning (Nascimento et al., 2018). In addition, by focusing on natural language text, other sources, such as policy documents (Lopez et al., 2022), were excluded. However, these may also be relevant for future research to explore whether they can be utilized to automatically extract Dutch legal text.
- *Prototype limitations:* The prototype does not replace human judgment due to the inherent uncertainty of deep learning models and is not meant for direct implementation. Legal stakeholders should collaborate with NLP experts to understand the prototype’s capabilities and interpret results appropriately.
- *BERTje results:* The high results (above 0.8) were influenced by ChatGPT-generated datasets and similar training and test data. This raises concerns such as model collapse, i.e., reduced diversity and increased repetitiveness in generated sentences. However, ChatGPT was chosen for its efficiency in

quickly generating sentences, with a focus on demonstrating the prototype's functionality. In addition, the pre-trained BERTje model already has a general understanding of the Dutch language.

- *No Dutch legal texts:* The absence of labeled Dutch legal texts, partly due to the authors' technological background, and the use of synthetic data in the first iteration of the DSR process highlights the need for future labeled datasets. These datasets are also needed for fine-tuning three corresponding BERT models trained on legal texts (e.g., RechtBERT (Looijenga, 2024)) in subsequent DSR iterations.
- *Manual parameter selection:* The parameters for fine-tuning were manually selected and may not be optimal. A grid search should be conducted when fine-tuning the three BERTje models for Dutch legal texts.
- *Prototype visualization:* The current prototype visualizes only part of the DMN. Future research should incorporate multiple decision tables using context-aware models to capture sentence context, as well as additional elements such as knowledge sources.

Furthermore, future research should assess whether the proposed prototype extension can handle the structural complexity of Dutch legal texts through implementation, and validation by legal professionals. Therefore, future research will involve the next iteration, refining the prototype, and evaluating it with legal analysts to have a specific legislation prototype that meets the proposed requirements.

8 Conclusion

To implement legislation, legal texts are translated into business rules, but the manual process is labor-intensive, time-consuming, and error-prone. An NLP-based approach can help address these challenges. Therefore, the following main question is addressed: “*In what way can a DMN model be automatically discovered from Dutch legal texts using NLP?*” From a theoretical viewpoint, this research: (1) contributes to the body of knowledge in BRM by providing a foundation for the automatic extraction of a DMN model from Dutch legal texts, adding to the *elicitation*, *design*, and *specification* capabilities. From a practical viewpoint, this research: (2) addresses several challenges in translating Dutch legal texts into executable business rules, such as reducing time and minimizing errors; (3) demonstrates how NLP can be applied

to analyze legislation; (4) makes complex Dutch legal texts more comprehensible to more people than a select few in society by modeling them using a more easy-to-understand DMN model; and (5) facilitates the implementation of business rules due to the executable nature of DMN in, e.g., IT systems, making it easier for rule and law analysts to implement them.

References

- Amnesty International. (2021, October 25). Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms. *Amnesty International*.
<https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/>
- Anonymous. (2025). *DMN Extraction Source Code and Pseudocode*.
https://osf.io/7jpyg/?view_only=618871b7ed3142b5b94ecf65287b21ae
- Arco, L., Nápoles, G., Vanhoenshoven, F., Lara, A. L., Casas, G., & Vanhoof, K. (2021). Natural language techniques supporting decision modelers. *Data Mining and Knowledge Discovery*, 35(1), 290–320. <https://doi.org/10.1007/s10618-020-00718-4>
- Bajec, M., Vavpotič, D., & Krisper, M. (2004). Tracking Business Rules from Their Source to Their Implementation to Support is Maintenance. In *Constructing the Infrastructure for the Knowledge Economy* (pp. 317–330). Springer US. https://doi.org/10.1007/978-1-4757-4852-9_23
- Bakker, R. M., de Boer, M. H. T., van Drie, R. A. N., & Vos, D. (2022). Extracting Structured Knowledge from Dutch Legal Texts: A Rule-based Approach. *EKA'W'22: Companion Proceedings of the 23rd International Conference on Knowledge Engineering and Knowledge Management*.
- Bakker, R. M., De Boer, M. H. T., Van Drie, R. A. N., & Vos, D. (2022). Extracting Structured Knowledge from Dutch Legal Texts: A Rule-based Approach. *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 448–457. <http://ceur-ws.org>
- Barale, C., Klaisoongnoen, M., Minervini, P., Rovatsos, M., & Bhuta, N. (2023). AsyLex: A Dataset for Legal Language Processing of Refugee Claims. *Proceedings of the Natural Legal Language Processing Workshop (NLLP 23)*.
- Bartolini, R., Lenci, A., Montemagni, S., Pirrelli, V., & Soria, C. (2004). Semantic Mark-up of Italian Legal Texts Through NLP-based Techniques. *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*.
http://www.itl.nist.gov/iaui/894.02/related_projects/muc.
- Biagioli, C., Francesconi, E., Passerini, A., Montemagni, S., & Soria, C. (2005). Automatic semantics extraction in law documents. *Proceedings of the 10th International Conference on Artificial Intelligence and Law*, 133–140. <https://doi.org/10.1145/1165485.1165506>
- Boella, G., Di Caro, L., & Robaldo, L. (2013). Semantic Relation Extraction from Legislative Text Using Generalized Syntactic Dependencies and Support Vector Machines. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8035, pp. 218–225). Springer Verlag. https://doi.org/10.1007/978-3-642-39617-5_20
- Boyer, J., & Mili, H. (2011). Agile Business Rule Development. In *Agile Business Rule Development* (pp. 49–71). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-19041-4_3
- De Maat, E., & Winkels, R. (2010). Suggesting Model Fragments for Sentences in Dutch Law. *Proceedings of the 4th Workshop on Legal Ontologies and Artificial Intelligence Techniques*.
<https://www.researchgate.net/publication/229476044>
- De Vries, W., van Cranenburgh, A., Bisazza, A., Caselli, T., van Noord, G., & Nissim, M. (2019). BERTje: A Dutch BERT Model. <http://arxiv.org/abs/1912.09582>

- Dragoni, M., Villata, S., Rizzi, W., & Governatori, G. (2018). Combining Natural Language Processing Approaches for Rule Extraction from Legal Documents. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10791, pp. 287–300). Springer Verlag. https://doi.org/10.1007/978-3-030-00178-0_19
- Etikala, V. (2021). Extracting Decision Model Components from Natural Language Text for Automated Business Decision Modelling. *RuleML+ RR (Supplement)*. https://link.springer.com/10.1007/978-3-030-66498-5_27
- Etikala, V., Van Veldhoven, Z., & Vanthienen, J. (2020). *Text2Dec: Extracting Decision Dependencies from Natural Language Text for Automated DMN Decision Modelling* (pp. 367–379). https://doi.org/10.1007/978-3-030-66498-5_27
- Goossens, A., Claessens, M., Parthoens, C., & Vanthienen, J. (2021). *Deep Learning for the Identification of Decision Modelling Components from Text* (pp. 158–171). https://doi.org/10.1007/978-3-030-91167-6_11
- Goossens, A., Claessens, M., Parthoens, C., & Vanthienen, J. (2022). *Extracting Decision Dependencies and Decision Logic from Text Using Deep Learning Techniques* (pp. 349–361). https://doi.org/10.1007/978-3-030-94343-1_27
- Goossens, A., De Smedt, J., & Vanthienen, J. (2023). Extracting Decision Model and Notation models from text using deep learning techniques. *Expert Systems with Applications*, 211, 118667. <https://doi.org/10.1016/j.eswa.2022.118667>
- Graham, I. (2007). *Business Rules Management and Service Oriented Architecture: A Pattern Language* (1st ed.). Wiley.
- Guisé, A., Lévy, F., & Nazarenko, A. (2012). *From Regulatory Texts to BRMS: How to Guide the Acquisition of Business Rules?* (pp. 77–91). https://doi.org/10.1007/978-3-642-32689-9_7
- Gusenbauer, M. (2019). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, 118(1), 177–214. <https://doi.org/10.1007/s11192-018-2958-5>
- Hasić, F., de Smedt, J., & Vanthienen, J. (2017). Towards Assessing the Theoretical Complexity of the Decision Model and Notation (DMN). *Business Process Modeling, Development, and Support*. <https://www.researchgate.net/publication/322364467>
- Hay, D., & Healy, K. A. (2000). *Defining Business Rules - What Are They Really?* (K. Healy, Ed.).
- Hendrycks, D., Burns, C., Chen, A., & Ball, S. (2021). CUAD: An Expert-Annotated NLP Dataset for Legal Contract Review. *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly: Management Information Systems*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
- Kacfar Emani, C. (2014). *Automatic Detection and Semantic Formalisation of Business Rules* (pp. 834–844). https://doi.org/10.1007/978-3-319-07443-6_57
- Klein Haarhuis, C. M., & Niemeijer, E. (2008). *Wet en werkelijkheid*.
- Leewis, S., Smit, K., & Zoet, M. (2020). Putting Decision Mining into Context: A Literature Study. In *Lecture Notes in Information Systems and Organisation* (Vol. 38, pp. 31–46). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-47355-6_3
- Looijenga, M. S. (2024). *RechtBERT: Training a Dutch Legal BERT Model to Enhance LegalTech*. <http://essay.utwente.nl/104811/>
- Lopez, V., Picco, G., Vejsbjerg, I., Hoang, T. L., Hou, Y., Sbodio, M. L., Segrave-Daly, J., Moga, D., Swords, S., Wei, M., & Carroll, E. (2022). *Envisioning a Human-AI collaborative system to transform policies into decision models*.
- Martin, L., Muller, B., Ortiz Suárez, P. J., Dupont, Y., Romary, L., de la Clergerie, É., Seddah, D., & Sagot, B. (2020). CamemBERT: a Tasty French Language Model. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7203–7219. <https://doi.org/10.18653/v1/2020.acl-main.645>

- Michel, M., Djurica, D., & Mendling, J. (2022, January 4). Identification of Decision Rules from Legislative Documents Using Machine Learning and Natural Language Processing. *Proceedings of the 55th Hawaii International Conference on System Sciences*.
<https://doi.org/10.24251/HICSS.2022.757>
- Mokashi, D., Vidyapeeth, B., Mahadik, P., Ajit, P., Patil, R., Pramod, P., Rahate, G., Prashant Mahadik, P., Neeraj, P., Gangurde, A., Jadhav, A. N., Rajendra, D., & Mohite, B. (2024). Legal Rules and Regulations Document Summarizer: Regulatory Compliance with NLP, ML, and LLMs. *Library Progress International*, 44(3). www.bpasjournals.com
- Morgan, T. (2002). *Business Rules and Information Systems: Aligning IT with Business Goals* (1st ed.). Addison-Wesley Educational Publishers Inc.
- Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Linguisticae Investigationes*, 30(1), 3–26. <https://doi.org/10.1075/li.30.1.03nad>
- 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>
- Nascimento, N., Alencar, P., Lucena, C., Cowan, D. D., Cheriton, D. R., & Cowan, D. (2018). A Context-Aware Machine Learning-based Approach. *Conference of the Centre for Advanced Studies on Collaborative Research*. <https://www.researchgate.net/publication/328874557>
- Nederlandse Omroep Stichting. (2019). *Broers krijgen celstraffen voor miljoenenfraude met pgb's*.
<https://nos.nl/artikel/2312387-broers-krijgen-celstraffen-voor-miljoenenfraude-met-pgb-s>
- Nederlandse Omroep Stichting. (2024, October 14). *Computersystemen tot 2029 niet klaar voor nieuwe belastingplannen*. <https://nos.nl/artikel/2540774-computersystemen-tot-2029-niet-klaar-voor-nieuwe-belastingplannen>
- Nelson, M. L., Rariden, R. L., & Sen, R. (2008). A Lifecycle Approach towards Business Rules Management. *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, 113–113. <https://doi.org/10.1109/HICSS.2008.25>
- Ng, V., & Cardie, C. (2001). Improving machine learning approaches to coreference resolution. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*, 104.
<https://doi.org/10.3115/1073083.1073102>
- Object Management Group. (2019a). *Decision Model and Notation (DMN) 1.3*.
<http://www.omg.org/spec/DMN/1.0/Beta1/>
- Object Management Group. (2019b). *Semantics of Business Vocabulary and Business Rules 1.5*.
- Palmirani, M., Governatori, G., Rotolo, A., Tabet, S., Boley, H., & Paschke, A. (2011). LegalRuleML: XML-Based Rules and Norms. In *LNC3* (Vol. 7018, pp. 298–312).
https://doi.org/10.1007/978-3-642-24908-2_30
- 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>
- Quishpi, L., Carmona, J., & Padró, L. (2021). *Extracting Decision Models from Textual Descriptions of Processes* (pp. 85–102). https://doi.org/10.1007/978-3-030-85469-0_8
- Ross, R. G. (2009). *Basic RuleSpeak Guidelines*.
- Schlosser, S., Baghi, E., Otto, B., & Oesterle, H. (2014). Toward a Functional Reference Model for Business Rules Management. *2014 47th Hawaii International Conference on System Sciences*, 3837–3846. <https://doi.org/10.1109/HICSS.2014.476>
- Sleimi, A., Ceci, M., Sannier, N., Sabetzadeh, M., Briand, L., & Dann, J. (2019). A Query System for Extracting Requirements-Related Information from Legal Texts. *2019 IEEE 27th International Requirements Engineering Conference (RE)*, 319–329. <https://doi.org/10.1109/RE.2019.00041>
- Sleimi, A., Sannier, N., Sabetzadeh, M., Briand, L., & Dann, J. (2018). Automated Extraction of Semantic Legal Metadata using Natural Language Processing. *2018 IEEE 26th International Requirements Engineering Conference (RE)*, 124–135. <https://doi.org/10.1109/RE.2018.00022>
- Smit, K. (2018). *Organization and governance of business rules management capabilities* [Doctoral Thesis]. Open Universiteit.

- Spinosa, P., Giardiello, G., Cherubini, M., Marchi, S., Venturi, G., & Montemagni, S. (2009). NLP-based metadata extraction for legal text consolidation. *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, 40–49. <https://doi.org/10.1145/1568234.1568240>
- Sunkle, S., Saxena, K., & Kulkarni, V. (2020). *Conceptual Model-driven Legal Insights for Stakeholder Decision Making*.
- Vayadande, K., Bhat, A., Bachhav, P., Bhoyar, A., Charoliya, Z., & Chavan, A. (2024). AI-Powered Legal Documentation Assistant. *2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN)*, 84–91. <https://doi.org/10.1109/ICPCSN62568.2024.00022>
- Von Halle, B., & Goldberg, L. (2009). *The Decision Model: A Business Logic Framework Linking Business and Technology* (1st ed.). Auerbach Publications.
- Wyner, A., & Peters, W. (2011). On Rule Extraction from Regulations. *Frontiers in Artificial Intelligence and Applications*, 235, 113–122. <https://doi.org/10.3233/978-1-60750-981-3-113>

BETTER TOGETHER: CONDITIONS FOR IMPLEMENTING INNOVATIONS TO IMPROVE QUALITY OF CARE IN HEALTH DELIVERY ORGANIZATIONS

PIETER KIEVIT, ROGER W.H. BONDS, NADINE ROIJAKKERS

Open Universiteit, Heerlen, the Netherlands

pieter.kievit@ou.nl, roger.bonds@ou.nl, nadine.roijackers@ou.nl

The necessary resilience of healthcare delivery organizations to cope with a changing and ever challenging socio-economic environment depends on continuous, effective innovation, based on the adoption of proven interventions. It is aimed at increasing output quality of care, thus better responding to demands of patients, staff and stakeholders. Although expectations of (digital) innovation in healthcare are generally high, the outcomes often do not meet these expectations. In a scoping review of 46 research papers on innovation, we analyzed which factors were reported as barriers and facilitators for success or failure of innovation initiatives. The majority of papers reported a variety of conditions and a limited use of available implementation frameworks. Furthermore, the underlying model of quality of care was often incomplete, thereby compromising the outcome of the initiative, hampering effective dissemination and implementation of interventions and the means for valid outcome research of innovation projects.

DOI

<https://doi.org/10.18690/um.fov.4.2025.36>

ISBN

978-961-286-998-4

Keywords:

innovation,
implementation,
healthcare quality,
health delivery
organizations,
patients



University of Maribor Press

1 Introduction

Healthcare delivery organizations are under an increasing amount of pressure to keep delivering the necessary quality of care in an environment, where the demand for healthcare services continues to grow, while the availability of staff to deliver them continues to decrease. Continuous innovation is a crucial condition for the resilience of organizations, proving vitality in such a changing and ever challenging socio-economic environment, complying with regulation and imposed standards and responding to varying demands of clients, staff and external stakeholders (Garrido-Moreno et al., 2024).

The authoritative Oslo Manual (OECD, 2018, p. 46) defines innovation as: “The implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations”. For the healthcare sector, a teleological dimension is added where innovation is “A new or improved solution with the transformative ability to accelerate positive health impact” (WHO, 2024, p. 2). Digital innovations in particular “are poised to significantly alter the healthcare environment offering a more sustainable, efficient and accessible healthcare ecosystem for future generations” (Thacharodi et al., 2024, p. 1).

Innovation in healthcare attains this positive health impact through proven interventions. They aim to improve quality of care, which can be defined by six interacting dimensions: effectiveness, efficiency, safety, equity, timeliness and patient-centeredness (Ayanian & Markel, 2016; Berwick & Fox, 2016; Busse et al., 2019). The World Healthcare Organization (WHO) and national healthcare agencies have embraced these dimensions in practice.

However, the outcomes of such innovation initiatives often fail to meet expectations (Hügler & Grek, 2023; Sony et al., 2023). Although innovation is vital for organizational continuity and healthcare improvement, dissemination and implementation (D&I) of even the most proven interventions often remains a challenge with uncertain outcome (Nilsen & Birken, 2020; Peden et al., 2019; Philp & Pitt, 2019). An obvious tension exists between the need to improve quality on the one hand and knowing how to do so on the other – a gap which cannot be closed by applying standard stage-gate models (Auerbach et al., 2007; Cooper, 1990).

Our research interest is to identify the conditions for successful dissemination and -implementation of interventions aimed at innovations that lead to quality improvement. We pose our research question as follows: Which factors are reported as determinants for implementation effectiveness and its outcome in terms of improved quality of care?

This study aims to determine, to what extent frameworks are in fact used in implementation projects for healthcare innovation. Given our aim to determine the scope of coverage of a body of literature on this topic, we will use a scoping review design in this research (Munn et al., 2018). Our study follows a five steps methodological framework. In the next section we explore the theoretical background and formulate a set of expectations (Step 1). Section 3 will discuss how studies were identified (Step 2) and selected (Step 3). Section 4 presents the mapping and interpretation of the data (Step 4). We conclude with the final step (5) of summarizing and reporting the results in Section 5.

2 Theoretical Background

In this section we explore the theoretical background on implementation frameworks and on quality in healthcare delivery organizations, leading to expectations on the role they play in innovation dissemination and implementation.

2.1 Implementation frameworks

To facilitate and support quality improvement through proven interventions, implementation science has developed several frameworks for innovation management and implementation research (Estabrooks et al., 2006; Kitson et al., 2008; Rycroft-Malone & Bucknall, 2010). Frameworks list and classify factors which influence the implementation process and its outcome in several domains to design better innovation projects, reduce the risk of unforeseen side-effects and manage uncertainty (Fagerberg et al., 2005). In total a few dozen frameworks are available, 18 of which were combined by Damschröder et al. into the Consolidated Framework for Implementation Research or CFIR, revised in 2022 based on users' feedback (2009; 2022), leading to one of the most highly cited frameworks (Skolarus et al., 2017).

Considering the need for planning, management and accountability of the necessary investments in quality improvement, it could be expected that these frameworks are widely used as crucial tools in project planning and management. The first focus of this paper's review will be to confirm whether the available resources and experience provided by implementation science are in fact applied and in what way.

2.2 Aspects of quality in healthcare delivery organizations

We define "quality in healthcare" as a healthcare delivery organization's performance on one or more of the six dimensions of quality mentioned in Section 1. However, there remain several complications to consider when addressing quality from an innovation perspective.

Firstly, in any model of quality, including the six dimensions model, there is the challenge of reckoning the interaction between dimensions when implementing an intervention in structure, process or outcome (Donabedian, 1980). There is not just the desired effect on the primary dimension to consider, but also the related intended or unintended interacting side effects on other dimensions (Busse et al., 2019).

Then there is the challenge of operationalizing and measuring the quality of care and the effects of the organizations' activities to improve it. Quality is whatever is measured in the interdependent dimensions at any moment. Crucial in this respect are the quality indicators which can be defined as quantitative measures providing information about the status of the six quality dimensions.

Finally, the problem of interacting effects between dimensions can be addressed by developing methods for combining different indicators into composite indicators or scores (Shwartz et al., 2015). These scores allow aggregation of different aspects into one outcome to give a clearer picture of the overall quality of care and changes therein.

Given the utmost importance of quality improvement as primary motive and success determinant for innovation, the expectation is that innovation initiatives would account for the intended primary effects on the six dimensions of quality, consider the interdependency effects and provide concrete outcome measurements. This will be the second focus of our review.

3 Identification and Selection of Sources

This section presents the process that was followed to identify and select the articles to be included for the review. Munn (2018) recommends the use of a standardized protocol to report on this process. We have summarized our approach according to the scoping review extension of the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement (Page et al., 2021).

3.1 Identifying studies

Healthcare is not the only sector of society being confronted by the need for innovation or experiencing success and failure in implementing quality improvement. This research into dissemination and implementation practice will include organizations in all areas of society facing similar challenges in deciding on interventions, selecting appropriate strategies and managing the process.

We executed our initial literature search in February 2024 on four databases:

- Web of Science as a database of peer reviewed journals focusing on social sciences and general management including journals dedicated to innovation.
- PubMed/MEDLINE for biomedical literature, life science journals and publications focusing on healthcare management.
- the Wiley Online Library as an additional source of publications from the areas of health, social sciences and the humanities on the subject of innovation and conditions for success or failure of quality improving interventions.
- the Cochrane Database of systematic research. Findings in the Cochrane database are systematic reviews, and they are primarily used as a possible source for relevant case studies, not included from the other databases.

The search strategy on the four databases consisted of a query with combinations of the search terms “innovation” and “enabler(s)” and/or “facilitator(s)” and/or “barrier(s)”. The terms were arranged in a Boolean query as ((Innovation AND (facilitator OR enabler) AND barrier)) over title, key words, abstract and full text of

publications ranging from the beginning of database registration till February 1st, 2024. The resulting list of research papers was then submitted to the set of inclusion/exclusion criteria to come to a final selection.

3.2 Select: eligibility criteria and study selection

As the literature research is focused on the conditions for success or failure of actual D&I projects and their outcomes, three inclusion criteria were applied. Studies must have appeared in international peer-reviewed journals in English (1). Case studies, reviews or quality evaluation studies, reporting on D&I projects are included, excluding theoretical and positioning papers (2). Innovations in clinical care or medical interventions in diagnostics, treatment and rehabilitation are excluded as they represent a different type of quality (3).

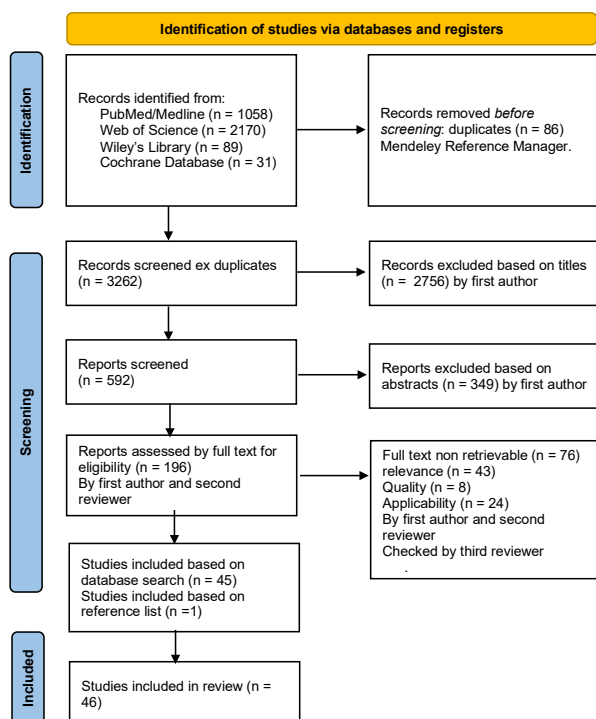


Figure 1: PRISMA Flow Diagram of the selection process

Source: adapted from Page et al., 2021

The retrieval of studies yielded a total of 3348 results. Figure 1 shows how the inclusion criteria reduced the original set to 46 studies for further review. In the end, 11 of the 46 included studies are systematic reviews (reporting on 10-110 primary studies), 9 were scoping reviews (16-81 primary studies) and 26 studies reported the outcome of qualitative (evaluative) case study research into innovation projects and programs. 36 out of 46 studies pertained to innovation in health care, the other 10 addressed conditions for innovation in other areas of society.

4 Results

We coded the 46 review papers for type of implementation frameworks applied, including the type of barriers and facilitators reported, and for quality enhancing interventions and measurements of quality dimensions. The open coding was executed independently by three researchers (the first author and two external reviewers) and combined into detailed tables listing the results per review paper. The remainder of this section presents the results in both areas.

4.1 Application of implementation frameworks

The review papers apply or infer implementation frameworks to identify, summarize and classify the barriers and facilitators. 30 of the 46 papers apply a validated general framework to subsume the barriers and facilitators reported in their primary studies or resulting from data collecting. Seven review papers applied CFIR, three papers used the PARiHS and 20 papers applied a different but validated framework, suited to either the intervention or the implementation environment. Eleven papers developed a specific framework from the primary research they reported on. Five papers analyzed the underlying projects by interpreting the outcome of interviews or focus groups with focus groups by means of discourse analysis or semantic classification, without developing a framework.

This means that most of the authors use or create a wide variety of frameworks, even the ones that do focus on healthcare innovations. Unclear is, whether this is because these authors are following other research paradigms or whether this means that there are relevant categories of barriers and facilitators not yet captured in the most used frameworks in healthcare innovation. However, looking at the dominant domains of the frameworks used or developed in the 41 papers, it seems to be

possible to semantically map the domains on the five established domains from the CFIR framework (characteristics of the intervention, internal context, external context, adopter skills and quality of the implementation project). The exception to this is the focus on end-users, patient outcome or related aspects, which is addressed in several review papers but, notably, not in CFIR or PARIHS.

Regarding the use of frameworks in the underlying primary studies, eleven of the 46 papers comment explicitly on the application of frameworks in the primary studies and cases, reporting only limited deployment. For instance, Allen et al. (2017, p. 1) state that “more than half (55%) of the studies did not articulate an organizational theory or conceptual framework guiding the inquiry”. Kelm et al. (2014, p. 1) go even further and conclude that “the 64 included studies were characterized by relatively poor research designs (and) insufficient reporting of intervention procedures”. Yeboah concludes that in only 38 papers out of 110 included in his review, some form of management theory was applied (Yeboah, 2023).

Table 1: Excerpt from data coding table – implementation framework

Source: <https://github.com/rogerbons/Bled2025Kievit-Bons-Roijakkers>

Reference	Framework Applied	Dominant Domains	Condition: Barriers	Condition: Facilitators
Abu-Odah (2022)	Innovative Care for Chronic Conditions (ICCC) model (WHO 2006)	Five domains: medical condition, internal context, external context, adopters and end-users.	Micro level: insufficient skills, lack of motivation, mistrust of value of research, inability to interpret findings. Meso level: poor dissemination, lack of resources, lack of time and staff, miscommunication between academy and health, lack of access to research. Macro level: policy makers' distrust of research value, lack of training and skills, gap between research and policy considerations.	Micro level: motivation of professionals. Meso level: dissemination findings, capacity building workshops, budget for research activities. Macro level: collaboration and connection, identifying right stakeholders, developing trust between policy makers and researchers.

This limited use of frameworks in the primary studies does not imply that these organizations do not recognize barriers and facilitators for their implementations. The 46 review papers and quality studies classify large numbers of barriers and facilitators from the primary studies, spanning a substantial number of domains and influencing the outcome of the quality improving initiatives they report. For instance, Abell et al. find 227 individual barriers and 130 individual facilitators (2023), while Consolo et al. (2023, p. 1) reports that “... a total of 245 codes or factors was generated”.

Figure 2 contains an excerpt of the findings, the full list of 46 entries has been made available in a github repository. Summarizing, we find that less than half the projects and programs in the primary studies make use of a validated and well-established framework to plan, manage and evaluate the innovation initiative. Furthermore, if a framework is used, in most cases only a limited number of domains and constructs are applied. In that way, interaction and mutual influence go unnoticed. Also, barriers and facilitators often do not match the factors in which the domains of a framework are operationalized. In these cases, authors of review papers design their own dedicated framework with factors to match the reported barriers and facilitators. In many of those cases, it seems it would have been possible to conform to CFIR, with the exception of the patient-oriented domain.

4.2 Aspects of quality

Each of the 46 papers included in the selection addresses conditions for dissemination and implementation of interventions aimed at quality improvement. However, the concepts of quality they used vary. Of the 36 papers dealing with innovation in healthcare specifically, 20 apply to dimensions of quality that fit within the six dimensions model mentioned in Section 2, the other 16 do not refer to any specific quality dimension, but to quality improvement in general. Eight of these 20 concentrate on initiatives to improve patient centeredness. Six papers report on efficiency. Four report on initiatives to increase safety and equity is main concern of two papers. Interestingly, the dimensions of effectiveness and timeliness of healthcare do not appear as leading in any of the 36 papers with a healthcare focus. Figure 3 shows an excerpt of the quality dimensions in our coding table.

Table 2: Excerpt from data coding table – quality dimensions.Source: <https://github.com/rogerbons/Bled2025Kievit-Bons-Roijackers>

Reference	Intervention	Quality Dimension
Abell (2023)	Implementation and utilization of Computerized Clinical Decision Support Systems (CDSS)	CDSS have the potential to improve various aspects of healthcare delivery, such as patient safety, clinical management, diagnostic support, cost management and administrative efficiency

Although all papers report on initiatives to improve quality of care through interventions on one (or more) of the quality dimensions, none of the reported cases used quantifiable outcome measures. The reported effects of implementing quality improving interventions are thus of a qualitative nature which leaves them open to debates about relevance, applicability and validity. A few examples across the set of included papers may illustrate this (Table 3).

Table 3: Examples of outcome measures on dimensions of quality

Patient Centeredness	“low readmission rates and better patient satisfaction” (Fakha et al., 2021, p. 19)
	“patient autonomy, patient participation” (Vogel et al., 2023, p. 2)
	“patient benefit” (Baines et al., 2020, p. 7)
Efficiency	“more value from resources allocation” (Ahumada-Canale et al., 2023, p. 1)
	“administrative efficiency, cost effectiveness, or users experience” (Barnett et al., 2011, p. 2)
	“patient health, care processes, use of resources and economic variables” (Niezen & Mathijssen, 2014, p. 152)
Safety	“knowledge of and adherence to routines as fundamental to maintain patient safety” (Dugstad et al., 2020, p. 10)
	“the quality, safety, patient-centeredness and cost-effectiveness of care” (Laukka et al., 2020, p. 1)
Equity	“health policies for vulnerable populations” (Daniels et al., 2022, p. 1)
	“efficient care” (Ghabour et al., 2023, p. 157)

The only readily quantifiable outcome measure occasionally mentioned in both the reviews and the underlying papers is the clinical patient outcome. This is defined by and results from medical treatment based on the “diagnosis and treatment combination” (DTC). DTCs specify the entirety of all steps required to treat a medical condition or illness, from the first consultation up to the final check-up, and they are standardized and reported on, for instance from healthcare delivery organizations to health insurers. As such, clinical patient outcomes depend on the

relevance and quality of medical interventions and though they reflect the quality of a health delivery system, they are not included in the six quality dimensions.

Another outcome of our investigation is the missing aspect of interacting effects of interventions on more than one quality dimensions. Several reviews conceive of quality as a combination of dimensions. The consequences of interacting effects of interventions on more than one dimension are not acknowledged, although there are sets of composite indicators available to recon and account for these effects, thus providing a comprehensive assessment of quality (Busse et al., 2019).

Although quantifying the effects of interventions on dimensions of quality is complicated and, in case of combined dimensions, requires application of complexity theory (Storkholm et al., 2019; Turner & Baker, 2019), quality improvement is hard to substantiate without such indicators.

5 Discussion and Conclusions

This literature review was motivated by an interest in the conditions for successful implementation of healthcare innovation. Our focus was to assess if frameworks from implementation science are used in (healthcare) practice and to determine to what extent the six standardized quality of healthcare dimensions are used and operationalized. We selected 46 review papers (36 focusing on healthcare) that analyzed a total of over a thousand underlying primary studies.

There is consensus that the primary studies made very limited use of the available frameworks. This could be due to insufficient local expertise or to a lack of access to these instruments. This complicates comparing projects and establishing quality requirements for design and management of initiatives. It hampers development and application of implementation research and development in practice.

Based on the interpretation of findings reported by 46 review papers and qualitative evaluation studies, the conjectures of Moullin et al. are confirmed, that frameworks "...are not used optimally to substantiate or advance implementation science and practice", thereby "...slowing the translation of research evidence into practice and limiting public health impact" (Moullin et al., 2020, p. 2). We suggest to conduct further research on how the knowledge from the implementation science domain

can be better disseminated to the teams that are responsible for the necessary innovations in the healthcare delivery organizations.

Furthermore, we noticed a diversity in frameworks used by the authors of the review papers, with only 10 of them applying the CFIR and PARiHS frameworks. We believe that it would have been possible for the others to map their framework domains to the five commonly used by CFIR, thus improving comparability and further validation of this framework. The one exception was the patient domain, which is not covered by CFIR, but is considered important in several of the review studies. While we obviously recognize patients as important stakeholders in healthcare innovations, we believe further investigation is needed to assess, if this warrants the extension of the CFIR framework with a patient oriented domain.

The standard model of healthcare quality was partially applied. Concentrating on one or two dimensions, ignoring side effects through interaction between dimensions and insufficient use of quantitative outcome measures may lead to compromising the effectiveness of the innovation and the rationale for future investments in the field. The fact that this seems to be ignored poses interesting research questions we hope to address in the future.

This also applies to the lack of quantified operationalizations of quality dimensions. Investing in quality improvement requires a business case where funding, commitment of people and organizational creativity is outweighed by projected return and anticipated improvements. This business case needs a comprehensive and accurate definition of quality and hard outcome measures to weigh necessary resources against expected yields in terms of improved indicators on relevant dimensions, taking into account their interaction towards better outcomes.

These findings also lead us to some practical recommendations to healthcare delivery organizations. Firstly, involving innovation specialists in their projects and programs should increase the quality of the decisional practice in the dissemination and implementation of innovation. We believe this is especially relevant for digital innovations, where the knowledge on technology and its integration into healthcare practices is even more scarce. Furthermore, we recommend that standard frameworks and quality measurements are used to encourage benchmarking by implementation teams. We encourage these findings to be shared with researchers,

so that the implementation science field can benefit and its knowledge on D&I processes can be further validated and grown.

Finally, as with any study there are some limitations to be mentioned. Our findings are limited by the fact that they are based on the evaluation of 46 reviews and quality evaluations and not on the actual field studies. We plan to investigate the application of artificial intelligence based reviewing techniques to the underlying 1000+ studies, but also to set up action research based research projects in which we actively participate in innovation projects. Another possible limitation in the generalizability of our findings lies in the localized focus of healthcare. Even though the standards we suggest to use are internationally recognized, healthcare is typically organized at the national, regional or local level, making it more challenging to transfer knowledge across borders.

Acknowledgements:

The authors wish to express their appreciation for the support of Piet Bartels and Marianne Schoorl, who acted as “Reviewer 2” and “Reviewer 3” in the qualitative content analysis and provided valuable feedback on this paper.

References

- Abell, B., Naicker, S., Rodwell, D., Donovan, T., Tariq, A., Baysari, M., Blythe, R., Parsons, R., & McPhail, S. M. (2023). Identifying barriers and facilitators to successful implementation of computerized clinical decision support systems in hospitals: a NASSS framework-informed scoping review. *Implementation Science*, 18(1), 32. <https://doi.org/10.1186/s13012-023-01287-y>
- Ahumada-Canale, A., Jeet, V., Bilgrami, A., Seil, E., Gu, Y., & Cutler, H. (2023). Barriers and facilitators to implementing priority setting and resource allocation tools in hospital decisions: A systematic review. *Social Science & Medicine*, 322, 115790. <https://doi.org/10.1016/j.socscimed.2023.115790>
- Allen, J. D., Towne, S. D., Maxwell, A. E., DiMartino, L., Leyva, B., Bowen, D. J., Linnan, L., & Weiner, B. J. (2017). Measures of organizational characteristics associated with adoption and/or implementation of innovations: A systematic review. *BMC Health Services Research*, 17(1), 591. <https://doi.org/10.1186/s12913-017-2459-x>
- Auerbach, A. D., Landefeld, C. S., & Shojania, K. G. (2007). The Tension between Needing to Improve Care and Knowing How to Do It. *New England Journal of Medicine*, 357(6), 608–613. <https://doi.org/10.1056/NEJMs070738>
- Ayanian, J. Z., & Markel, H. (2016). Donabedian’s Lasting Framework for Health Care Quality. *New England Journal of Medicine*, 375(3), 205–207. <https://doi.org/10.1056/NEJMp1605101>
- Baines, R., Tredinnick-Rowe, J., Jones, R., & Chatterjee, A. (2020). Barriers and Enablers in Implementing Electronic Consultations in Primary Care: Scoping Review. *Journal of Medical Internet Research*, 22(11), e19375. <https://doi.org/10.2196/19375>
- Barnett, J., Vasileiou, K., Djemil, F., Brooks, L., & Young, T. (2011). Understanding innovators’ experiences of barriers and facilitators in implementation and diffusion of healthcare service

- innovations: a qualitative study. *BMC Health Services Research*, 11(1), 342. <https://doi.org/10.1186/1472-6963-11-342>
- Berwick, D., & Fox, D. M. (2016). "Evaluating the Quality of Medical Care": Donabedian's Classic Article 50 Years Later. *The Milbank Quarterly*, 94(2), 237–241. <https://doi.org/10.1111/1468-0009.12189>
- Busse, R., Klazinga, N., Panteli, D., & Quentin, W. (eds). (2019). Improving healthcare quality in Europe Characteristics, effectiveness and implementation of different strategies. *Health Policy Series*, 53.
- Consolo, L., Colombo, S., Basile, I., Rusconi, D., Campa, T., Caraceni, A., & Lusignani, M. (2023). Barriers and facilitators of electronic patient-reported outcome measures (e-PROMs) for patients in home palliative cancer care: a qualitative study of healthcare professionals' perceptions. *BMC Palliative Care*, 22(1), 111. <https://doi.org/10.1186/s12904-023-01234-0>
- Cooper, R. G. (1990). Stage-gate systems: a new tool for managing new products. *Business Horizons*, 33(3), 44–54.
- Damschroder, L. J., Aron, D. C., Keith, R., Kirsh, S. R., Alexander, J., & Lowery, J. (2009). *Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science*. Implement Sci.
- Damschroder, L. J., Reardon, C. M., Widerquist, M. A. O., & Lowery, J. (2022). The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation Science*, 17(1), 75. <https://doi.org/10.1186/s13012-022-01245-0>
- Daniels, S. I., Cheng, H., Gray, C., Kim, B., Stave, C. D., & Midboe, A. M. (2022). A scoping review of implementation of health-focused interventions in vulnerable populations. *Translational Behavioral Medicine*, 12(9), 935–944. <https://doi.org/10.1093/tbm/ibac025>
- Donabedian, A. (1980). The definition of quality and approaches to its assessment. *Exploration in Quality Assessment and Monitoring*, 1.
- Dugstad, J., Sundling, V., Nilsen, E. R., & Eide, H. (2020). Nursing staff's evaluation of facilitators and barriers during implementation of wireless nurse call systems in residential care facilities. A cross-sectional study. *BMC Health Services Research*, 20(1), 163. <https://doi.org/10.1186/s12913-020-4998-9>
- Estabrooks, C. A., Thompson, D. S., Lovely, J. J. E., & Hofmeyer, A. (2006). A guide to knowledge translation theory. *Journal of Continuing Education in the Health Professions*, 26(1), 25–36. <https://doi.org/10.1002/chp.48>
- Fagerberg, J., Mowery, D. C., & Nelson, R. R. (2005). *The Oxford Handbook of Innovation*, Oxford (Fagerberg J., Mowery D., & Nelson R., Eds.). Oxford University Press.
- Fakha, A., Groenvynck, L., de Boer, B., van Achterberg, T., Hamers, J., & Verbeek, H. (2021). A myriad of factors influencing the implementation of transitional care innovations: a scoping review. *Implementation Science*, 16(1), 21. <https://doi.org/10.1186/s13012-021-01087-2>
- Garrido-Moreno, A., Martín-Rojas, R., & García-Morales, V. J. (2024). The key role of innovation and organizational resilience in improving business performance: A mixed-methods approach. *International Journal of Information Management*, 77, 102777. <https://doi.org/10.1016/j.ijinfomgt.2024.102777>
- Ghabour, M., Wilby, K. J., Morris, C. J., & Smith, A. J. (2023). Overview of factors influencing successful implementation of non-medical prescribing. *Journal of Pharmacy Practice and Research*, 53(4), 155–170. <https://doi.org/10.1002/jppr.1868>
- Hügler, T., & Grek, V. (2023). Digital transformation of an academic hospital department: A case study on strategic planning using the balanced scorecard. *PLOS Digital Health*, 2(11), e0000385. <https://doi.org/10.1371/journal.pdig.0000385>
- Kelm, Z., Womer, J., Walter, J. K., & Feudtner, C. (2014). Interventions to cultivate physician empathy: a systematic review. *BMC Medical Education*, 14(1), 219. <https://doi.org/10.1186/1472-6920-14-219>
- Kitson, A. L., Rycroft-Malone, J., Harvey, G., McCormack, B., Seers, K., & Titchen, A. (2008). Evaluating the successful implementation of evidence into practice using the PARiHS

- framework: theoretical and practical challenges. *Implementation Science*, 3(1), 1. <https://doi.org/10.1186/1748-5908-3-1>
- Laukka, E., Huhtakangas, M., Heponiemi, T., & Kanste, O. (2020). Identifying the Roles of Healthcare Leaders in HIT Implementation: A Scoping Review of the Quantitative and Qualitative Evidence. *International Journal of Environmental Research and Public Health*, 17(8), 2865. <https://doi.org/10.3390/ijerph17082865>
- Moullin, J. C., Dickson, K. S., Stadnick, N. A., Albers, B., Nilsen, P., Broder-Fingert, S., Mukasa, B., & Aarons, G. A. (2020). Ten recommendations for using implementation frameworks in research and practice. *Implementation Science Communications*, 1(1), 42. <https://doi.org/10.1186/s43058-020-00023-7>
- Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(1), 143. <https://doi.org/10.1186/s12874-018-0611-x>
- Niezen, M. G. H., & Mathijssen, J. J. P. (2014). Reframing professional boundaries in healthcare: A systematic review of facilitators and barriers to task reallocation from the domain of medicine to the nursing domain. *Health Policy*, 117(2), 151–169. <https://doi.org/10.1016/j.healthpol.2014.04.016>
- Nilsen, P., & Birken, S. A. (2020). *Handbook on implementation science*. Edward Elgar Publishing.
- OECD. (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation* (T. and I. A. The Measurement of Scientific, Ed.; 4th edition). OECD Publishing. <https://doi.org/10.1787/9789264304604-en>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>
- Peden, C. J., Stephens, T., Martin, G., Kahan, B. C., Thomson, A., Rivett, K., Wells, D., Richardson, G., Kerry, S., Bion, J., & others. (2019). Effectiveness of a national quality improvement programme to improve survival after emergency abdominal surgery (EPOCH): a stepped-wedge cluster-randomised trial. *The Lancet*, 393(10187), 2213–2221.
- Philp, M. M., & Pitt, H. A. (2019). Emergency general surgery: can we do better? *The Lancet*, 393(10187), 2178–2180. [https://doi.org/10.1016/S0140-6736\(18\)32982-9](https://doi.org/10.1016/S0140-6736(18)32982-9)
- Rycroft-Malone, J., & Bucknall, T. (Eds). (2010). *Models and Frameworks for Implementing Evidence-Based Practice: Linking Evidence to Action*. Oxford, John Wiley & Sons.
- Shwartz, M., Restuccia, J. D., & Rosen, A. K. (2015). Composite measures of health care provider performance: a description of approaches. *The Milbank Quarterly*, 93(4), 788–825.
- Skolarus, T. A., Lehmann, T., Tabak, R. G., Harris, J., Lecy, J., & Sales, A. E. (2017). Assessing citation networks for dissemination and implementation research frameworks. *Implementation Science*, 12(1), 97. <https://doi.org/10.1186/s13012-017-0628-2>
- Sony, M., Antony, J., & Tortorella, G. L. (2023). Critical Success Factors for Successful Implementation of Healthcare 4.0: A Literature Review and Future Research Agenda. *International Journal of Environmental Research and Public Health*, 20(5), 4669. <https://doi.org/10.3390/ijerph20054669>
- Storkholm, M. H., Mazzocato, P., & Savage, C. (2019). Make it complicated: a qualitative study utilizing a complexity framework to explain improvement in health care. *BMC Health Services Research*, 19(1), 842. <https://doi.org/10.1186/s12913-019-4705-x>
- Thacharodi, A., Singh, P., Meenatchi, R., Tawfeeq Ahmed, Z. H., Kumar, R. R. S., V, N., Kavish, S., Maqbool, M., & Hassan, S. (2024). Revolutionizing healthcare and medicine: The impact of modern technologies for a healthier future—A comprehensive review. *Health Care Science*, 3(5), 329–349. <https://doi.org/10.1002/hcs2.115>

- Turner, J. R., & Baker, R. M. (2019). Complexity Theory: An Overview with Potential Applications for the Social Sciences. *Systems*, 7(1), 4. <https://doi.org/10.3390/systems7010004>
- Vogel, A., Guinemer, C., & Fürstenau, D. (2023). Patients' and healthcare professionals' perceived facilitators and barriers for shared decision-making for frail and elderly patients in perioperative care: a scoping review. *BMC Health Services Research*, 23(1), 197. <https://doi.org/10.1186/s12913-023-09120-4>
- WHO. (2024). *Innovation for urban health*. <https://www.who.int/westernpacific/initiatives/innovation-for-health-impact>
- Yeboah, A. (2023). Knowledge sharing in organization: A systematic review. *Cogent Business & Management*, 10(1). <https://doi.org/10.1080/23311975.2023.2195027>

AN ANALYSIS OF SCIENTIFIC REQUIREMENTS ON ARTIFICIAL INTELLIGENCE GOVERNANCE

MARIUS GOEBEL,¹ ROGER W. H. BONNS,¹
RÜDIGER BUCHKREMER²

¹ Open University, Department of Information Science, Heerlen, the Netherlands
maris.goebel@ou.nl, roger.bons@ou.nl

² FOM Hochschule für Oekonomie & Management - Institute of IT Management and
Digitization Research, Düsseldorf, Germany
ruediger.buchkremer@fom.de

The emergence of AI technology has prompted the need for standardization and governance due to the potential societal risks associated with its use. However, there is currently no common concept for AI standardization, that considers a broad range of social and ethical subject areas. International cooperation is necessary to address the possible threats, and various nations and organizations have already made initial efforts in this direction. Our overall research question investigates, to what extent requirements based on scientific insights have been addressed in international standards and what new insights standardization efforts can offer to science. In this paper we report the findings of an extensive systematic literature review of 482 scientific articles, using a hybrid analysis process combining manual coding with generative AI supported triangulation steps. The resulting 17 requirements will be used as a basis for a thematic analysis of the most relevant AI standards currently being developed and deployed globally.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.37](https://doi.org/10.18690/um.fov.4.2025.37)

ISBN

978-961-286-998-4

Keywords:

artificial intelligence,
global AI governance,
AI governance components,
AI standardization,
scientific requirements



University of Maribor Press

1 Introduction

The advent of Artificial Intelligence (AI) has been a long and winding road, arguably starting with either Asimov's novel "I, robot" in 1943 or McCulloch and Pitts' scientific article "A Logical Calculus of the Ideas Immanent in Nervous Activity" in 1943, according to (Toosi et al., 2022). They describe the history of the field and in particular the "winters and summers" of AI depression and euphoria respectively and suggest, that we were already in the midst of a summer back in 2021. Since then, Microsoft, Google and Open AI launched several generative AI tools such as Dall-E in 2019 and ChatGPT in 2022, bringing AI to the attention of the general public. Whether or not this means, that we have now moved into a permanent summer, or the prelude of an AI ice-age, is yet to be seen, but it is apparent that unprecedented numbers of organizations, companies, employees and citizens are now looking for ways to make use of AI based tools.

Ethical and societal issues resulting from the adoption of AI, such as bias of training data and its resulting applications, unexplainable decisions by black-box neural networks and lack of human measure in automated decision making, have been discussed in scientific literature for decades. The generative capabilities of large language models are just one of many AI categories, but due to their exposure and availability to the general public, they are at the forefront of the public debate on AI and have raised public awareness and concerns about the ethical use of AI.

One of the consequences of this development and the societal concerns resulting from this, is that governments and other (inter)national bodies are looking to regulate and/or standardize the way AI has to be developed, implemented and used. Perhaps most prominently, the European Commission and the Parliament of the European Union have recently endorsed the EU AI Act, with the aim to "foster trustworthy AI in Europe" (European Commission, 2024). Other national and international bodies have also made initial efforts to counteract possible threats posed by AI standardization and regulation initiatives (Butcher & Beridze, 2019; Schmitt, 2022; Lorenz, 2020). The possible negative and positive impact of AI on society, organizations and individuals has such a magnitude, that it is understandable that these initiatives press on with urgency. But this also means, that unforeseen consequences may arise, due to the combination of still evolving technological capabilities and the discovery of new application domains on the one hand and the

“need for speed” on the other. For instance, the classification of “emotion recognition” as a prohibited application area in the EU AI Act (European Commission, 2025) has the - probably unforeseen and undesirable - consequence, that certain applications in the medical and educational domains can no longer bring their potential benefits to society (Iren, 2025).

The standardization and regulation of AI is a young field in the IT industry (Chen, 2021) and literature describing the landscape of international AI standardization and governance activities is rare. We therefore ask ourselves, to what extent the scientific body of knowledge formulates requirements on AI governance and standardization that can help finding this balance between speed and rigor.

On the other hand, in the development of standards and regulations, it should be common practice to involve all stakeholders in the decision-making process. So even if it would be the case, that these safeguards are not always met when such processes need to be expedited, we feel it is fair to assume that the standards and regulations that are being put in place do contain a significant degree of stakeholder input, leading us to a second research question, to what extent AI standards and regulations offer relevant new contributions to the scientific body of knowledge.

This paper focuses on answering the first of these two research questions. The next section provides a background from literature on AI governance and the context of standardization. Section 3 then discusses the methodological approach that we have followed for the systematic literature review. The results of the review are presented in Section 4. Finally, Section 5 discusses the results and presents the conclusions, including the limitations of this study.

2 Prior research

Butcher and Beridze (2019) give an overview about stakeholder and their kicked-off activities from the private sector, public sector, research, and multi-stakeholder organizations. The authors identify disagreements over how to effectively regulate and implement AI, challenges for legal systems, as well as regulation up-sides with incentivization or risk minimization. They name the European Union’s General Data Protection Regulation as an example of an already implemented AI regulation vehicle. Butcher and Beridze (2019) conclude that AI governance is still an

indeterminate for many states and that first regulation frameworks for specific use cases are needed before a global governance approach can be targeted. The authors designate aviation safety regulation as an inspiring blueprint for global AI governance.

Schmitt (2022) builds on the research of Butcher and Beridze. Schmitt divides the stakeholders in two groups based on whether the stakeholders are trying to regulate AI within existing frameworks or not. The author points out that stakeholders tend to adapt existing frameworks. Schmitt observes that an initial consolidation is taking place and that some stakeholders are increasingly aligning themselves with the Organization for Economic Co-Operation and Development (OECD). The author concludes that international cooperation is lagging legally binding agreements and actors from the public sector are effectively absent from discussions around global AI governance. Further, Schmitt sees a risk that geopolitical considerations as well as different perceptions of human-centric AI values will interfere global standardization.

In its report Lorenz (2020) gives a comprehensive overview on organizations developing AI governance standards. The report gives insights about the category, main topics, goals, and members of each identified organization. Lorenz states global governance is largely defined by political actors and standard developing organizations. Lorenz addresses the need for a mechanism to identify the most powerful international AI governance organization, to help other actors in understanding and participation.

Djeffal, Siewert and Wurster (2022) conducted a study to help understand how states approach to regulate AI. The authors identified four different approaches: Self-regulation-promotion state, market-oriented state, entrepreneurial state, and regulatory state. They identify responsibility as a crosscutting subject that runs through all four approaches.

In our observation there is limited attention in literature for AI Governance and standardization. There also seems to be a lack of literature that attempts to map current activities with a certain level of detail or that matches these with requirements from research and science. This paper's aim is to fill the first of these two gaps.

3 Research Methodology

Our aim is to derive requirements on AI governance and -standardization from the scientific body of knowledge, without putting a priori restrictions on the scientific field nor the application domain. This yields a relatively broad scope, for which we require an inductive yet systematic approach to maximize the transparency of our selection and synthesis process. We have followed the guidelines (Wolfswinkel et al., 2013), who propose a combination of a systematic selection process, combined with methods based on grounded theory. Following their structure, this chapter first discusses the definition of search, then reports about the search and select process, and concludes with a description of the analysis process, the results of which will be presented in Chapter 4.

3.1 Define

As stated before, our goal is to maintain a broad scope throughout the literature review. This requires us to deviate slightly from the guidelines offered by Wolfswinkel et al. to “efficiently perform a systematic literature review”, which are aimed at narrowing both the scope of the review as well as the “fields” of research. Instead, we intend to keep our search broad, and while being precise in the definition of our search criteria and sources and the documentation of our findings, we used an alternative approach to keeping the resulting analytical work manageable, both in terms of data retrieval and the subsequent qualitative content analysis (which we will discuss in Section 3.4).

The first element for this approach is that we have chosen is to not limit the selection and thus work with a relatively large set of articles, coming from different fields and disciplines. The size of the dataset allows us to focus on the most important requirement that the respective authors present, which is typically found in the abstract of the paper. This eliminates the need to obtain and analyze the full papers, making the retrieval of the data trivial and the execution of the study feasible. From other research with large datasets, we have seen that the actual impact of the limitation to abstracts on the validity of the results only has a relatively small impact, if the number of papers analyzed is sufficiently large (Buchkremer et al., 2019).

After reviewing the prior literature research documents, a keyword design was first developed based on the research question. The topic term “artificial intelligence” was set as the initial keyword. Based on the objective of the work, terms from the field of AI standardization were selected as sub-keywords (“standards”, “governance”, “principles”), which were identified during the preliminary literature research combined with forward and backward search. Further keywords during this step were used as the respective sub-categories (e. g. “policies”, “ethics” or frameworks”). From the keyword design developed, corresponding search queries for the database search were designed, considering the operators “AND” and “OR” (etc.). Preliminary searches were carried out via search tests and based on the available search options and accessibility of the publications using the following databases: IEEE Xplorer, Bielefeld Academic Search Engine, Emerald Insights and Web of Sciences. A corresponding search string was then compiled based on these test search queries (Figure 1).

```
AB=(("artificial intelligence" OR "AI" ) NEAR/15 ("govern*" OR "standard*") AND ("principle*" OR "polic*" OR "ethic*" OR "framework*" OR "responsibilit*" OR "evaluation*" OR "certification*" OR "auditabilit*")))) AND TI=("artificial intelligence")
```

Figure 1: Search query used in the Web of Science database

Source: Own research

We selected Web of Science (WoS) as the data source for our search, given that it is well-maintained and offers easy exporting of the raw abstract texts. (Falagas et al., 2008; Martín-Martín et al., 2018) In addition, only peer-reviewed articles that were not older than 2012 were considered.

3.2 Search and Select

The search string triggered 482 hits, which were stored in a large Microsoft Excel table for further processing. As stated before, we did not make any further reductions on this set. We made a manual analysis of the application areas for each of the articles, resulting in overview of Figure 2. The full list of articles and their labeling into application areas is available through an online appendix¹.

¹ <https://github.com/rogerbons/Bled2025Goebel-Bons-Buchkremer>

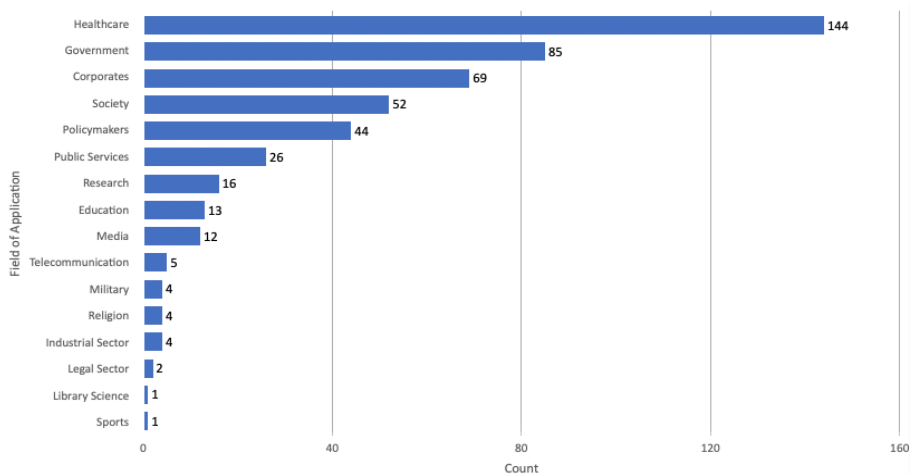


Figure 2: Distribution of the articles across application areas
Source: own research

3.3 Analyze

In the analysis stage of the process, we have made the second change to the guidelines of Wolfswinkel et al., in the way the grounded theory was applied. The inductive open coding of text is inherently subjective, when performed by a single researcher, where conscious and unconscious bias might influence the result. This is typically solved by having multiple researchers (albeit often from the same research team) execute the various steps and document the findings in a transparent way. Even after having reduced the scope of the analysis to the abstract of papers only, the set of 482 posed challenges in the feasibility of a multi-researcher full analysis.

Therefore, we used a combination of manual coding with various generative AI models to triangulate the potentially subjective findings in each of the coding stages. Our methodology for analyzing abstracts, as opposed to full-text analysis, aligns with contemporary comparative studies on abstract screening that utilize large language models (LLMs), such as those conducted by Li et al. (2024) and Dennstädt et al. (2024). This approach, while efficient, also encompasses its inherent limitations. While generative AI models, particularly the large public ones, most likely will have biases of their own, it is highly unlikely that these biases are *specific* towards the prompts we ask them to execute as part of the research. By disclosing the prompts used, we can give maximum transparency of the process. In the next chapter, we will

first present how we did this for each of the stages and then present the outcome of our analysis. Figure 3 summarizes the methodological steps taken, including the coding stages we will further detail in the next section.

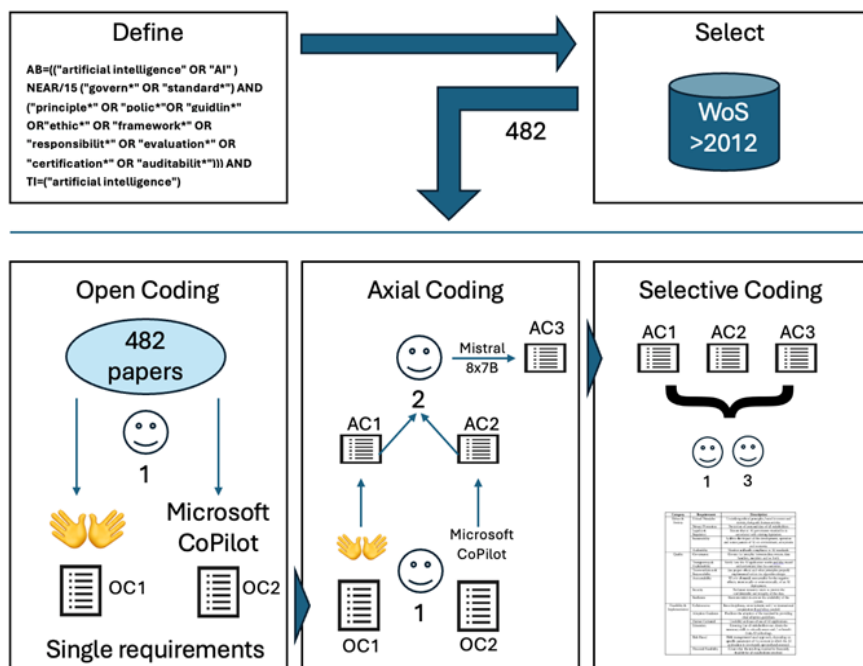


Figure 3: The qualitative content analysis stage of our study

Source: own work, adapted from Wolfswinkel et al., 2013

4 Findings

4.1 Open coding

In the first coding stage, the aim was to identify the core requirement for each of the papers in our selection. On the set of 482 abstracts, we therefore did execute a single manual inductive content analysis, following the classical coding principles defined in (Corbin, 2011), but were executed by a single researcher. This researcher first identified codes in each abstract, clustered the results and finally prioritized the most prominent components encountered. Instead of repeating this manual process with a second researcher, we then collectively deployed a Generative AI model (Microsoft

Co-Pilot version as of June 2024) on the same dataset of 482 abstracts. We prompted the GenAI to only identify a single requirement per abstract and to report back the original terms from the document. The prompts used in the process have been included in the online appendix.

4.2 Axial coding

In this stage, our aim was to reduce and harmonize the long list of requirements identified from the manual and the GenAI based process. We followed a four steps process for this part of the analysis. First, we continued with the dual approach by having a manual reduction of the list we manually created, combining similar terms into single terms wherever possible, while keeping track of the number of underlying abstracts that formed the basis of the categories. In parallel, we again prompted the GenAI we used for the open coding to combine the original set of codes coming from the GenAI open coding version, also included in the online appendix.

For the third step we involved a second researcher, who was provided with the top ranking categories from both resulting lists. He combined the terms and then performed a thematic analysis on the 482 abstracts using a different GenAI model (Mistral version 8x7B), but this time allowing for multiple requirements identified by the model as well as allowing for additional themes to be added to the list. The step confirmed the categorized codes that were specifically looked for and identified a long list of themes that were found in addition.

4.3 Selective coding

In the first step we generated a list of clustered themes present in at least two of the three sets we found in the previous stage, keeping track of how originating concepts of the three sets were collided into the encompassing term. After that, themes that exclusively occurred in only a single set were added to the cluster list for transparency reasons, but without assigning a term to them. To conclude the selective coding, the 19 resulting requirements were critically reviewed by Researchers 1 and 3 on their usefulness. We eliminated the category “Standardization & Guidelines” as a separate requirement, as we consider their appearance in the results primarily as a confirmation that they are indeed needed, but not as a requirement on the process to define them. The resulting list is depicted in Table 1.

Table 1: Requirements on AI Governance Standardization

Category	Requirement	Description
Ethics & Society	Ethical Principles	Underlying ethical principles, based in norms and morals, that guide human activity.
	Privacy Protection	Protection of personal data of all stakeholders.
	Legality & Regulation	Ensure that an AI governance standard is in accordance with existing legislation.
	Sustainability	Address the impact of the development, operation and consequences of AI on ecology, society and economy.
	Auditability	Monitor and audit compliance to AI standards.
Quality	Governance	Govern the interplay between data owners, data handlers, modelers and so forth.
	Transparency & Explainability	Verify how the AI application works. Trace and understand the outcomes.
	Trustworthiness & Responsibility	Are proper ethical and other principles properly implemented within the algorithm design.
	Accountability	Who is ultimately accountable for the negative effects, intentionally or unintentionally, of an AI deployment.
	Security	Technical measures taken to protect the confidentiality and integrity of the data.
	Resilience	Measures taken to ensure the availability of the system.
Feasibility & Implementation	Collaborative	Interdisciplinary, cross-industry and / or international cooperation if and when needed.
	Adoption Guidance	Facilitate the adoption of the standard by providing clear adoption guidelines.
	Human Centered	Useability and ease of use of AI applications.
	Education	Ensuring that all stakeholders can obtain the skills to critically assess and / or benefit from AI technology.
	Risk-Based	Risk-management based approach, depending on specific parameters of the context in which the AI application is developed, operated and assessed.
	Financial Feasibility	Ensure that the resulting standard is financially feasible for all stakeholders involved.

4.4 Discussion on the requirements

We conclude our section on findings with a brief review of each of the resulting requirements and our interpretation thereof. It is not our aim to define the resulting concepts, that is left to the developers of the standards, regulations and legislations. For readability and easy access, we have clustered the requirements into four main categories: Ethics & Society, Quality, Feasibility & Implementation.

In the “Ethics & Society” cluster, we have combined the requirements that stem from addressing ethical / societal demands on AI applications, which may or may not have been also captured by other legislation. The “ethical principles” are the first requirements in this category, which we interpret to be the underlying ethical principles, based in norms and morals, that guide human activity. These principles can yield a range of specific requirements, for instance aiming at safeguarding human rights, fairness and non-discrimination, or put negatively, aiming at avoiding gender bias, age bias and so forth. “Privacy protection” can be seen as a result of addressing the ethical principles, but has such profound impact on citizens, that we use it as a separate requirement, referring to the protection of personal data of all stakeholders of the AI application. To ensure that a standard developed for AI governance can work in harmony with existing legislation, we have added “legality and regulation” as a requirement on such standards as well. The realization of the UN’s Sustainable Development Goals is another important dimension of the ethics and society cluster and the “sustainability” requirement reflects the need to explicitly address the impact of the development, operation and consequences of AI applications on environment, ecosystems and economy. In order to achieve all these goals, it must also be possible to monitor and audit the compliance of any standard in use, which is why the “auditability” requirement has been included in this category as well.

The “quality” cluster contains properties of the standards, that are needed for stakeholders to assess the quality of the AI applications satisfying those requirements. Many of them are direct consequences from ethical and legal considerations, but the difference with that category is that the “quality” requirements aim at properties of the resulting standards themselves, rather than the considerations that have to be safeguarded in the standardization process. Firstly, one of the key areas to be addressed in the standards is the “Governance” aspect. The interplay between data owners, data handlers, modelers and so forth makes the

governance of data, algorithms, access and so forth an important aspect to be addressed in the standards. “Transparency and Explainability” refers to the possibility that stakeholders can verify how the AI application works and that outcomes of them can be traced and understood. “Trustworthiness & Responsible AI” on the other hand refers to the intentions behind the AI application and whether they are proper implementations of the ethical and other principles at the foundation of the development. In other words, the question whether these properties have been taken account already during the design stage of the application. “Accountability” follows these requirements by looking at the question, if there is somebody ultimately accountable for the negative effects, intentionally or unintentionally, of an AI deployment. Given the increasing dependence that society will have on these systems, the requirements “security” and “resilient” close this category, with the former referring to the technical measures taken to protect the confidentiality and integrity of the data and the latter to the measures taken to ensure the availability of the system, even in case of unforeseen errors or acts of God.

In the cluster “feasibility & implementation” we have combined those requirements that aim to ensure the societal and financial feasibility of the standards that are to be developed. A first requirement in this category is the need for a “collaborative” approach, where various stakeholders work together on defining the standard. This includes the requirement for interdisciplinary, cross-industry and / or international cooperation if and when needed. The “adoption guidance” requirements refers to the need to facilitate the adoption of the standard by providing clear adoption guidelines. The feasibility of an implementation is also strongly connected to the useability and ease of use of the AI application under consideration, which we cover with the “human centered” requirement. We have also put “Education” in this category, which we interpret as ensuring that all stakeholders can obtain the necessary skills to critically assess and / or benefit from AI technology. With the requirement “Risk-based approach” we refer to a type of standardization that allows for a risk-management based approach, which allows for a diversified approach depending on specific parameters of the context in which the AI application is developed, operated and assessed. Ensuring that the standards are based on and fed into scientific research and the innovation agenda of companies helps with the efficiency of the standard development process. A final element in this category is the “financial feasibility”, covering all financial aspects that might be covered to ensure that the resulting standard is financially feasible for all stakeholders involved.

5 Conclusion

This study addressed the question, to what extent the scientific body of knowledge formulates requirements on AI governance and standardization that can help finding a balance between speed and rigor. We have identified 17 requirements, which we derived from analyzing 482 scientific sources, using an approach that combined manual and AI-supported coding technique and was designed to be able to analyze a broad body of literature.

The practical relevance of this paper is twofold. Firstly, even though this was not our original aim for the research, we believe that our results can also help organizations that are preparing to develop and/or implement AI applications in identifying key areas for the governance they need to address as part of their implementation process. The areas we identified might serve as guidance for those activities, but of course more empirical research would be needed to verify this claim. Secondly, it enables us to investigate to what extent these requirements are reflected in standards and guidelines that are currently available and thus possibly contribute to their continuous improvement and applicability in practice.

The validity of our results may have been impacted by the decision to just use abstracts, in lieu of being able to take a wide perspective on the field. The impact is determined largely by how mature the field of research is and 482 results, based on fairly broad search terms suggest, that the field is indeed not yet very mature. Also, the resulting selection of the primary requirement per article might have an limiting impact. We think that the size of our sample has a mitigating effect on both these aspect, but our analysis of standards used in practice should provide a more precise estimate of the impact, perhaps leading additional requirements. In parallel we are evaluating options to perform an AI-supported full-text analysis, to enhance the depth and reliability of the research approach.

We are acutely aware of the potential risks and limitations associated with employing generative AI tools such as CoPilot and Mistral, therefore we documented the use and output of the tools according to best practices to maintain transparency and accountability insofar as possible. Our upcoming analysis of current standards will be a first indication of the practical relevance of our findings.

Acknowledgements

The authors gratefully acknowledge feedback and insightful comments provided by Thorsten Weber and Johan Versendaal.

Additional references

The references to the 482 scientific articles analyzed are accessible via:
<https://github.com/rogerbons/Bled2025Goebel-Bons-Buchkremer>

References

- Buchkremer R. et al. (2019) The Application of Artificial Intelligence Technologies as a Substitute for Reading and to Support and Enhance the Authoring of Scientific Review Articles, *IEEE Access* 7 (2019) 65263–65276. doi:10.1109/ACCESS.2019.2917719.
- Butcher J. & Beridze I. (2019) What is the State of Artificial Intelligence Governance Globally?, *RUSI Journal* 164 (5-6) (2019) 88–96. doi: 10.1080/03071847.2019.1694260.
- Chen B. et al. (2021) Security Awareness and Protection System for 5G Smart Healthcare Based on Zero-Trust Architecture, *IEEE Internet of Things Journal* 8 (13) (2021) 10248–10263. doi:10.1109/JIOT.2020.3041042.
- Corbin J. (2011) Grounded theory methodology, *vu.fernuni-hagen.de*. URL <https://vu.fernuni-hagen.de/lvuweb/lvu/file/FcU/KSW/2016SS/33800/oeffentlich/33800-vorschau.pdf>
- Dennstädt, F., Zink, J., Putora, P.M. et al. (2024) Title and abstract screening for literature reviews using large language models: an exploratory study in the biomedical domain. *Syst Rev* 13, 158 (2024). <https://doi.org/10.1186/s13643-024-02575-4>
- Djefal C., Siewert M. B., Wurster S. (2022) Role of the state and responsibility in governing artificial intelligence: a comparative analysis of AI strategies, *Journal of European Public Policy* 29 (11) (2022) 1799– 1821. doi:10.1080/13501763.2022.2094987.
- European Commission (2025) Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act). URL file:///Users/marius/Downloads/C_2025_884_1_EN_annexe_acte_autonome_cp_part1_v4_4YAkRZf5bMIWOuMqiKMIGySPo_112367.pdf
- European Commission (2024) Regulation (EU) 2024_1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300_2008, (EU) No 167_2013, (EU) No 168_2013, (EU). URL moz-extension://3f666933-3ba3-4dca-8eb5-4c50d89d2f8e/enhanced-reader.html?openApp&pdf=https%3A%2F%2Feur-lex.europa.eu%2Flegal-content%2FEN%2FTXT%2FPDF%2F%3Furi%3DOJ%3AL_202401689
- Falagas M. E. et al. (2008) Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses, *The FASEB Journal* 22 (2) (2008) 338–342. doi: 10.1096/FJ.07-9492LSF.
- Felländer A. et al. (2022) Achieving a data-driven risk assessment methodology for ethical AI, *Digital Society* 13 (1) (2022) 1–27. doi:10.1007/s44206-022-00016-0. URL <https://link.springer.com/article/10.1007/s44206-022-00016-0>
- Iren D. (2025) AI Act & Guidelines on Prohibited AI Practices: An Analysis for the Emotion Recognition Field, *LinkedIn Article*, URL <https://www.linkedin.com/pulse/ai-act-guidelines-prohibited-practices-analysis-field-iren-phd-pmp-y3the>, accessed on March 21st
- Li, M., Sun, J. & Tan, X. (2024) Evaluating the effectiveness of large language models in abstract screening: a comparative analysis. *Syst Rev* 13, 219 (2024). <https://doi.org/10.1186/s13643-024-02609-x>

- Lorenz P. (2020) AI Governance through Political Fora and Standards Developing Organizations Mapping the actors relevant to AI governance, Tech. rep., Stiftung neue Verantwortung, Berlin (2020). URL <https://ssrn.com/abstract=3731354>
- Martín-Martín A. et al. (2018) Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories, *Journal of Informetrics* 12 (4) (2018) 1160–1177. doi:10.1016/j.joi.2018.09.002.
- Novelli C. et al. (2024) AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act, *Digital Society* 3 (1). doi:10.1007/S44206-024-00095-1.
- Schmitt L. (2022) Mapping global AI governance: a nascent regime in a fragmented landscape, *AI and Ethics* 2 (2) (2022) 303–314. doi: 10.1007/s43681-021-00083-y.
- Toosi A. (2022) A brief history of AI: how to prevent another winter (a critical review), *pet.theclinics.com* doi:10.1016/j.cpet.2021.07.001. URL [https://www.pet.theclinics.com/article/S1556-8598\(21\)00053-5/abstract](https://www.pet.theclinics.com/article/S1556-8598(21)00053-5/abstract)
- Wolfswinkel et al. (2013) Using grounded theory as a method for rigorously reviewing literature (2013). doi:10.1057/ejis.2011.51.
- Zicari R. et al. (2021) Z-Inspection®: A Process to Assess Trustworthy AI, *IEEE Transactions on Technology and Society* 2 (2) (2021) 83–97. URL <https://ieeexplore.ieee.org/abstract/document/9380498/>

BUILDING GREEN GENERATIVE AI: AN ECOSYSTEM-WIDE APPROACH TO ENVIRONMENTAL SUSTAINABILITY

FABIAN HELMS, KAY HÖNEMANN, MANUEL WIESCHE

TU Dortmund University, Department of Business & Economics, Dortmund, Germany
fabian.helms@tu-dortmund.de, kay.hoenemann@tu-dortmund.de,
manuel.wiesche@tu-dortmund.de

Generative AI's (GenAI) rapid growth raises environmental concerns due to high energy consumption. Despite accelerating technological advancements, understanding how different stakeholders in the GenAI ecosystem can contribute to environmental sustainability remains limited. We address this gap with a taxonomy of actions for environmentally sustainable GenAI ecosystems. Our taxonomy, developed through a design science approach combining literature review and case analysis, categorizes environmental sustainability interventions across resources, models, and usage. We identify key stakeholders (hardware manufacturers, cloud providers, model developers, application providers) and map their roles in implementing these actions. The taxonomy reveals trade-offs between performance, cost, and environmental sustainability, highlighting the need for context-specific strategies. Through an illustrative vignette, we demonstrate how GenAI application providers can systematically implement sustainability measures. We provide a framework for researchers and practitioners to develop environmentally responsible GenAI solutions, fostering coordinated action to ensure GenAI benefits without compromising environmental well-being.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.38](https://doi.org/10.18690/um.fov.4.2025.38)

ISBN

978-961-286-998-4

Keywords:

generative AI,
sustainability,
environment,
design science,
taxonomy,
energy



University of Maribor Press

1 Introduction

By 2026, data centers and AI systems will consume approximately 1000 TWh of electricity annually, equivalent to Japan's total consumption (IEA, 2024), making energy the "primary bottleneck" for AI development (Lacey & Phillips, 2024). This surge, driven by increasingly complex AI models and popular GenAI tools like ChatGPT, strains energy infrastructure while increasing carbon emissions and water usage (Zuccon et al., 2023), creating tension between innovation and environmental sustainability.

While GenAI capabilities advance rapidly, understanding how different ecosystem stakeholders¹ can improve environmental sustainability remains limited. Current research primarily addresses technical optimizations (Bai et al., 2024; Jiang et al., 2024; Yu et al., 2024), yet environmentally sustainable deployment requires coordinated action across stakeholders. The complex relationships and roles within the GenAI ecosystem and their environmental sustainability opportunities are not yet clearly mapped.

To address this gap, we developed a taxonomy using Nickerson et al.'s (2013) design science-informed approach, combining conceptual-to-empirical and empirical-to-conceptual strategies. Taxonomies aim to conceptualize objects within a domain of interest to assist researchers and practitioners in deepening their understanding. Through a structured literature review (vom Brocke et al., 2015; Webster & Watson, 2002) and analysis of real-world cases across the ecosystem, we created a comprehensive framework that categorizes stakeholder actions and responsibilities to facilitate targeted environmental sustainability strategies throughout the GenAI ecosystem.

2 Background

Generative AI (GenAI), a technology capable of creating content like text, images, and code, relies on a complex ecosystem of interconnected components (Banh & Strobel, 2023). We can view this ecosystem as four interacting layers: the hardware layer (specialized chips like Nvidia GPUs); the cloud provider layer (Infrastructure-

¹ Stakeholders are non-human entities in the GenAI ecosystem, mostly organizations.

as-a-service offerings, e.g., AWS, Microsoft Azure, Google Cloud); the models layer (foundation models like GPT-4o to specialized models); and the applications layer (user-facing tools like ChatGPT and Midjourney). These layers interact dynamically, with each stage consuming energy as users engage with applications that utilize models hosted on cloud platforms running on specialized hardware.

Therefore, the GenAI capabilities come with a significant environmental cost, primarily through energy consumption projected to reach 1000 TWh by 2026 (IEA, 2024). This energy demand is driven by several factors: the computationally intensive training of large models, the ongoing energy required for inference, and the energy used for data storage and transfer. It is estimated that GPT-3 consumed 1,287MWh for training, while total inference demands required higher energy consumption (de Vries, 2023). Carbon intensity varies with electricity sources, which can be reduced by switching to sustainable energy sources or offsetting carbon emissions through certificates (Schwartz et al., 2020). Another aspect is the choice of models and their size, which significantly impact the energy consumption for training and inference (Argerich & Patino-Martinez, 2024; Everman et al., 2023). GenAI's rapid growth, with OpenAI's ChatGPT reaching 300 million weekly users by December 2024, further intensifies these concerns (Roth, 2024). Beyond energy, water usage is also a significant concern (Zucon et al., 2023). Current research efforts in environmental GenAI sustainability have mostly focused on specific parts of the GenAI ecosystem (Verdecchia et al., 2023), such as model training (McDonald et al., 2022), inference (Samsi et al., 2023), or benchmarking (Hodak et al., 2024). A holistic, ecosystem-level perspective is largely missing.

3 Research Design and Methodology

We developed a taxonomy following the method proposed by Nickerson et al. (2013), which has become the de facto standard for taxonomy development in Information Systems (IS) research (Szopinski et al., 2019). This approach enables systematic classification of dimensions relevant to environmental GenAI sustainability, valuable for structuring complex domains while revealing relationships between elements and their theoretical foundations (Bailey, 1994; Schöbel et al., 2020). We adapted the methodology of Kundisch et al. (2022) to ensure rigorous development of our taxonomy.

Our taxonomy aims to equip GenAI ecosystem stakeholders with a tool for developing environmentally sustainable solutions based on real-world cases, benefiting both researchers and practitioners seeking guidance on environmental sustainability and cost-influencing activities. Thus, as a first step, it is necessary to identify a meta-characteristic, which serves as the overarching purpose of the taxonomy. In our case, the meta-characteristic stated as follows: “Features and properties of GenAI ecosystems that impact environmental sustainability. Then, the next step in a taxonomy development process is to define the ending conditions that end the iterative taxonomy building process. Nickerson et. al (2013) proposed five subjective ending conditions and eight objective ending conditions, which we adopted in our building process. We employed a hybrid approach combining conceptual-to-empirical, where characteristics and dimensions are derived from literature, researcher’s existing knowledge, and individual judgment, and empirical-to-conceptual, where real-world objects are analyzed and grouped based on their shared characteristics. (Nickerson et al., 2013). Based on this approach we developed meta-dimensions (MDs), dimensions, characteristics, and their relevant stakeholders to formulate our taxonomy.

First, we conducted a structured literature review following vom Brocke et al. (2015) and Webster & Watson (2002), searching for sustainability-related GenAI publications across recognized databases (Scopus, AIS eLibrary, arXiv) with a post-2021 timeframe. We included arXiv to capture the most recent publications in this rapidly evolving field. We limited the timeframe to post-2021, as the launch of ChatGPT in 2022 marked the primary onset of GenAI ecosystem development and research. Our search strategy encompassed titles, abstracts, and keywords using the search string: ("sustainab*" OR "climat*" OR "energ*" OR "environmental") AND ("generative AI" OR "generative artificial intelligence" OR "genai" OR "llm" OR "large language model" OR "text-to-image"). After screening 1,671 articles, we identified 53 relevant publications, with four additional papers from forward and backward searching. We focused on articles addressing the direct environmental sustainability aspects of GenAI operations. Articles discussing the application of GenAI for sustainability in other fields, such as energy research (Kench et al., 2024) or enhancing climate literacy (Atkins et al., 2024), were excluded.

Table 1: Meta-Dimensions (MDs) and Dimensions

Meta-Dimensions (MDs)	Dimensions	Sources
Resources	Compute Resources	(Liu & Yin, 2024)
	Energy Resources	(Dodge et al., 2022)
Models	Model Size	(Argerich & Patino-Martinez, 2024)
	Training & Fine-tuning	(Albalak et al., 2024; Bai et al., 2024)
Usage	Inference	(Stojkovic et al., 2024)
	Input/Output	(Husom et al., 2024)
	User Interface	(Ren et al., 2023)

Source: Own

In the second iteration, to further ground our proposed taxonomy with practical relevance, we followed an empirical-to-conceptual approach by analyzing real-world cases from 20 key ecosystem stakeholders identified in an industry report, including hardware manufacturers, cloud computing providers, model providers, and application providers (Artificial Analysis, 2024). For those 20 stakeholders we conducted an internet search and analyzed the relevant documentations and applications for further information for the taxonomy building process. We identified no new characteristics in the second iteration. The taxonomy was finalized after verifying all ending conditions were met (Nickerson et al., 2013). After reaching saturation, we abstracted from the identified dimensions and characteristics to group them under meta-dimensions, which are presented in Table 1 (Nickerson et al., 2013; Strobel et al., 2024).

4 Results

The final taxonomy and their relevant stakeholders are presented in Table 2. Dimensions and characteristics are presented in the following.

Table 2: Taxonomy of Environmental Sustainability Actions in GenAI Ecosystems

MDs	Dimensions	Characteristics	Stakeholders			
Resources	Compute Resources	Hardware Choice	Hardware	Cloud	Model	Application
		Partitioning	Hardware	Cloud		
		Power Capping	Hardware	Cloud		
	Energy Resources	Tracking Emissions		Cloud	Model	Application
		Time & Location Shifting		Cloud	Model	Application
Models	Model Size	Model Compression			Model	Application
		Model Choice			Model	Application
		Specialized Models			Model	Application
	Training & Finetuning	Algorithms			Model	Application
		Data Management			Model	Application
Usage	Inference	Batching				Application
		Caching				Application
	Input/Output	Input Optimization				Application
		Output Optimization				Application
	User Interface	Energy Usage Display				Application

Source: own

4.1 Resources

4.1.1 Compute Resources

Hardware update strategies significantly reduce energy usage, as shown by the transition from NVIDIA T4 to A100 GPUs, which cuts carbon emissions by 83% for equivalent workloads (Liu & Yin, 2024). Novel architectures like memristor crossbar chips achieve 69% energy reduction compared to traditional systems (Wang et al., 2024). Several companies offer application specific integrated circuit (ASIC) chips specialized for AI computational efforts at greater energy efficiency, e.g. AWS Trainium and Inferentia or Google Tensor Processing Units (TPU) (Jouppi et al., 2023).

GPU partitioning effectively manages smaller workloads, reducing energy demand by up to 33% and enabling 55% faster fine-tuning while using 42% less energy, though with 2-9.5x slower computation (Amazon Web Services, 2025).

The power consumption of GPUs can be capped to reduce energy consumption and temperature of the chips, which can extend the lifespan of GPUs (Zhao et al., 2023). Power capping reduces GPU energy consumption during inference by 23.21% with only 6.7% increased inference time (Samsi et al., 2023). Similarly, training a BERT model with a 150W cap (vs. 250W) needs 108.5% of the time but only 87.7% of the energy (McDonald et al., 2022). A 20% lower frequency cap for the GPUs, which saves about 20% of power, can support a medium load without lowering the latency or throughput of inference (Stojkovic et al., 2024).

4.1.2 Energy Resources

Tools like FootPrinter help assess datacenter carbon footprints (Niewenhuis et al., 2024). Time and location shifting leverages varying renewable energy availability, reducing emissions for training GenAI models (Dodge et al., 2022; Jagannadharao et al., 2023). Routing inference requests to greener datacenters can reduce carbon emissions by 35% while maintaining acceptable latency (Chien et al., 2023). Leading cloud providers offer tools which allow users to track carbon emissions of their usage and inform users about the energy mix at their chosen locations.

4.2 Models

4.2.1 Model Size

Models with smaller number of parameters use less energy (Argerich & Patino-Martinez, 2024; Liu & Yin, 2024). Quantization techniques like GPT-Generated Unified Format (GGUF) reduce energy usage by decreasing numerical precision and thus computational requirements (Rajput & Sharma, 2024). A 4 bit quantized model uses less than half of the energy of a 16-bit model, while reducing inference latency 3x (Argerich & Patino-Martinez, 2024). Distillation is another approach to reduce model size by transferring knowledge from larger to smaller models for specific use cases (Alzoubi & Mishra, 2024). Some cloud providers offer services to distill models for customers' use cases.

Model selection should balance efficiency with performance requirements (Argerich & Patino-Martinez, 2024; Everman et al., 2023). Developers of end-user apps can also utilize techniques to choose the right model for the requested task by employing

model cascading or model routing. Model cascading starts the inference process with the smallest, most efficient model and only escalates requests to larger ones when necessary. This enables smaller models to answer simple tasks and only employs larger models for more complex tasks, reducing energy consumption (L. Chen et al., 2023; L. Chen & Varoquaux, 2024). Model routing assesses the complexity of the requested task to decide on the right model to answer successfully (L. Chen & Varoquaux, 2024). End-user applications like Perplexity and ChatGPT allow the user to choose a model for specific tasks. Perplexity also features an “auto mode” to select appropriate models based on users’ prompts.

Developers have to decide whether to use a base model or fine-tune a model to their specific use case, requiring additional energy usage. Fine-tuning should be used judiciously, as specialized models for medical and financial domains often don't outperform base models (Jeong et al., 2024; X. Li et al., 2023). Specialized models still outperform large general models in classification tasks such as sentiment, approval/disapproval, emotions, and party positions (Bucher & Martini, 2024). While large general models can achieve many of the tasks of specialized models in a zero-shot manner with no further fine-tuning, the energy impact of larger models is greater than those of smaller models (Luccioni et al., 2024).

4.2.2 Training & Fine-tuning

Algorithms for fine-tuning models can be optimized to reduce the computational need of fine-tuning models. Fine-tuning optimization frameworks like GreenTrainer achieve 64% reduction in computation by adaptively selecting tensors based on importance and backpropagation cost, demonstrating energy-efficient training without performance sacrifice (Huang et al., 2024). Additional efficiency approaches during model training include data parallelism (dividing datasets across nodes), model parallelism (distributing model layers across nodes), and mixed-precision training (reducing floating-point types) (Bai et al., 2024).

For LLMs trained on massive text corpora, improving dataset efficiency significantly reduces energy demands. The quality of large text datasets vary which calls for careful selection of data to train a capable model and reduce energy consumption (Albalak et al., 2024). Data pruning estimates the importance of data points in the

dataset to prioritize those data points during training (D. Chen et al., 2024; Zhuang et al., 2023).

4.3 Usage

4.3.1 Inference

Batch processing improves GPU utilization and energy efficiency, particularly for smaller models (Argerich & Patino-Martinez, 2024). Reducing batch size during low demand can save up to 15% of energy consumption (Stojkovic et al., 2024). Most API providers offer services to process requests in batches to save cost, compute, and thus energy usage.

Caching requested answers in databases reduces energy usage by retrieving stored responses for similar prompts rather than re-running inference, particularly beneficial for LLM-powered search engines (Betti et al., 2024). Most API providers offer prompt caching capabilities that store parts of prompts or documents so that in future requests models can access this cached information without reprocessing.

4.3.2 Input/Output

For multimodal LLMs, selecting only important image tokens can reduce token count by 79%, decreasing processing time by 67% and memory usage by 30% (Betti et al., 2024). Similarly, simplifying text prompts and concatenating multiple queries into single prompts reduces computational demands (L. Chen et al., 2023; Husom et al., 2024).

Image generation is more energy-intensive than text generation (Luccioni et al., 2024). When generating images with diffusion models, reducing the number of inference steps and image resolution lowers the energy usage (Seyfarth et al., 2025). Early detection of unsuccessful image generations can reduce inference time by 12% (Betti et al., 2024). For text generation, energy usage scales with the length of the generated text (Husom et al., 2024). Instructing larger models to provide concise answers can reduce carbon footprint by 40%, though with reduced accuracy for complex reasoning tasks (B. Li et al., 2024). This can be realized through most API providers by adapting the system message to generate shorter answers or defining a

maximum number of tokens to return. End-user applications such as ChatGPT and Claude also allow users to define system preferences or styles which could include an instruction to generate shorter answers.

4.3.3 User Interface

A way to save energy through reducing the use of GenAI tools is by presenting the user with information about the environmental sustainability of the continued use of the tool. Displaying energy usage information through visualizations effectively promotes sustainable usage, with users responding positively to transparency about environmental impact (Ren et al., 2023).

5 Application

We illustrate the taxonomy's practical application with a vignette describing choices and events (Miles et al., 2020). Consider "Imaginarium," a hypothetical text-to-image generation service similar to DALL-E or Midjourney. Using the taxonomy, Imaginarium can systematically analyze its operations to identify environmental sustainability opportunities across different ecosystem layers.

At the resource level (cloud provider), Imaginarium could select providers offering energy-efficient hardware (newer GPUs, TPUs) and utilize tools to track emissions. They could strategically choose data centers with more renewable energy and explore time/location shifting for training and inference.

At the model level (model/API providers, application developers), Imaginarium could offer a smaller, efficient diffusion model, potentially accepting a minor image quality trade-off for substantial energy reductions. Techniques like quantization could further minimize the model's memory and computational demands.

Within the usage layer (application developers, end-users), Imaginarium could offer batch inference for predictable loads or cache generated images for frequent prompts. Features encouraging smaller images or fewer inference steps could be implemented, with transparent feedback on energy implications. This example shows how applying the taxonomy helps Imaginarium identify concrete environmental sustainability actions across multiple GenAI ecosystem layers.

6 Discussion

We presented a taxonomy of stakeholders and activities in the GenAI ecosystem specifically designed to identify environmental sustainability opportunities. This holistic framework is crucial for understanding GenAI's complex environmental footprint. We demonstrate that responsibility for environmental GenAI sustainability cannot be assigned to a single stakeholder or confined to a specific layer within the ecosystem. It needs to be a coordinated effort, encompassing hardware manufacturers, cloud providers, model developers, application developers, and even end-users. Given the great importance of platforms for the development, access, and deployment of GenAI capabilities, the interconnectedness of these platforms and their diverse stakeholders means that actions and challenges at one level impact others, necessitating coordinated approaches for environmental sustainability (Heimburg et al., 2025).

We show inherent trade-offs between performance, cost, and environmental sustainability. For instance, larger models offer higher performance but consume significantly more energy (Argerich & Patino-Martinez, 2024; Liu & Yin, 2024). Conversely, model compression can reduce energy consumption, but may slightly decrease performance (Rajput & Sharma, 2024). The optimal approach is highly context-dependent, contingent upon the specific application, its performance requirements, and the acceptable environmental impact.

We identify numerous opportunities for enhancing efficiency at various stages, from hardware choice and data center operations to model optimization and application-level choices. These efficiency improvements not only mitigate environmental impact but can also lead to significant cost savings, creating a mutually beneficial scenario for both environmental sustainability and economic viability.

We show that the ecosystem is closely connected and measures to achieve environmental sustainability rely on different stakeholders to be achieved successfully. Transparent reporting standards of the environmental effects of GenAI operations are therefore crucial. Hardware manufacturers and cloud providers should offer standardized reports on the energy usage and environmental sustainability of their offerings. Model providers can then reliably calculate and report environmental factors for training and inference. This ecosystem-wide

transparency enables application providers and end-users to make informed choices about their GenAI usage and its environmental impact. Hugging Face, a platform for hosting GenAI models, has created a promising initiative to increase transparency in the reporting of environmental impact of different GenAI models (Hugging Face, 2025).

While we provide a comprehensive taxonomy, we acknowledge limitations and outline promising directions for future research. The GenAI ecosystem is characterized by rapid and continuous evolution. New models, hardware advancements, and innovative techniques are constantly emerging. Consequently, the taxonomy will require regular updates and refinements to maintain its relevance and accuracy in this dynamic landscape. A recent example is the release of reasoning models such as OpenAI o1 (OpenAI, 2024) and DeepSeek R1 (DeepSeek-AI et al., 2025), which, by design, produce a greater amount of tokens during inference and thus have higher energy demands. Furthermore, the taxonomy identifies opportunities but doesn't precisely quantify environmental impact for each action, focusing primarily on technical and operational aspects.

Future research should explore the role of economic incentives, such as carbon pricing mechanisms, and policy regulations to encourage sustainable practices across the ecosystem. Understanding user behavior is another critical area for future investigation. Further research is needed on how user interface design, information presentation, and user education can influence user choices and promote environmentally sustainable usage patterns.

In conclusion, we present a taxonomy promoting a holistic, environmentally sustainable approach to GenAI development and deployment. By promoting coordinated action across the entire ecosystem, we contribute to the responsible and environmentally sustainable development of this transformative technology. The application vignette demonstrates its practical value, offering a roadmap for companies improving their environmental sustainability. The goal is to ensure that the substantial benefits of GenAI can be realized without compromising long-term environmental well-being.

References

- Albalak, A., Elazar, Y., Xie, S. M., Longpre, S., Lambert, N., Wang, X., Muennighoff, N., Hou, B., Pan, L., Jeong, H., Raffel, C., Chang, S., Hashimoto, T., & Wang, W. Y. (2024). A Survey on Data Selection for Language Models (No. arXiv:2402.16827). arXiv. <http://arxiv.org/abs/2402.16827>
- Alzoubi, Y. I., & Mishra, A. (2024). Green artificial intelligence initiatives: Potentials and challenges. *Journal of Cleaner Production*, 468, 143090. <https://doi.org/10.1016/j.jclepro.2024.143090>
- Amazon Web Services. (2025). AI Accelerator—AWS Trainium—AWS. Amazon Web Services, Inc. <https://aws.amazon.com/ai/machine-learning/trainium/>
- Argerich, M. F., & Patino-Martinez, M. (2024). Measuring and Improving the Energy Efficiency of Large Language Models Inference. *IEEE Access*, 12, 80194–80207. Scopus. <https://doi.org/10.1109/ACCESS.2024.3409745>
- Artificial Analysis. (2024). Artificial Analysis AI Review—2024 Highlights. <https://artificialanalysis.ai/downloads/ai-review/2024/Artificial-Analysis-AI-Review-2024-Highlights.pdf>
- Atkins, C., Girgente, G., Shirzaei, M., & Kim, J. (2024). Generative AI tools can enhance climate literacy but must be checked for biases and inaccuracies. *Communications Earth and Environment*, 5(1). Scopus. <https://doi.org/10.1038/s43247-024-01392-w>
- Bai, G., Chai, Z., Ling, C., Wang, S., Lu, J., Zhang, N., Shi, T., Yu, Z., Zhu, M., Zhang, Y., Yang, C., Cheng, Y., & Zhao, L. (2024). Beyond Efficiency: A Systematic Survey of Resource-Efficient Large Language Models (No. arXiv:2401.00625; Version 3). arXiv. <https://doi.org/10.48550/arXiv.2401.00625>
- Bailey, K. D. (1994). *Typologies and Taxonomies: An Introduction to Classification Techniques*. SAGE Publications. <https://books.google.de/books?id=1TaYulGjhLYC>
- Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, 33(1), 63. <https://doi.org/10.1007/s12525-023-00680-1>
- Betti, F., Baraldi, L., Baraldi, L., Cucchiara, R., & Sebe, N. (2024). Optimizing Resource Consumption in Diffusion Models through Hallucination Early Detection (No. arXiv:2409.10597; Version 1). arXiv. <https://doi.org/10.48550/arXiv.2409.10597>
- Bucher, M. J. J., & Martini, M. (2024). Fine-Tuned “Small” LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification (No. arXiv:2406.08660). arXiv. <https://doi.org/10.48550/arXiv.2406.08660>
- Chen, D., Huang, Y., Ma, Z., Chen, H., Pan, X., Ge, C., Gao, D., Xie, Y., Liu, Z., Gao, J., Li, Y., Ding, B., & Zhou, J. (2024). Data-Juicer: A One-Stop Data Processing System for Large Language Models. *Companion of the 2024 International Conference on Management of Data*, 120–134. <https://doi.org/10.1145/3626246.3653385>
- Chen, L., & Varoquaux, G. (2024). What is the Role of Small Models in the LLM Era: A Survey (No. arXiv:2409.06857; Version 3). arXiv. <https://doi.org/10.48550/arXiv.2409.06857>
- Chen, L., Zaharia, M., & Zou, J. (2023). FrugalGPT: How to Use Large Language Models While Reducing Cost and Improving Performance (No. arXiv:2305.05176; Version 1). arXiv. <https://doi.org/10.48550/arXiv.2305.05176>
- Chien, A. A., Lin, L., Nguyen, H., Rao, V., Sharma, T., & Wijayawardana, R. (2023). Reducing the Carbon Impact of Generative AI Inference (today and in 2035). *2nd Workshop on Sustainable Computer Systems, HotCarbon 2023*. Scopus. <https://doi.org/10.1145/3604930.3605705>
- de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, 7(10), 2191–2194. <https://doi.org/10.1016/j.joule.2023.09.004>
- DeepSeek-AI, Guo, D., Yang, D., Zhang, H., Song, J., Zhang, R., Xu, R., Zhu, Q., Ma, S., Wang, P., Bi, X., Zhang, X., Yu, X., Wu, Y., Wu, Z. F., Gou, Z., Shao, Z., Li, Z., Gao, Z., ... Zhang, Z. (2025). DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning (No. arXiv:2501.12948). arXiv. <https://doi.org/10.48550/arXiv.2501.12948>

- Dodge, J., Prewitt, T., Tachet des Combes, R., Odmark, E., Schwartz, R., Strubell, E., Luccioni, A. S., Smith, N. A., DeCario, N., & Buchanan, W. (2022). Measuring the Carbon Intensity of AI in Cloud Instances. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 1877–1894. <https://doi.org/10.1145/3531146.3533234>
- Everman, B., Villwock, T., Chen, D., Soto, N., Zhang, O., & Zong, Z. (2023). Evaluating the Carbon Impact of Large Language Models at the Inference Stage. 150–157. *Scopus*. <https://doi.org/10.1109/IPCCC59175.2023.10253886>
- Heimburg, V., Schreieck, M., & Wiese, M. (2025). Complementor Value Co-Creation in Generative AI Platform Ecosystems. *Journal of Management Information Systems*, 42, 2.
- Hodak, M., Ellison, D., Van Buren, C., Jiang, X., & Dholakia, A. (2024). Benchmarking Large Language Models: Opportunities and Challenges. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 14247 LNCS, 77–89. *Scopus*. https://doi.org/10.1007/978-3-031-68031-1_6
- Huang, K., Yin, H., Huang, H., & Gao, W. (2024). TOWARDS GREEN AI IN FINE-TUNING LARGE LANGUAGE MODELS VIA ADAPTIVE BACKPROPAGATION. 12th International Conference on Learning Representations, ICLR 2024. *Scopus*.
- Hugging Face. (2025). AI Energy Score. AI Energy Score. <https://huggingface.github.io/AIEnergyScore/>
- Husom, E. J., Goknil, A., Shar, L. K., & Sen, S. (2024). The Price of Prompting: Profiling Energy Use in Large Language Models Inference (No. arXiv:2407.16893; Version 1). *arXiv*. <https://doi.org/10.48550/arXiv.2407.16893>
- IEA. (2024, January 24). Electricity 2024. IEA. <https://www.iea.org/reports/electricity-2024>
- Jagannadharao, A., Beckage, N., Nafus, D., & Chamberlin, S. (2023). Timeshifting strategies for carbon-efficient long-running large language model training. *Innovations in Systems and Software Engineering*. *Scopus*. <https://doi.org/10.1007/s11334-023-00546-x>
- Jeong, D. P., Garg, S., Lipton, Z. C., & Oberst, M. (2024). Medical Adaptation of Large Language and Vision-Language Models: Are We Making Progress? (No. arXiv:2411.04118). *arXiv*. <https://doi.org/10.48550/arXiv.2411.04118>
- Jiang, P., Sonne, C., Li, W., You, F., & You, S. (2024). Preventing the Immense Increase in the Life-Cycle Energy and Carbon Footprints of LLM-Powered Intelligent Chatbots. *Engineering*, 40, 202–210. <https://doi.org/10.1016/j.eng.2024.04.002>
- Jouppi, N. P., Kurian, G., Li, S., Ma, P., Nagarajan, R., Nai, L., Patil, N., Subramanian, S., Swing, A., Towles, B., Young, C., Zhou, X., Zhou, Z., & Patterson, D. (2023). TPU v4: An Optically Reconfigurable Supercomputer for Machine Learning with Hardware Support for Embeddings (No. arXiv:2304.01433). *arXiv*. <https://doi.org/10.48550/arXiv.2304.01433>
- Kench, S., Squires, I., Dahari, A., Brosa Planella, F., Roberts, S. A., & Cooper, S. J. (2024). Li-ion battery design through microstructural optimization using generative AI. *Matter*, 7(12), 4260–4269. *Scopus*. <https://doi.org/10.1016/j.matt.2024.08.014>
- 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. <https://doi.org/10.1007/s12599-021-00723-x>
- Lacey, S., & Phillips, N. (2024, May 22). Energy is now the ‘primary bottleneck’ for AI. *Latitude Media*. <https://www.latitudemedia.com/news/energy-is-now-the-primary-bottleneck-for-ai>
- Li, B., Jiang, Y., Gadepally, V., & Tiwari, D. (2024). Toward Sustainable GenAI using Generation Directives for Carbon-Friendly Large Language Model Inference (No. arXiv:2403.12900; Version 1). *arXiv*. <https://doi.org/10.48550/arXiv.2403.12900>
- Li, X., Chan, S., Zhu, X., Pei, Y., Ma, Z., Liu, X., & Shah, S. (2023). Are ChatGPT and GPT-4 General-Purpose Solvers for Financial Text Analytics? A Study on Several Typical Tasks (No. arXiv:2305.05862). *arXiv*. <http://arxiv.org/abs/2305.05862>
- Liu, V., & Yin, Y. (2024). Green AI: Exploring carbon footprints, mitigation strategies, and trade offs in large language model training. *Discover Artificial Intelligence*, 4(1). *Scopus*. <https://doi.org/10.1007/s44163-024-00149-w>

- Luccioni, S., Jernite, Y., & Strubell, E. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency, 85–99. <https://doi.org/10.1145/3630106.3658542>
- McDonald, J., Li, B., Frey, N., Tiwari, D., Gadepally, V., & Samsi, S. (2022). Great Power, Great Responsibility: Recommendations for Reducing Energy for Training Language Models. 1962–1970. Scopus.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). Qualitative data analysis: A methods sourcebook (Fourth edition). SAGE.
- 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. <https://doi.org/10.1057/ejis.2012.26>
- Niewenhuis, D., Talluri, S., Iosup, A., & De Matteis, T. (2024). FootPrinter: Quantifying Data Center Carbon Footprint. 189–195. Scopus. <https://doi.org/10.1145/3629527.3651419>
- OpenAI. (2024). Learning to reason with LLMs. <https://openai.com/index/learning-to-reason-with-llms/>
- Rajput, S., & Sharma, T. (2024). Benchmarking Emerging Deep Learning Quantization Methods for Energy Efficiency. 238–242. Scopus. <https://doi.org/10.1109/ICSA-C63560.2024.00049>
- Ren, Y., Sivakumaran, A., Niemelä, E., & Jääskeläinen, P. (2023). How to Make AI Artists Feel Guilty in a Good Way? : Designing Integrated Sustainability Reflection Tools (SRTs) for Visual Generative AI. ICCI International Conference of Computational Creativity, Waterloo, Canada, 19-23 June, 2023. <https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-327918>
- Roth, E. (2024, December 4). ChatGPT now has over 300 million weekly users. The Verge. <https://www.theverge.com/2024/12/4/24313097/chatgpt-300-million-weekly-users>
- Samsi, S., Zhao, D., McDonald, J., Li, B., Michaleas, A., Jones, M., Bergeron, W., Kepner, J., Tiwari, D., & Gadepally, V. (2023). From Words to Watts: Benchmarking the Energy Costs of Large Language Model Inference. 2023 IEEE High Performance Extreme Computing Conference, HPEC 2023. Scopus. <https://doi.org/10.1109/HPEC58863.2023.10363447>
- Schöbel, S. M., Janson, A., & Söllner, M. (2020). Capturing the complexity of gamification elements: A holistic approach for analysing existing and deriving novel gamification designs. *European Journal of Information Systems*, 29(6), 641–668. <https://doi.org/10.1080/0960085X.2020.1796531>
- Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63. <https://doi.org/10.1145/3381831>
- Seyfarth, M., Dar, S. U. H., & Engelhardt, S. (2025). Latent Pollution Model: The Hidden Carbon Footprint in 3D Image Synthesis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 15187 LNCS, 146–156. Scopus. https://doi.org/10.1007/978-3-031-73281-2_14
- Stojkovic, J., Choukse, E., Zhang, C., Goiri, I., & Torrellas, J. (2024). Towards Greener LLMs: Bringing Energy-Efficiency to the Forefront of LLM Inference (No. arXiv:2403.20306; Version 1). arXiv. <https://doi.org/10.48550/arXiv.2403.20306>
- Strobel, G., Banh, L., Möller, F., & Schoormann, T. (2024). Exploring Generative Artificial Intelligence: A Taxonomy and Types. *Hawaii International Conference on System Sciences 2024 (HICSS-57)*. https://aisel.aisnet.org/hicss-57/in/platform_ecosystems/7
- Szopinski, D., Schoormann, T., & Kundisch, D. (2019, June 13). Because your taxonomy is worth it: Towards a framework for taxonomy evaluation. ECIS.
- Verdecchia, R., Sallou, J., & Cruz, L. (2023). A systematic review of Green AI. *WIREs Data Mining and Knowledge Discovery*, 13(4), e1507. <https://doi.org/10.1002/widm.1507>
- vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems*, 37(1). <https://doi.org/10.17705/1CAIS.03709>

- Wang, Z., Luo, T., Liu, C., Liu, W., Goh, R. S. M., & Wong, W.-F. (2024). Enabling Energy-Efficient Deployment of Large Language Models on Memristor Crossbar: A Synergy of Large and Small. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–17. IEEE Transactions on Pattern Analysis and Machine Intelligence. <https://doi.org/10.1109/TPAMI.2024.3483654>
- Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii.
- Yu, Y., Wang, J., Liu, Y., Yu, P., Wang, D., Zheng, P., & Zhang, M. (2024). Revisit the environmental impact of artificial intelligence: The overlooked carbon emission source? *Frontiers of Environmental Science and Engineering*, 18(12). Scopus. <https://doi.org/10.1007/s11783-024-1918-y>
- Zhao, D., Samsi, S., McDonald, J., Li, B., Bestor, D., Jones, M., Tiwari, D., & Gadepally, V. (2023). Sustainable Supercomputing for AI: GPU Power Capping at HPC Scale. 588–596. Scopus. <https://doi.org/10.1145/3620678.3624793>
- Zhuang, B., Liu, J., Pan, Z., He, H., Weng, Y., & Shen, C. (2023). A survey on efficient training of transformers. *Proceedings of the 32nd International Joint Conference on Artificial Intelligence, IJCAI 2023*, 6823–6831. <https://research.monash.edu/en/publications/a-survey-on-efficient-training-of-transformers>
- Zuccon, G., Scells, H., & Zhuang, S. (2023). Beyond CO2 Emissions: The Overlooked Impact of Water Consumption of Information Retrieval Models. 283–289. Scopus. <https://doi.org/10.1145/3578337.3605121>

CAUSAL EFFECTS OF ADDRESSING AND PSI ON COMPLIANCE WITH A SOCIAL MEDIA INFLUENCER'S RECOMMENDATIONS

JUSSI NYRHINEN,¹ EERO RANTALA,¹ JESSE TUOMINEN,²
MARKUS MAKKONEN,³ TERHI-ANNA WILSKA,⁴
LAURI FRANK⁵

¹ Lappeenranta-Lahti University of Technology, Faculty of Engineering Sciences,
Lappeenranta, Finland
jussi.nyrhinen@lut.fi, eero.rantala@lut.fi

² University of Jyväskylä, Jyväskylä, Finland
jesse.o.tuominen@jyu.fi

³ Tampere University, Faculty of Management and Business, Tampere, Finland
markus.makkonen@tuni.fi

⁴ University of Jyväskylä, Faculty of Humanities and Social Sciences, Jyväskylä, Finland
terhi-anna.wilka@jyu.fi

⁵ University of Jyväskylä, Faculty of Information Technology, Jyväskylä, Finland
lauri.frank@jyu.fi

Our between-subjects experimental study on US citizens aged 18-35-years ($n = 401$) revealed that the young people are more likely to comply with recommendations by influencers who address the viewers of their YouTube videos. The viewers also experienced higher levels of parasocial interaction if an influencer addressed them verbally and physically in the video. The experience of parasocial interaction increased the likelihood to comply with the influencer's recommendations. Our findings emphasize the role of humane and social factors in content marketing.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.39](https://doi.org/10.18690/um.fov.4.2025.39)

ISBN
978-961-286-998-4

Keywords:
addressing,
compliance,
experiment,
parasocial interaction,
social media influencer



University of Maribor Press

1 Introduction

The present study delves into the role of parasocial interaction (PSI) in social media influencer marketing. PSI is a viewer's illusionary experience of being engaged in social interaction with media performers (Horton & Wohl, 1956). Despite the relationship is unilateral, viewers may experience that the influencer is aware of their presence, pays attention to them, or reacts to their behavior (Dibble, Hartmann, & Rosaen, 2016; Horton & Wohl, 1956). Therefore, PSI evokes feelings of closeness and connection to the performers, such as social media influencers (Aw & Labreque, 2020). The experience of PSI is also utilized in influencer marketing as the consumers who experience PSI are likely to purchase the products the influencers promote (Lee & Lee, 2021) and make impulsive purchases on social media platforms (Xiang et al., 2016). It is shown that influencers' verbal and bodily addressing lead viewers to experience PSI (Cummins & Cui, 2014, Hartmann & Goldhoorn, 2011; Tukachinsky & Sangalang, 2016). Also, addressing effects viewer's attitudes and behavioral intentions (Wei et al., 2019). Bodily addressing refers to facial expressions and eye-contact. Prior research has shown that eye-contact ignites automatic mindreading i.e., inferring others' mental state (Malle & Hodges, 2005) which may lead to a parasocial experience (Hartmann & Goldhoorn, 2011). Also, for instance, a smile is seen as an invitation for cooperative behavior and (Horstmann, 2003). Verbal addressing occurs when a performer implicitly or explicitly verbally addresses the viewers (e.g., speaks to and engages the audience using personal pronouns) (Cummins & Cui, 2014).

Despite the importance of the topic for influencer marketing, there is scarce evidence on whether addressing causally affects behavioral outcomes through the experience of PSI. The present study addresses this research gap by conducting a between-subjects experiment in which randomization of addressing is used as instrument variable for PSI. The purpose of the experiment was to tease out the causal effects addressing to PSI and consumers' decisions to listen to a recently published song in YouTube when the influencer recommends listening to it. The paper examines the topic in the context of young consumers, because they are highly exposed to influencer marketing and are susceptible to peer influence and social norms within the realm of social media (Nyrhinen et al, 2024). Furthermore, it is topical to examine parasocial interaction in social media, because recent research has associated parasocial interaction with the experience of immersion in contexts such

as streaming commerce (Liao et al., 2023). Immersion is considered as important when using interactive digital technology (Paananen et al., 2024)

We also contribute to the methodologies on research on social media marketing by applying advanced causal modelling with experimental data. To the best of our knowledge, we are among the first in the field to appropriately conduct analysis with instrumental variable from non-parametric perspective without linear assumptions. However, we can get a precise estimate of a causal effect of a subgroup of population while linear modelling techniques like regression modelling assume effect homogeneity (i.e., they rely on assumption that all individuals in population have similar homogenic effect), which is not supported in social sciences.

2 Hypothesis

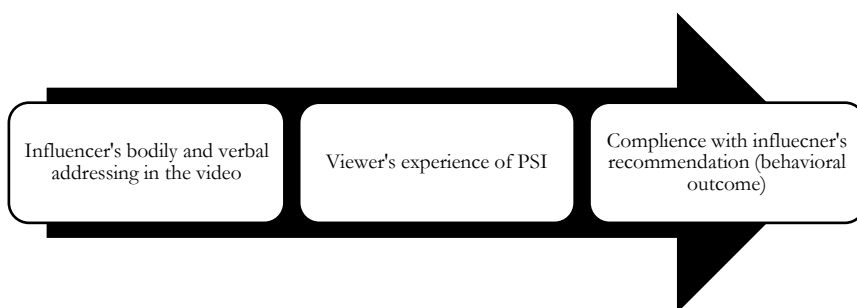


Figure 1: Illustration of the conceptual model

Research indicates that consumers who experience PSI with influencers are more inclined to follow the influencers' recommendations. This is because the experience PSI with influencer enhances consumer confidence, particularly when they are uncertain about their preferences (Penttinen et al., 2022). More precisely, when influencers employed both bodily and verbal addressing, the videos resulted in significantly elevated experiences of PSI among viewers, in contrast to instances where such addressing was absent (Hartmann & Goldhoorn, 2011). On the other hand, there is evidence that addressing alone effects viewer's attitudes and behavioral intentions (Wei et al., 2019). For this reason, it is worth investigating whether there is causal relationship between an influencer's bodily and verbal addressing and

compliance with influencers recommendations through the experience of PSI (see Figure 1). Therefore, we postulate:

H1. Addressing is causally related to compliance with a social media influencer's recommendation through the experience of PSI.

3 Methodology

3.1 Participants and Procedure

Data were collected as a part of a research project (see Tuominen et al., 2025) from 401 US citizens aged 18–35 years recruited from a crowd sourcing platform. The participants were assigned to a treatment in a between-subjects experimental design with a short video in which a nano-influencer recommended listening to a song. We manipulated whether the influencer addressed the participants both verbally and bodily or not. Hence, addressing was used as a randomization as instrumental variable for a latent construct of PSI that cannot be randomized. This was done to estimate the effect of PSI to the behavioral outcome.

The study included two videos: first in which the influencer recommended a song titled “This Year Be” by the band Left Vessel with high addressing, a second in which the influencer recommended the same song with low addressing. After watching the video, participants decided whether they wanted to listen to “This Year Be” or a song titled “Eclipse” by the band Wages that the influencer did not mention in her video. “Eclipse” was described as having fewer views than “This Year Be” so that it would appear a tempting alternative to “This Year Be”. After that, participants reported their level of parasocial interaction with a single item “While watching the video clip, I had the feeling that the speaker knew I was there” (0 = Disagree, 1 = Agree). The item was drawn from the Experience of Parasocial Interaction Scale (Dibble et al., 2016; Hartmann & Goldhoorn, 2011).

The addressing was manipulated with a method that has been validated in previous studies (Dibble et al., 2016; Hartmann & Goldhoorn, 2011). In the high addressing treatment, influencer smiled and performed (bodily) directly to camera (see Figure 2). She also addressed viewers in the first persona, wished them welcome, greeted and thanked them.

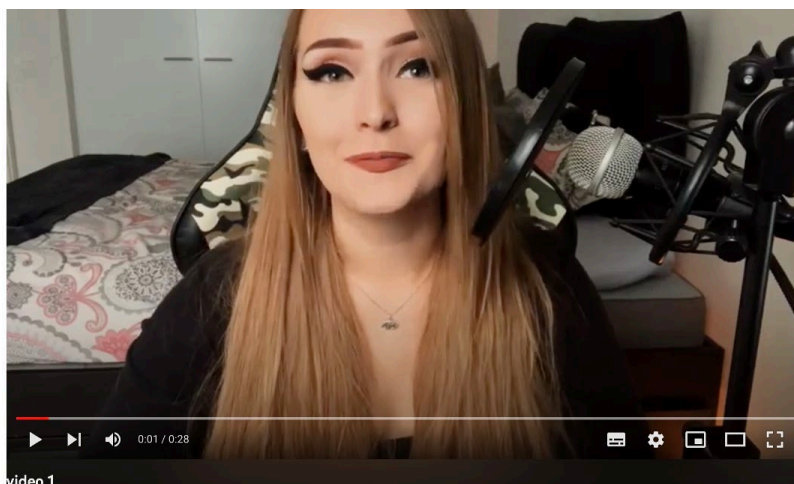


Figure 2: An illustration of high bodily addressing

Source: Own

In the low addressing treatment, the influencer did not smile and looked away from the camera (see Figure 3). Also, unlike in the high addressing treatment, the influencer did not address the viewers personally, thanked, or greeted them.

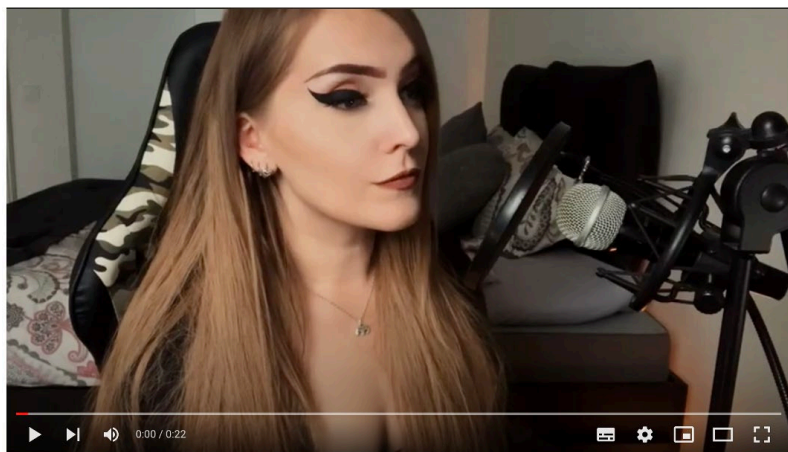


Figure 3: An illustration of low bodily addressing

Source: Own

3.3 Analysis Strategy

We estimated the causal effect of PSI to compliance with an influencer's recommendation with an experimental research design with nonparametric approach. This approach relies on two main assumptions: monotonicity and no direct effect of instrument to behavioral outcome (see more Imbens & Rubin, 2015). We chose this analysis strategy, because PSI is an observed latent variable that cannot be randomized.

4 Results

Complier average causal effect (late) is defined as follows (see Imbens & Rubin, 2015). Intention to treat effects (ITT) are the effect of treatment for each variable. In case of our study y variable is the behavioral outcome and w variable is the experience of PSI. We get expected value ($\mathbb{E}$) for effect of PSI to behavioral to outcome in the subpopulation of compliers ($G_i = \text{co}$) by dividing the ITT_y by ITT_w . This causal effect can be estimated following the equation below with the observed sample analogs of ITT effects (i.e., the mean differences = $[Y_i(1) - Y_i(0)]$).

$$\tau_{\text{late}} = \frac{ITT_y}{ITT_w} = \mathbb{E}[Y_i(1) - Y_i(0) | G_i = \text{co}] \dots \dots \dots (1)$$

Our results support H1 by revealing that if a participant experiences PSI, it increases probability for compliers to follow influencers recommendation 2.5726 percentage points (see Table 1). Here the members of population who react to treatment as intended are referred as compliers. See about compliance status and generality of estimates from Imbens & Rubin (2015).

Table 1: Estimated average causal effect

Variable	Treatments	Means	Quotient	Result
Behavioral Outcome (Y)	Low Addressing	.3604		
	High Addressing	.5392	.1788	
PSI (W)	Low Addressing	.7148		
	High Addressing	.7843	.0695	
			.1788/.0695	2.5726

Note: Experience of PSI (0 = low, 1 = High); Behavioral outcome = complying with the influencer's recommendation (0 = did not comply, 1 = complied)

5 Conclusion

Our findings supported the hypothesis by revealing that the besides directly increasing probability to comply with influencer's recommendations, addressing has also causal effect to the behavioral outcome through the experience of PSI. In line with previous studies (Malle & Hodges, 2005; Hartmann & Goldhoorn, 2011; Horstmann, 2003, Cummins & Cui, 2014), our findings supported the notion that direct addressing (e.g., eye contact, smile, use of personal pronouns) increased the viewers' illusion that the influencer was aware of their presence and especially probability to comply with the influencer's recommendations. However, in contrast to prior studies, the participants of the present study experience high levels of PSI regardless the treatment of addressing. Even though the high level of addressing slightly elevated the experience of PSI in comparison with the treatment with low addressing, our findings indicate that just a presence of a human on a video can create an experience of PSI.

5.1 Managerial Implications

Our results emphasize the role of social and humane factors in social media influencer marketing. Especially, we highlight the importance of the experience of PSI as a metric when assessing the content of influencer marketing. Also, regarding the measurement, our pioneering research elucidates how randomization of treatment can be used accurately as an instrument variable for latent observed variables such as viewers' experiences.

5.2 Limitation and Avenues for Future Research

The present study had certain limitations that open avenues for future research. First, the study focused on addressing in general, but the effects of bodily and verbal addressing could be also studied separately. Future studies could also examine the effect on valence of the video to PSI and likelihood to comply with the influencer's suggestions.

Second, only one of six items in the EPSI scale (Dibble et al., 2016; Hartmann & Goldhoorn, 2011) was found to measure the experience of PSI in the present study. The item measured whether the respondents had an illusion that the influencer was

aware of their presence. There were no significant differences in the other EPSI items between the treatments. It is worth investigating whether the present study was focused solely on a certain aspect of PSI or whether more research related to metrics of PSI is needed.

Third, the focus of the present study was a video mediated human-to-human communication in an online video streaming platform, namely YouTube. The future studies could also examine addressing and PSI in the context of augmented and virtual reality. In such context it would be beneficial to delve into the causal effects of these variable on experience of immersion.

Fourth, the present study supported the earlier studies by showing that the positive causal relationships between addressing, PSI, and behavioral response. However, due to the emerge of AI based virtual assistants, it worth investigating if addressing by a virtual assistant can evoke an experience of PSI or is it considered intrusive or uncanny instead.

References

- Aw, E. C.-X., & Labrecque, L. I. (2020). Celebrity Endorsement in Social Media Contexts: Understanding the Role of Parasocial Interactions and the Need to Belong. *Journal of Consumer Marketing*, 37(7), 895–908.
- Cummins, R. G., & Cui, B. (2014). Reconceptualizing address in television programming: The effect of address and affective empathy on viewer experience of parasocial interaction. *Journal of Communication*, 64(4), 723–742.
- Dibble, J. L., Hartmann, T., & Rosaen, S. F. (2016). Parasocial Interaction and Parasocial Relationship: Conceptual Clarification and a Critical Assessment of Measures. *Human Communication Research*, 42(1), 21–44.
- Hartmann, T., & Goldhoorn, C. (2011). Horton and Wohl Revisited: Exploring Viewers' Experience of Parasocial Interaction. *Journal of Communication*, 61(6), 1104–1121.
- Horstmann, G. (2003). What Do Facial Expressions Convey: Feeling States, Behavioral Intentions, or Actions Requests? *Emotion*, 3(2), 150–166.
- Horton, D., & Wohl, R. R. (1956). Mass Communication and Para-Social Interaction. *Psychiatry*, 19(3), 215–229.
- Imbens, G. W., & Rubin, D. B. (2015). *Causal inference for statistics, social, and biomedical sciences: An introduction*. Cambridge University Press.
- Lee, M., & Lee, H.-H. (2022). Do Parasocial Interactions and Vicarious Experiences in the Beauty YouTube Channels Promote Consumer Purchase Intention? *International Journal of Consumer Studies*, 46(1), 235–248.
- Liao, J., Chen, K., Qi, J., Li, J., & Yu, I. Y. (2023). Creating immersive and parasocial live shopping experience for viewers: the role of streamers' interactional communication style. *Journal of research in interactive marketing*, 17(1), 140–155.

- Malle, B., & Hodges, S. (2005). *Other Minds: How Humans Bridge the Divide between Self and Others*. New York: The Guilford Press.
- 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.
- Paananen, T., Kemppainen, T., Makkonen, M., Holkkola, M., Tyrväinen, O., & Frank, L. (2024). Towards Techno-Psychological Immersion: A Narrative Literature Review of Immersion and Its Related Concepts. In A. Pucihar, M. Kljajić Borštnar, S. Blatnik, R. W. H. Bons, K. Smit, & M. Heikkilä (Eds.), *37th Bled eConference: Resilience Through Digital Innovation : Enabling the Twin Transition* (pp. 577-596). University of Maribor.
- Penttinen, V., Ciuchita, R., & Čaić, M. (2022). YouTube It Before You Buy It: The Role of Parasocial Interaction in Consumer-to-Consumer Video Reviews. *Journal of Interactive Marketing*, 57(4), 561–582.
- Tukachinsky, R., & Sangalang, A. (2016). The Effect of Relational and Interactive Aspects of Parasocial Experiences on Attitudes and Message Resistance. *Communication Reports*, 29(3), 175–188.
- Tuominen, J., Rantala, E., Nyrhinen, J., Reinikainen, H., Warren, C., Wilska, T. (2025 in press). When does parasocial interaction make influencers more influential? *Human Technology*.
- Wei, L., Ferchaud, A., & Liu, B. (2019). Endorser and bodily addressing in public service announcements: Effects and underlying mechanisms. *Communication Research Reports*, 36(1), 1-13.
- Xiang, L., Zheng, X., Lee, M. K. O., & Zhao, D. (2016). Exploring Consumers' Impulse Buying Behavior on Social Commerce Platform: The Role of Parasocial Interaction. *International Journal of Information Management*, 36(3), 333–347.

FROM SIMPLE SANDBOX PROCESS TO REGULATORY SANDBOX FRAMEWORK: SERVING THE DUAL OBJECTIVES OF AI REGULATION

JUKKA HEIKKILÄ,¹ ALTTI LAGSTEDT,² SEPPO HEIKURA,¹
EEVA KAAKKURIVAARA,¹ ANASTASIIA DARDYKINA,³
ALICE TEILHARD DE CHARDIN⁴

¹ University of Turku, Turku School of Economics, Turku, Finland
jups@utu.fi, seppo.heikura@utu.fi, eeva.e.kaakkurivaara@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
anastasii.a.dardykina@uwasa.fi

⁴ ALLAI, Amsterdam, the Netherlands
alice.teilhard-de-chardin@allai.nl

In this paper, we build on a simplified process to evaluate AI systems compliance with AI Act. As AI Act's risk-based regulatory measures simultaneously strive for providing innovative, but trustworthy AI systems to the market, we extend the process to an AI Regulatory Sandbox Framework (AIRSF), which is connected with the EU's innovation ecosystem. The study builds on a design science approach combined with action research sandbox as a boundary object for multistakeholder sandbox development. We first analyse the options for regulating AI Systems. In doing so, we explore the regulatory sandbox as a means aiming to remedy the known limitations of regulatory sandboxes, especially from the viewpoint for SMEs. We conclude by reflecting upon the findings during the sandbox development process.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.40](https://doi.org/10.18690/um.fov.4.2025.40)

ISBN

978-961-286-998-4

Keywords:

AI act,
regulatory sandbox,
innovation ecosystem,
compliance,
evaluation,
design science,
action research



University of Maribor Press

1 Introduction

The European Union (EU) is seeking to boost the development and deployment of AI systems by attempting to strike a balance between innovation and regulation. One proposed mechanism for achieving this balance is the implementation of AI Regulatory Sandboxes to improve business viability and compliance, particularly with fundamental rights.

However, AI regulatory sandboxes are not a silver bullet to this end and multiple challenges currently stand. First, the AI Act's requirements to identify and mitigate risks to fundamental rights are often perceived as distant from the practical realities faced by AI providers, deployers, and regulators (Orwat et al., 2024). In addition, the role of Regulatory Sandboxes remains unclear. Second, while the AI Act (AIA) mandates free access to supervision and testing environments for small and medium-sized enterprises (SMEs) and start-ups, supported by various EU and Member State innovation initiatives, the specifics of their implementation is still evolving. Thirdly, whilst introducing an AI System on the European Common Market is possible without sandbox participation, AI providers and deployers may expose themselves to non-compliance risks. For example, if they are unable to demonstrate adequate risk assessment and mitigation measures prior to market entry. Finally, the objectives of regulation seem contradictory (e.g., Lanamäki et al., 2024) and the practical deployment of sandboxes is still under development.

2 Research Questions and Method

The core research question underpinning this paper is: **should and could regulatory sandboxes be practically implemented for AI as a Europe-wide mechanism?** We divide this into further sub-questions:

- What does the AIA require from sandbox solutions to balance the promotion of innovative AI uses with compliance requirements?
- What factors must be considered in designing an attractive sandbox, particularly for SMEs?
- What insights have we gained during the development of such a process?

To elucidate these questions, we adopt a constructive research approach of design science research by van Aken and Romme, 2009). We aim to develop a generalizable design proposition to address the identified challenges in practice or identified in prior research when implementing AI regulatory sandboxes. The effectiveness of our proposed design solution will be evaluated over the course of our on-going project. During the project we engage a range of relevant experts, regulators, companies and vested parties in the co-designing of eventual sandbox.

Our intervention in the sandbox process is conceptualised as a Stage-Gate model (see e.g. Cooper, 1990), providing a modular process that allows for iterative incorporation of details, and connected services. Our Stage-Gate model serves as a shared design artefact to classify, evaluate and elaborate on risk factors associated with AI in use cases. Methodologically, our research approach integrates elements of Action Research with Design Science Research (Iivari and Venables, 2009), because the proposed solution design is adaptable to diverse AI implementation contexts, industries, and domains. (e.g. Buocz et al., 2023; Birkstedt et al., 2024; von Steffen, 2025), The purpose of our baseline process is to serve as a boundary object for further development of Sandbox in co-operation with various stakeholders.

3 Regulatory Approaches for AI Innovations

The EU has gained a reputation as the preeminent technological regulator, initially establishing regulation for unified internal digital markets before expanding to regulate international technology corporations to promote a more competitive internal market. This dual approach aims to foster innovation while simultaneously safeguarding user and worker rights and ensuring fair competition. The resulting landscape comprises complex set of overlapping and sometimes contradicting regulation that impact the operation of technology-driven companies by imposing constraints on data collection, processing and sharing practices, as well as on interactions with internet users, or businesses (Bradford, 2024).

Meanwhile, some have claimed the dominance of less restrictive AI policies in the United States, innovation-focused policies in China have cast a shadow over the European ideal of 'control-first, innovate-thereafter'. There are growing pressures (Draghi, 2024) to avoid what critics have deemed the confusing, contradicting over-regulation of the market and instead to boost AI-related innovations. In response,

the EU has established several innovation-fostering initiatives including the European Digital Innovation Hubs (EDIH), sectoral Testing and Experimentation Facilities (TEFs), and AI Factories, which are complemented by national business incubators and start-up ecosystems. These developments underscore the tension made evident by the introduction of the AIA: the question is no longer whether to regulate AI, but how to do so in balance with the initiatives and activities that support innovation.

One approach relies on leveraging voluntary *ethical principles* as an alternative to regulation, to complement sole business logic. However, as Buhmann and Fieseler (2021) illustrate, the opacity of AI makes obscure the decisions made by designers. As the laymen are not involved, they cannot determine the outcome before it is too late – they cannot afford to rely on designers' following voluntary ethical principles only. By implementing such principles, innovators are positioned as proactive representatives of public discourse, tasked with discovering to ethically sound solutions through responsible innovation processes. Concerns persist regarding whether they really can properly represent the perspectives of all stakeholders affected by the influence of an AI system (Brown and Pironka, 2021). The EU is not alone in its concerns that technology companies cannot fully foresee how their AI solutions might compromise individuals' fundamental rights or undermine established democratic institutions and practices. Indeed, the OECD AI policy observatory, reflects an emerging consensus that AI development prioritises societal benefit while ensuring equitable conditions for market actors and participants thereby justifying and necessitated coordinated global regulatory frameworks. (OECD, 2022).

Regulatory implementation varies depending on the underlying regulatory strategy (these are mutually exclusive), which may involve a) comprehensive overhauling updates to existing laws, b) creation of novel exclusive regulations, or c) engagement in experimental/incremental regulation (the category which the EU AIA belongs) (Sloane and Wüllhorst, 2025). The latter is characterised as *middle-out regulation* because it is positioned between conventional 'hard law' and industry self-regulation (Pagallo et al., 2019). In a middle-out approach, stakeholders engage in mutual learning and iterative adaption of the regulation. This approach is getting preferred by scholars for regulating rapidly evolving technology in society (e.g., Ranchordas, 2021).

Regulatory sandboxes are a primary means for implementing middle-out regulation and are gaining traction (Zetsche et al., 2017; Longo and Bagni, 2025) as general means to provide guidance for the development and testing of innovations under conditions isolated from the markets. Sandboxes have been used for these purposes since the financial crisis in FinTech (e.g., Zetsche et al., 2017) and for evaluating compliance with privacy regulation (GDPR).

According to Sloane and Wüllhorst (2025), Dardykina (2024), and Seferi's (2025) analyses, contemporary sandboxes are implemented in various ways. At their core, regulatory sandboxes are defined as *"a controlled framework established by a competent authority in which participants – whether public or private – can develop, validate, and test innovative products under regulatory supervision for a limited period of time"* (Longo and Bagni, 2025, p. 28). However, initial implementation has demonstrated a *"The regulatory sandbox alone as presently structured is typically too limited in scope and scale to promote further meaningful innovation"* (Zetsche et al., 2017; c.f. Dardykina, 2024, p. 4), providing the motivation for our research.

Therefore, addressing our second research question, we analyze whether a sandbox suits our situation. Recent AI regulatory sandbox trials (Truby et al., 2022; Buocz, 2023; von Steffen, 2025), have found the main challenge for AIA sandboxes centred around access to the sandbox and the dual objectives of fostering innovation while ensuring regulatory compliance. Frameworks for balancing these goals have been proposed (e.g. Janssen et al., 2022; Baldini and Frances, 2024), as well as software tools (OECD, 2022). As Trusby notes: *"... it [a sandbox] exposes the developer to compliance and setup costs, an added layer of regulatory supervision, and it exposes the AI technology to regulators and third parties. In exchange for these risks, a regulatory sandbox should minimise regulatory constraints, lower the risk of regulatory enforcement and liability and provide ongoing guidance from regulators."* (Truby, 2021, p. 291). At present, it remains uncertain whether this intended mutual exchange of value will materialize.

Nevertheless, AIA act requires the duality of balance. Díaz-Rodríguez et al. (2024) propose an ideal for trustworthy AI by dividing regulatory actions to *ex ante* evaluation and *ex post* Market Surveillance for meeting the dual objectives. Their drafted procedure leaves many questions open for practical implementation, such as when is the system ready for compliance evaluation? What if it gets rejected, can it return some day after further development and revision? What if there are domain

specific regulations to be considered beyond AI Act requirements? Can an AI System candidate be altered to meet the requirements while in the sandbox?

Considering the AIA's objectives, we argue greater attention should be paid to sandboxes as enablers for improving AI systems' market viability. To this end, sandboxes should relate to broader innovation support ecosystems such as funding programmes, peer support and expert guidance to provide varied forms of compliance evidence, especially from the National Supervisory Authorities.

3.2 Special Requirements of SMEs

Our understanding of the needs of SMEs for the features of sandboxes emerge from both the literature and through our collaboration with authorities and experts. Despite the well-intentioned initiatives of the AIA, SMEs often struggle to adopt or create AI innovations due to the rapidly evolving regulatory environment, complex and inconsistent policies, and the serious legal consequences of non-compliance (Muminova et al., 2024; Ardito et al. (2025)). These challenges are burdensome for SMEs because they typically work with limited legal and administrative resources (Iyelolu et al., 2024; Wolf-Brenner et al., 2024; Timan et al., 2021).

Over half of European SMEs have reported they view regulation and administrative obligations among top challenges to their business (EC, 2020). Regulatory sandbox has been proposed as a potential solution because they encourage mutual learning between regulators and AI providers (Moraes, 2024). In addition to reducing regulatory uncertainty, sandboxes could also aid SMEs by connecting them with financiers and funding sources (Muminova et al., 2024). Regulatory sandboxes include valuable information about regulatory risks and requirements, too (OECD, 2023.)

However, for SMEs reap these benefits, they must be made aware of their potential value, and information about how they are to be provided. Indeed, there is a risk that SMEs may lack knowledge of regulation and an understanding of how related sandbox processes work, which could deter participation.

In turn, if sandboxes are to fulfil their proposed value, the issue of sandbox capacity should be solved. Indeed, there is concern stemming from the potential limited capacity of sandboxes to accommodate the volume of potential users. There are over 20 million entrepreneurs in Europe and even if 10% are involved in AI development or application, the scale of demand would be significant. These potential numbers highlight the need for the sandbox process to be streamlined, scalable, and eventually digitalized to ensure the efficient handling of potential participation.

As such, an AI regulatory sandbox should be enhanced with analysis and testing capabilities. It should thus be designed to support interaction among internal and external stakeholders, while addressing the capacity challenges, potentially through a service desk and process automation. With the previously outlined considerations these bring to light three additional design requirements for an AI regulatory sandbox: simplicity, transparency, and efficiency.

Finally, it is important to note that the AIA mandates compliance evaluation (for high-risk AI Systems), but sandbox participation is voluntary. We argue this creates potential unintentional risk exposure for SMEs which may lack the resources to fully identify and address compliance issues outside the sandbox. Meanwhile, large companies can avoid this risk internally through established corporate governance structures and resources. They have additional incentives to invest in compliance for fear of negative reputation and sanctions (Birkstedt et al., 2024) without the need for sandboxes. Therefore, to attract and serve less resource-rich SMEs, sandboxes must offer similar value, such as guidance, compliance tools, access to funding and connections to support initiatives.

4 Baseline Sandbox Process as a Stage-Gate -Model

To support the practical implementation of AI regulatory sandboxes, we propose a simple baseline process. This effort is motivated by the need to avoid the common pitfalls identified in above. To define the scope, the borders, interfaces and design the functions of AI Sandbox, the following design requirements must be addressed. These will be validated against our ongoing EU-wide survey results:

1. Eligibility Criteria: clear and consistent criteria are needed to determine which AI systems, providers, and deployers can participate in the sandbox. This has been a challenge in past sandbox implementations (Dardykina, 2024; von Steffen, 2025).
2. Integration with innovation ecosystems: the sandboxes should be connected with the innovation ecosystems to assist in peer support, expert advice and funding access (Draghi, 2024; Muminova et al., 2024).
3. Market access facilitation: the sandbox should support viable AI systems reach the market by providing expert validation and assistance in obtaining the necessary documentation for conformity registration and assessment (AIA objective).
4. Regulatory Engagement: Guidance from and interactions with National Supervisory (NSAs) and Competent Authorities (NCAs) to comply with the AIA and related regulations. (AIA requirement for the competent authorities).
5. Re-entry Criteria: criteria must be established for rejected or suspended AI systems to re-enter the compliance evaluation processes after necessary revisions (our on-going discussion with authorities and providers).
6. Iterative flexibility: guidelines should specify when and how the original AI system can be altered during sandbox participation. (to be discussed with the authorities, consultants and providers).
7. Procedural aspects: the sandbox process should be transparent, predictable, simple, cost-efficient with the ultimate aim to reduce time to market (conclusions by the researchers).

As the above requirements are intertwined, we identify the need is to provide initial guidelines and transparent steps for AI providers. Just as importantly, the sandbox should be communicated and developed as a shared artefact so that EU innovation ecosystem stakeholders, NCAs, NSAs, industry and domain experts, AI developers and deployers can build a shared understanding to assess and articulate how sandboxes can facilitate the journey from concept to candidate AI system. This is done through compliance evaluation, leading finally to the successful registration and market entry of CE-marked, trustworthy and innovative AI System for the market.

4.1 Getting prepared for sandbox

The process of an AI System's sandbox participation starts with self-evaluation by the AI provider leading towards entry gate. This involves documenting whether the AI system meets the practices in their intended context environment, i.e. its use cases and whether it initially complies with AI Act. If there is uncertainty regarding compliance, an AI system becomes a candidate for evaluation, and is lead towards industry specific, domain specific, or general evaluation of compliance (see fig. 1). The first stage is therefore to identify the AI system's core use case and determine the scope of its intended application.

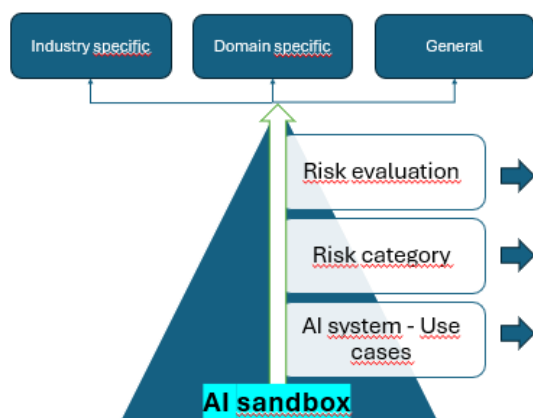


Figure 1: AI Sandbox process builds up from the bottom

Use cases can be assessed along two key dimensions:

- *Number of AI models/systems*: The number of AI models/systems deployed significantly influences the effort required to evaluate compliance. Different models within the AI system may fall into distinct risk categories, and when multiple models are used, each may require distinct evaluation methods (e.g. OECD's evaluation tools).
- *Use Across Domains /Industries*: Deploying an AI System across different domains or industries increases the need for regulatory scrutiny for each domain/industry. Each domain may be governed by sector-specific rules and standards. Our estimations suggest applying an AI system across multiple domains could double the effort required for compliance due to

the need for replicated evaluation processes. These two sets of use cases span the following matrix.

The initial division is explained in Appendix 1. It is of use in estimating the readiness and effort required for successful compliance evaluation. Thus, whenever a system involves multiple risk categories it should undergo scrutiny through participation in the AI sandbox.

4.2 The process

The actual sandbox process begins with the registration of the AI-system candidate (e.g., EU unique *AI System identifier*, *case process #*, *NCA-id*, and *applicant id*) and documentation for classification and related documentation (data sets, GDPR-compliance, AI model in use etc., see e.g. guidelines in (OECD, 2022)). This information should be made publicly available or accessible among NCAs for two reasons. First, to prevent forum shopping/redundant sandboxing processes. Secondly, it signals the AI System as in the evaluation process (Zarra, 2025) for changed risk status, exit reports/declaration of conformity to enter to market - or halted evaluation.

The eligibility for sandbox entry depends on the AI system's maturity and its risk classification. Since the AIA refers to practices and use cases the AI system should be assessed early to ensure whether it does not contradict with safety, livelihoods and rights of people making it *unacceptable to the market* independently of its use case. If the AI System passes this first stage, it should be evaluated for risk classification according to documentation requested by authorities (based on rules of conduct of AI Office) in the AI system's use case(s). If the eligibility criteria (proper documentation, and description of use cases) are met, AI Systems can be undergoing further scrutiny according to its preliminary risk category. If the AI system is *high risk*, it is subject to strict obligations including risk assessment and mitigation, quality of datasets, data logs, access to compliance assessment information, human oversight across all use cases. If the AI system is in the *limited risk category*, it is subject to disclosure obligations across all use cases. These initial stages (referred to as Gates 0 and 1) are depicted in the Stage-Gate model process flow (Fig. 2, in Appendix 2).

Accurate risk classification often requires further evaluation of the use cases, which may involve consultation with other sectoral authorities to supplement NCAs guidance (e.g. privacy, etc.). At this stage – Gate 2 it is likely necessary to engage with the innovation ecosystem to align its development with regulatory compliance (Janssen et al., 2022). If the use case specific evaluation reveals industry, or domain specific concerns, a more granular evaluation of the AI System would additionally be required (Gate 3). The specific use case(s) provide the basis for assessing the AI system’s data, model and outputs, particularly focussing on their content and quality to pass Gates four and five. Special attention must be given to the full data lifecycle, intellectual property considerations, and the integration of data from various repositories used to generate the system’s outputs. These stages may reveal potential non-compliant features in the candidate AI-systems. In such cases, we see there two potential remedies: first, introducing value-added services from the innovation ecosystem/consultants to address deficiencies in the data or model (Pagallo et al. 2019), or secondly, allowing some systemic uncertainty (Novelli et al., 2024) that can later be simulated in testing environments (Gate 6).

The goal of our tentative sandbox process is to ensure only a minority of AI system candidates will proceed to the final technical testing environment. Most evaluations should occur in the early stages of the sandbox through iterative learning and feedback (see Fig. 2, Appendix 2) to meet both the eligibility and compliance criteria outlined in the AIA. Only a smaller proportion of AI solutions are expected to progress to phase 2, which requires a more detailed risk analysis. At that stage, innovation ecosystem stakeholders can provide legal, technical and business support to help revise and improve the candidate AI system. While the final risk classification of the AI system will ultimately depend on the specific use case(s), quality of the data and models, the likelihood of achieving proof of compliance increases through the support from the innovation ecosystem and accumulated regulatory insights from NSAs and NCAs.

5 Considerations for Further Development

The proposed candidate AI systems are assisted through the process in collaboration with the existing EU AI innovation ecosystem. This process can be supplemented with mechanisms and tools for ethical and trustworthy AI assessment, many of

which are already available on international platforms such as OECD.AI⁶ The tools can be adapted to align with the to the basic process stages of our framework.

To support the simultaneous and practical implementation of those two objectives, regulators could introduce a dedicated service desk. This service desk would manage sandbox operations but also serve as a function for the accumulation and dissemination of regulatory knowledge, which is important for ensuring the interoperability of sandboxes across member states. Also, the final report from sandbox process should be compatible between member state sandboxes and preferably utilise (Fig 2., Appendix 2) certificate levels to reflect how far a candidate AI system has progressed through the sandbox. These levels can serve as indicators during funding rounds, helping stakeholders assess the system's regulatory maturity and readiness.

6 Conclusion

In sum, this paper presents a baseline process for evaluating AI system compliance with the EU AIA, addressing critical gaps in the practical implementation of AI regulatory sandboxes, particularly for SMEs and start-ups that face challenges in navigating complex regulatory environments. The proposed sandbox process also addresses the "opacity problem" of AI systems previously outlined in Chapter 3 by systematically examining both data and algorithmic components within specific use contexts with stakeholders. By categorizing AI systems according to deployment complexity (single vs. multiple models) and application scope (single domain vs. cross-domain), we enable more accurate risk assessment and tailored regulatory scrutiny.

Furthermore, by extending the process baseline with its stage-gates to AI Regulatory Sandbox Framework (AIRSF) that balances dual objectives: enabling AI innovation while ensuring compliance with risk-based regulatory requirements. The AIRSF is designed to be linked directly with innovation ecosystem stakeholders and services, offering a structured pathway from early AI system development to market entry through clearly defined steps towards regulatory compliance and legal certainty. As a result, our framework overcomes several limitations identified in previous sandbox implementations by providing clear eligibility criteria, transparent evaluation steps, and graduated certification levels. The baseline process is designed to scale

efficiently, with most AI systems expected to reach to market readiness through early-stage support, while only complex or high-risk systems progress to a more comprehensive evaluation in later stages utilizing also the innovation ecosystems to full potential

The AIRSF can be supported, managed and utilised for regulatory learning through service desk-style mechanisms, but the final design is subject to further iterations with stakeholders in the design scientific sense and its actual implementation in the sense of action research by the actors. This will take a lot of time and effort to become generally deployable in all EU member states, being its main limitation.

Acknowledgements

The authors wish to thank Fabio Seferi, Arto Lanamäki, and Karin Väyrynen for their insightful and constructive comments.

References

- van Aken J.E., and Romme G., (2009). Reinventing the future: adding design science to the repertoire of organization and management studies, *Organization Management Journal*, 6:1, 5-12. <http://dx.doi.org/10.1057/omj.2009.1>
- Ardito, L., Filieri, R., Raguseo, E. and Vitari, C. (2024). Artificial intelligence adoption and revenue growth in European SMEs: synergies with IoT and big data analytics. *Internet Research*, Vol. ahead-of-print No. ahead-of-print. <https://doi-org.ezproxy.utu.fi/10.1108/INTR-02-2024-0195>.
- Bagni F., and Seferi F., eds., (2025). *Regulatory Sandboxes for AI and Cybersecurity. Questions and answers for stakeholders*. Cybersecurity National Lab in collaboration with Security and Rights in the Cyberspace. White Paper edited by Bagni and Seferi. ISBN 9788894137378, February 2025. 231 pages. <https://www.aigl.blog/content/files/2025/04/Regulatory-Sandboxes-for-AI-and-Cybersecurity.pdf>
- Birkstedt T., Minkinen M., Tandon A. & Mäntymäki M. (2023) AI Governance: themes, knowledge gaps and future agendas. *Internet Research*, Vol 33 (7). <https://doi.org/10.1108/INTR-01-2022-0042>
- Bradford A., (2024). The False Choice Between Digital Regulation And Innovation. *Northwestern Law University Review*. Vol. 119, No. 2., pp. 377-454. <https://scholarlycommons.law.northwestern.edu/nulr/vol119/iss2/3/>
- Buhmann A., and Fieseler C., (2024). Towards a deliberative framework for responsible innovation in artificial intelligence. *Technology in Society*, 64:2021 101475. <https://doi.org/10.1016/j.techsoc.2020.101475>
- Cooper R., (1990). Stage-Gate Systems : A New Tool for Managing New Products. *Business Horizons* 33(3):44-54. [https://doi.org/10.1016/0007-6813\(90\)90040-I](https://doi.org/10.1016/0007-6813(90)90040-I)
- Dardykina, A., (2024). Is There Enough Sand in The Sandbox? (October 01, 2024). forthcoming in the *European Business Law Review*, 2026, Available at SSRN: <https://ssrn.com/abstract=5126862> or <http://dx.doi.org/10.2139/ssrn.5126862>
- Díaz-Rodríguez N., Del Ser Javier., Coeckelbergh M., López de Prado M., Herrera-Viedma E., and Herrera F. (2024). Connecting the dots in trustworthy Artificial Intelligence: From AI

- principles, ethics, and key requirements to responsible AI systems and regulation. <https://doi.org/10.1016/j.inffus.2023.101896>
- Draghi M., et al. (2024). “The future of European competitiveness – A competitiveness strategy for Europe”. https://commission.europa.eu/topics/strengthening-european-competitiveness/eu-competitiveness-looking-ahead_en#paragraph_47059.
- EC, (2020). Survey confirms the need to support small and medium-sized businesses on their path towards digitalisation and sustainability <https://ec.europa.eu/newsroom/growth/items/688053/en>
- Iivari J., and Venables J.R. (2009). Action Research and Design Science Research – Seemingly similar but decisively dissimilar. In the Proceedings of 17th European Conference on Information Systems (ECIS 2009, Verona, Italy, <https://aisel.aisnet.org/ecis2009/73>.
- Janssen, H., Seng Ah Lee, M. & Singh, J. (2022). International Journal of Law and Information, vol. 30 (2), pp: 200-232, DOI: <https://doi.org/10.1093/ijlit/eaac018>
- Iyelolu, T., V., Agu, E., E., Idemudia, C. & Ijomah, T., I. (2024). Driving SME innovation with AI solutions: overcoming adoption barriers and future growth opportunities. International Journal of Science and Technology Research Archive, 07(01), 36–54. <https://doi.org/10.53771/ijstra.2024.7.1.0055>.
- Lanamäki A., Väyrynen K., Vainionpää F., (2024). The European Union's Regulatory Challenge: Conceptualizing Purpose in Artificial Intelligence. In the Proceedings of the Thirty Second European Conference on Information Systems (ECIS 2024), Paphos, Cyprus. https://aisel.aisnet.org/ecis2024/track04_impactai/track04_impactai/1
- Longo E., and Bagni F., (2025). From Legal Experimentation to Regulatory Sandboxes: The EU's Pioneering Approach to Digital Innovation and Regulation. In Bagni and Seferi (eds.), pp 18-28.
- Moraes, T. (2024). Ethical AI Regulatory Sandboxes: Insights from cyberspace regulation and Internet governance. In Proceedings of the Second International Symposium on Trustworthy Autonomous Systems (TAS '24). Association for Computing Machinery, New York, NY, USA, Article 17, pp. 1–10. <https://doi.org/10.1145/3686038.3686049>
- Muminova, E., Ashurov, M., Akhunova, S., and Turgunov, M. (2024). AI in Small and Medium Enterprises: Assessing the Barriers, Benefits, and Socioeconomic Impacts. 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS), Chikkaballapur, India, pp. 1-6, <https://doi.org/10.1109/ICKECS61492.2024.10616816>
- Novelli C., Casolari F., Rotolo A., Taddeo M., and Floridi L. (2024). AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act. Digital Society (2024) 3:13. <https://doi.org/10.1007/s44206-024-00095-1>
- OECD, (2022). OECD Framework For the Classification of AI Systems. OECD Digital Economy Papers, No. 323. OECD Publishing, February, 2022). 80p. Available at https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/02/oecd-framework-for-the-classification-of-ai-systems_336a8b57/cb6d9eca-en.pdf
- OECD, (2023). Regulatory sandboxes in artificial intelligence. Available at https://www.oecd.org/en/publications/regulatory-sandboxes-in-artificial-intelligence_8f80a0e6-en.html
- Orwat, C., Bareis, J., Folberth, A., Jahnel, J., and Wadephul, C., (2024) Normative Challenges of Risk Regulation of Artificial Intelligence. Nanoethics Vol. 18 (11). <https://doi.org/10.1007/s11569-024-00454-9>.
- Pagallo U., Casanovas P., and Madelin R., (2019) The middle-out approach: assessing models of legal governance in data protection, artificial intelligence, and the Web of Data, The Theory and Practice of Legislation, 7:1, 1-25, <https://doi.org/10.1080/20508840.2019.1664543>
- Papagiannidis E., Mikalef P., and Conboy K. (2025). Responsible artificial intelligence governance: A review and research framework. The Journal of Strategic Information Systems 34(2). <https://doi.org/10.1016/j.jsis.2024.101885>

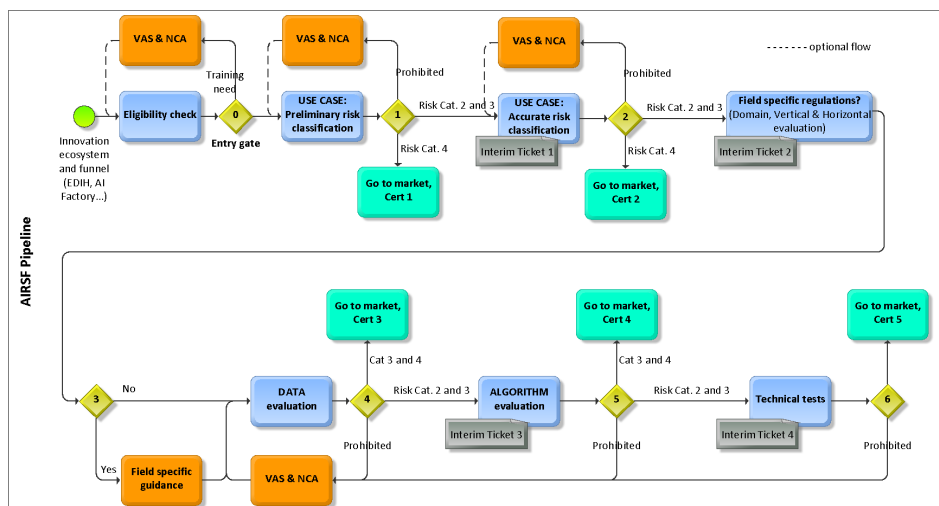
- Ranchordas S. (2021). Experimental Regulations and Regulatory Sandboxes: Law without Order? University of Groningen Faculty of Law Research Paper No. 10/2021, Available at SSRN: <https://ssrn.com/abstract=3934075> or <http://dx.doi.org/10.2139/ssrn.3934075>
- Seferi, F., (2025). A Comparative Analysis of Regulatory Sandboxes from Selected Use Cases: Insights from Recurring Operational Practices. In Bagni and Seferi (eds.), pp. 145-176.
- Sloane M., and Wüllhorst E., (2025). A systematic review of regulatory strategies and transparency mandates in AI regulation in Europe, the United States, and Canada. *Data & Policy*, 7: e11, <https://doi.org/10.1017/dap.2024.54>
- Timan, T., van Oirsouw, C., & Hoekstra, M. (2021). The role of data regulation in shaping AI: an overview of challenges and recommendations for SMEs. *The Elements of Big Data Value*, 355-376. https://doi.org/10.1007/978-3-030-68176-0_15
- Wolf-Brenner, C., Pammer-Schindler, V., & Breitfuss, G. 2024. How Do Professionals in SMEs Engage With AI and Regulation? An Interview Study in Austria. In *Proceedings of Mensch und Computer 2024 (MuC '24)*. Association for Computing Machinery, New York, NY, USA, 646–650. <https://doi.org/10.1145/3670653.3677514>
- Zarra A., (2025). Operationalizing AI Regulatory Sandboxes: A look at the incentives for Participating Start-Ups and SMEs beyond compliance. In Bagni and Seferi (eds.), pp. 101-115.
- Zetsche, D. A., Buckley R. P., Barberis J. N., Arner D. W. (2017). Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation. *Fordham Journal of Corporate & Financial Law*, Vol. 23 (1). <https://ir.lawnet.fordham.edu/jcfl/vol23/iss1/2/>

Appendix 1: Categories of AI Systems' use cases

Based on this, we categorise AI systems into the following types:

- **Embedded Single-Purpose AI System:** These involve a single AI model evaluated against a specific business purpose. For example, an AI tool that analyses item quantities. The evaluation within the AI sandbox is straightforward due to the clear boundaries of the use case.
- **Single-Target AI Systems with Multiple Functions:** These AI systems use several AI models to fulfil a specific business need. An example is a traffic monitoring solution that identifies traffic volume (AI model 1) and predicts future traffic based on historical data (AI model 2). The AI sandbox must analyse the compliance of both models as a single AI system against the overarching purpose.
- **General AI Solutions with Single Function:** These systems are more complex due to the open-ended potential of the use case. For example, an object identification system for video analytics may serve functions as diverse as identifying car brands or individuals. The lack of a specific focus complicates the creation of relevant test data and the compliance analysis of an AI System.
- **General AI Solutions with Multiple Functions:** This category includes general-purpose AI (GPAI) systems like ChatGPT⁷, which involve multiple layers of AI models and systems. For the scope of this paper, we exclude GPAI due to the some practical (e.g. the decisive role of AI Office) and fundamental issues around ambiguity of the AIA in relation to GPAI regulation (Lanamäki et al., 2024; Papagiannidis et al., 2025).

Appendix 2: Tentative Stage-Gate –process



Key:

- Innovation ecosystem and funnel (EDIHs, TEFs, AI Factories...) entry to eligibility check
- NCA = National Competent Authority (also Supervisory Authority)
- VAS = Value Added Service (e.g. Innovation Ecosystem (accredited) experts in AI, business, AI and related regulation.
- #0...#6 Stage Gates
- Prohibited = Risk Cat #1
- Risk Cat #2...#4 AI Act risky practices categories
- Cert #1 ... Cert #5 Registering of high-risk AI Systems to EU-database, granting certificates for Risk Cat #2 and #3.
- Interim Ticket #1...#4 For indicating the status of the AI System progress in the process, for performance assessment of the sandbox throughput and signal providers, financiers and potential deployers. Also useful, when switching from sandbox to another.

DATA-DRIVEN ECOSYSTEM BUSINESS MODELS IN AGRICULTURE WITH FOCUS ON SUSTAINABILITY: A SYSTEMATIC LITERATURE REVIEW

SHAHIN REZVANIAN

University of St. Gallen, Institute for Media and Communication Management,
St. Gallen, Switzerland
shahin.rezvanian@unisg.ch

Digital transformation is reshaping agriculture through data-driven business models leveraging emerging technologies. Understanding these models' sustainability contributions is crucial given agriculture's challenges with climate change, resource constraints, and food security. Following PRISMA, Scopus and Web of Science were searched, yielding 1538 articles. After screenings, 80 papers were analyzed thematically. 32 distinct data-driven ecosystem business models were identified, categorized into three primary groups: Technology-Focused Models, Value Chain Integration Models, and Data & Governance Models. These models contribute to economic sustainability through resource optimization and new revenue streams; environmental sustainability through precision management and emissions reduction; and social sustainability through knowledge sharing and community development. Implementation challenges include technical integration, organizational adoption barriers, data governance concerns, and policy gaps. These models show significant potential for enhancing agricultural sustainability. Trust emerges as fundamental for implementation, while power dynamics remain critical concerns. Future research should focus on governance frameworks, user-centric design, and impact assessment.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.41](https://doi.org/10.18690/um.fov.4.2025.41)

ISBN
978-961-286-998-4

Keywords:
digital transformation,
sustainability,
data-driven agriculture,
ecosystem business models,
value chains



University of Maribor Press

1 Introduction

Digital transformation is reshaping agriculture through data-driven business models that leverage emerging technologies to create value within ecosystem—driving the twin transitions of digitalization and sustainability. Despite growing research on digital agriculture, there remains limited synthesis regarding how specific business models leverage data to enhance sustainability. This review addresses that gap by examining how data-driven ecosystem business models (DDEBMs) contribute to economic, environmental, and social sustainability, while identifying implementation challenges and future directions. DDEBMs in agriculture can be defined as interconnected value creation systems that utilize digital technologies, data analytics, and collaborative networks to generate, process, and exchange agricultural data for economic, environmental, and social benefits. These models typically involve multiple stakeholders working together through platforms or networks to create and capture value from agricultural data (Vial & Tedder, 2017; Graf-Drasch et al., 2023; Rijanto, 2021; Gowri & Ramachander, 2024).

Sustainability—meeting present needs without compromising the ability of future generations to meet their own (Brundtland Commission, 1987)—is critical as agriculture faces mounting challenges from climate change, resource constraints, and food security. Agricultural output must increase 70% by 2050 despite shrinking farmland and declining workforce (Cavazza et al., 2023). Digital agriculture offers a path forward by enabling informed, data-driven decisions that address sustainability across value chains (Hrustek, 2020). This review addresses three research questions: What types of data-driven ecosystem business models exist in agriculture? How do these models contribute to sustainability (environmental, social, or economic)? What are the key challenges, limitations, and future directions for implementing these models? The goal is to synthesize current knowledge base on DDEBMs, identify and categorize model types, assess their contributions to various sustainability dimensions, and highlight implementation challenges and future research directions.

2 Methods

This systematic literature review followed PRISMA guidelines and searched Scopus and Web of Science up to January 2025, using terms related to digital technologies, business models, ecosystems, and agricultural contexts. The selection process

involved two stages. First, titles, abstracts, and keywords were screened using five criteria: digital technologies, business models, ecosystems, agriculture, and sustainability. Studies meeting at least four criteria advanced to full-text review. Second, full texts were evaluated for relevance to the research questions, with those scoring at least four out of five included, resulting in 80 papers. To ensure coding reliability, the results were triangulated by independently coding 20 papers (25% of the sample), reaching 81% agreement. Discrepancies were resolved through discussion before final coding. Figure 1 illustrates the review process, including identification, screening, eligibility, and inclusion.

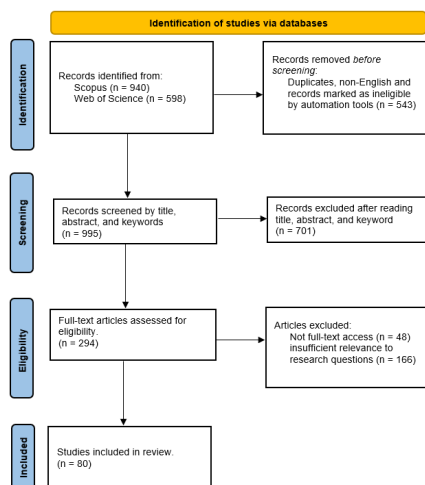


Figure 1: PRISMA Flow Diagram

Source: Own

Data were extracted on publication metadata, technologies, business model types, sustainability impacts, implementation challenges, and research directions. Business model categorization combined deductive and inductive approaches. Deductive coding used established frameworks such as Osterwalder's Business Model Canvas and the Triple Bottom Line (Elkington, 1997), while inductive analysis identified emergent value creation patterns through open coding, cross-case comparison, and iterative refinement. Sustainability impacts were assessed using the triple-bottom-line framework—economic, environmental, and social—and rated as strong, moderate, or limited based on explicit statements, quantitative evidence, frequency of mention, and comparative emphasis across studies.

3 Results

The literature on data-driven ecosystem business models in agriculture shows growing interest, with 72.5% of publications appearing between 2020-2025, reflecting the accelerating twin transition in the sector (Figure 2).

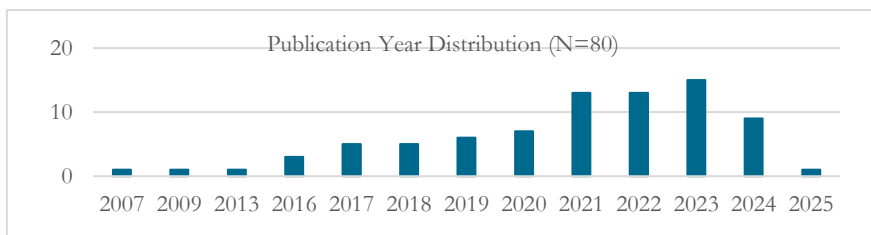


Figure 2: Temporal distribution of publications on data-driven ecosystem business models in agriculture (2007-2025)

Source: Own

Studies are geographically diverse, with major contributions from Asia (25%), Europe (25%), and Africa (20%), especially India, China, Indonesia, and various European countries (Figure 3). Developed economies often emphasize advanced digital infrastructures and regulation, while developing regions focus on smallholder inclusion and mobile-based solutions.

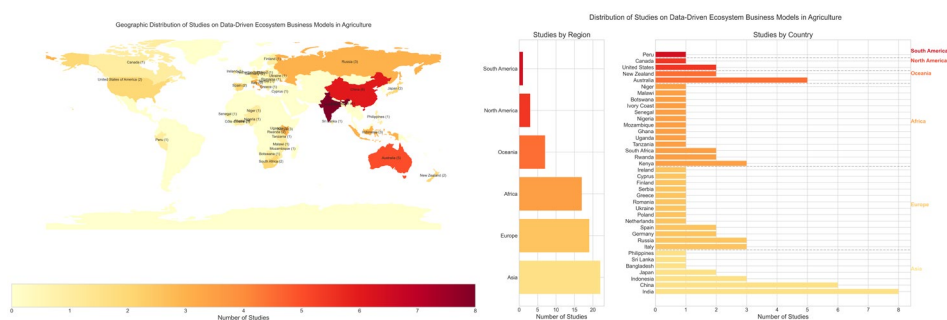


Figure 3: Geographical Distribution of Studies

Source: Own

Empirical studies dominate (55%), suggesting the field remains exploratory. From a technology perspective, IoT and sensor systems are most common (29.2%), followed by AI (17.2%), data analytics (16.7%), and blockchain (11.2%). Environmental sustainability is addressed most frequently (82%), with climate action and resource efficiency as dominant themes. Economic sustainability appears in 70% of studies, and social sustainability in 55%. Key implementation challenges include organizational barriers (37%), technical issues (29.5%), knowledge gaps (13%), financial constraints (11%), and data governance concerns (9.5%). The urban–rural digital divide is a cross-cutting issue.

3.1 Typology of Data-Driven Ecosystem Business Models in Agriculture:

Our analysis identified 32 distinct data-driven ecosystem business models in agriculture. These models were classified into three overarching categories (Figure 4): Technology-Focused, Value Chain Integration, and Data & Governance Models.

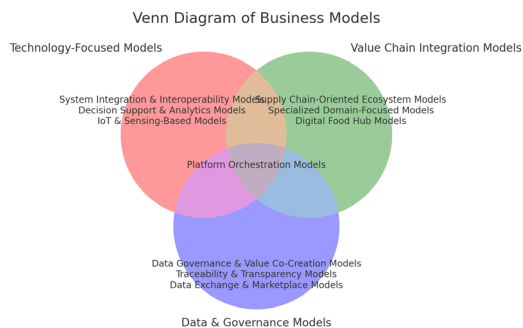


Figure 4: Venn Diagram of Data-Driven Ecosystem Business Model Categories

Source: Own

The classification was developed through an iterative coding process combining deductive and inductive reasoning. Deductively, the analysis drew on established frameworks such as Osterwalder’s Business Model Canvas and the Triple Bottom Line. Inductively, emergent value creation patterns were identified by analyzing model characteristics, technologies, stakeholder configurations, and value mechanisms. Several models span multiple categories. At the center of the typology are Platform Orchestration Models, which integrate technological infrastructure, value chain coordination, and data governance.

Technology-Focused Models emphasize technological solutions and infrastructure as their primary value proposition, organized into three sub-categories: **IoT & Sensing-Based Models** leverage Internet of Things technologies for data collection and processing, creating digital counterparts of physical entities. Examples include IoT-enabled livestock management (Alves et al., 2021), IoT-blockchain field integration through layered architectures (Tasic & Cano, 2024), and autonomous robot swarms for precision farming (Braun et al., 2018). **Decision Support & Analytics Models** process data to generate actionable insights, exemplified by platforms that analyze data to provide decision support (Kampker et al., 2018), systems minimizing greenhouse gas emissions in livestock (Bălănescu et al., 2020), and AI applications in food sorting, quality control, and vertical farming (Di Vaio et al., 2020; Cavazza et al., 2023). **System Integration Models** focus on connecting diverse agricultural systems through common semantic data models (Brewster et al., 2017), multi-partner interoperability frameworks (Huber & Markward, 2021), and administrative burden reduction systems (Pope et al., 2021).

Value Chain Integration Models connect different components of the agricultural supply chain: **Supply Chain-Oriented Models** reshape operational decisions through Big Data (Issa et al., 2024), establish multi-level collaboration frameworks (Braun et al., 2018), and create closed vertical networks with high ICT investment, as exemplified by ITC's eChoupal in India (Rao, 2007). **Digital Food Hub Models** function as "intermediary organizations based on an innovative digital strategy followed by small farms forming cooperation networks" (Berti et al., 2018, p. 427), creating shared value through quality, sustainability, and locality differentiation. **Specialized Domain-Focused Models** address specific sectors through digital livestock tools (Daum et al., 2021), bundled climate information services (Kagabo et al., 2025), holistic digital agriculture frameworks (Cook et al., 2021), and integrated renewable energy-agriculture systems (Hu et al., 2022).

Data & Governance Models focus on data exchange, management, and governance issues: **Data Exchange & Marketplace Models** enable farmers to share data for collective value creation (Vial & Tedder, 2017) and establish data broker platforms that aggregate agricultural data for external customers or create comprehensive marketplaces (Kampker et al., 2018). **Traceability & Transparency Models** leverage blockchain for wine supply chain integration (Malisic et al., 2023), connect sustainability certification with market access in sugar production (Kealley

et al., 2022), track products from production to consumption (Kosior & Mlodawska, 2024), and build on cooperative structures for intelligent traceability (Giagnocavo et al., 2017). **Data Governance & Value Co-Creation Models** identify distinct ecosystem roles in value creation (Azkan et al., 2022), emphasize fairness, accountability, and transparency principles (Stitzlein et al., 2021), enable progressive trust development between farmers and buyers (Kumarathunga & Ginige, 2022), and propose precompetitive platforms with standardized vocabularies (Holden et al., 2018).

Platform Orchestration Models operate at the intersection of all three domains, integrating technological capabilities, value chain coordination, and data governance. They include comprehensive frameworks connecting multiple stakeholders (Gebresenbet et al., 2023), complex architectures with device, network, and application layers (Grabher, 2020), and digital orchestrators like AgriCircle that bring together diverse stakeholders while specializing in modeling and decision support (Huber & Markward, 2021). For organizational purposes in the remainder of this paper, these cross-cutting models will be grouped under the Value Chain Integration Models category. Having established this typology, the following section examines how these business models contribute to economic, environmental, and social sustainability dimensions.

3.2 Sustainability Impacts of Data-Driven Ecosystem Business Models

To provide a comprehensive overview of how each business model type contributes to different sustainability dimensions, a matrix framework developed that maps the 32 identified business models against their economic, environmental, and social sustainability impacts (Figure 5). The matrix reveals several important patterns regarding how digital business models contribute to sustainability: **1) Strong Economic Focus:** Most business models demonstrate strong economic impacts, particularly Platform Orchestration Models, Data Exchange Models, and Supply Chain Integration Models; **2) Varying Environmental Impact:** Environmental sustainability shows greater variation, with the strongest contributions from IoT-Blockchain Integrated Systems, Renewable Energy Agricultural Integration, and Climate Information Services Models; **3) Less Pronounced Social Dimension:** Social sustainability received comparatively less emphasis in the literature, though several models demonstrate strong social impacts, including Cooperative-Based

Traceability Systems, Trust-Based Big Data Ecosystems, and Digital Food Hubs; **4) Balanced Models:** Some models show relatively balanced contributions across all three dimensions, particularly Platform Orchestration Models, Sustainability Credential Systems, and System-of-Systems Architecture Models; **5) Specialized Models:** Other models demonstrate specialized sustainability profiles, such as Trust-Based Blockchain Market Models (strong emphasis on economic and social but limited emphasis on environmental). This matrix highlights how the twin transition of digital and sustainability transformations manifests in agriculture.

Business Model			Chain	Economic	Environ- mental	Social	Tech	Key Citations
A. TECHNOLOGY- FOCUSED MODELS	A.1 IoT & Sensing-Based	A.1.1 Smart Livestock Management Systems	🐄	...	..	..	B,H	Alves et al. (2021); Issa et al. (2024)
		A.1.2 IoT-Blockchain Integrated Systems	📡	..	...	..	A,B	Tasic & Cano (2024)
		A.1.3 Autonomous Agricultural Systems	🚀	...	...	..	B,G	Braun et al. (2018)
	A.2 Decision Support & Integratio	A.2.1 Farm-Level Decision Support Systems	🧠	...	...	..	C	Bălănescu et al. (2020)
		A.2.2 AI-Driven Agricultural Models	🤖	...	..	..	C	Cavazza et al. (2023); Di Vaio et al. (2020)
		A.2.3 Collaborative Decision Support Frameworks	📡	..	...	...	C	Beaumont De Oliveira et al. (2021)
	A.3 System Integratio	A.3.1 System-of-Systems Architecture Models	📡	..	..	..	E	Brewster et al. (2017)
		A.3.2 Interoperability-Focused Ecosystems	📡	...	..	..	E	Huber & Markward (2021)
		A.3.3 Administrative Burden Reduction Systems	📡	...	..	..	E	Poppe et al. (2021)
B. VALUE CHAIN INTEGRATION MODELS	B.1 Platform Orchestrat	B.1.1 Integrated Technology Platforms	📡	...	...	..	B,C,E	Grabher (2020); Gebresenbet et al.(2023)
		B.1.2 Ecosystem Orchestrators	📡	...	...	..	D,E	Huber & Markward (2021)
		B.1.3 Digital Agriculture Value Platforms	📡	...	..	..	D,E	Cook et al. (2021)
	B.2 Supply Chain-Oriented Ecosystem	B.2.1 Big Data-Driven Supply Chain Models	📡	...	..	..	D	Issa et al. (2024)
		B.2.2 Blockchain-Based Supply Chain Models	📡	..	...	..	A	Malisic et al. (2023)
		B.2.3 Multi-Level Collaborative Supply Chain Models	📡	...	..	..	B	Braun et al. (2018)
	B.3 Specialized Domain-Focused	B.2.4 Closed Vertical Supply Chain Networks	📡	...	..	..	R	Rao (2007)
		B.3.1 Digital Livestock Management Models	🐄	...	..	..	F	Daum et al. (2021)
		B.3.2 Climate Information & Smart Agriculture Models	🌱	..	...	..	H	Kagabo et al. (2025)
C. DATA & GOVERNANCE MODELS	C.1 Data Exchange & Marketpla	B.3.3 Vertical Farming Decision Support Systems	📡	...	...	..	C	Beaumont De Oliveira et al. (2021)
		B.3.4 Renewable Energy Agricultural Integration	🌞	...	...	..	H	Hu et al. (2022)
		C.1.1 Data Exchange Communities	📡	...	..	...	D	Vial & Tedder (2017)
		C.1.2 Data Marketplaces	📡	...	..	..	D	Kampker et al. (2018)
	C.2 Traceability & Transpare	C.1.3 Digital Food Hubs	📡	...	..	..	F	Berti et al. (2018)
		C.1.4 B2B Agricultural E-Marketplaces	📡	...	..	..	F	Zheng et al. (2009)
		C.2.1 Sustainability Credential Systems	📡	..	...	..	A	Kcalley et al. (2022)
		C.2.2 Digital Food Passport Systems	📡	..	..	..	H	Kosior & Młodawska (2024)
	C.3 Data Governance & Value Co-	C.2.3 Cooperative-Based Traceability Systems	📡	..	...	..	G	Giagnocavo et al. (2017)
		C.2.4 Trust-Based Blockchain Market Models	📡	...	..	...	A	Kumarathunga & Ginige (2022)
		C.3.1 Data Ecosystem Value Co-Creation Models	📡	...	..	..	D	Azkan et al. (2022)
		C.3.2 Trust-Based Big Data Ecosystems	📡	..	..	..	D	Stitzlein et al. (2021)
	C.3.3 Data Space and Transparency Models	C.3.3 Data Space and Transparency Models	📡	..	..	..	A	Klug & Prinz (2023)
		C.3.4 Internet of Food Ecosystem Models	📡	...	...	..	D	Holden et al. (2018)

Legend: Sustainability Impact: Strong emphasis, Moderate emphasis, Limited emphasis (based on relative emphasis and evidence in the literature); Value Chain Position: 🐄 Production, 🤖 Processing, 🚚 Distribution, 🏪 Retail, 📡 Cross-value chain; Technologies: A:Blockchain, B:IoT, C:AI/ML, D:Big Data, E:Cloud, F:Mobile, G:Drones, H:Edge Computing.

Figure 5: Matrix Framework of Data-Driven Ecosystem Business Models and Their Sustainability Impacts

Source: Own

Data-driven ecosystem business models contribute to economic sustainability through multiple pathways: Improved access to financing via blockchain platforms connecting farmers with financial institutions (Rijanto, 2021); Resource optimization enabling more precise resource use and cost reduction (Vial & Tedder, 2017); New revenue streams through data monetization platforms (Kampker et al., 2019); Enhanced decision-making efficiency, illustrated by disease detection two years earlier than human discovery in the Höcklistein vineyard case (Huber & Markward, 2021); Increased yields, demonstrated by a 5% yield increase per hectare using digital twins for potato harvesting (€560 additional revenue/hectare) (Kampker et al., 2019); Diversified revenue sources through integrated systems, exemplified by a 200 MW PV fishery project with a net present value of ¥352.1253 million (Hu et al., 2022).

The digital transformation enables environmental benefits through: Water reduction, exemplified by rice irrigation systems reducing consumption by 10-25% (Routis et al., 2022); Emissions reduction, with a 200 MW digital PV fishery project reducing emissions by approximately 119,241 tons annually (Hu et al., 2022); Early intervention capabilities enabling targeted rather than widespread interventions (Huber & Markward, 2021); Waste reduction through B2B marketplaces addressing the 20 billion pounds of "ugly" produce lost annually in the US (Vlachopoulou et al., 2021); Sustainable land use through integrated systems enhancing land usage by 32.2% compared to traditional approaches (Hu et al., 2022); Circular economy approaches establishing closed-loop farming ecosystems (Abdillah et al., 2023).

Social sustainability is addressed through: Empowerment of small-scale producers, with digital food hubs positioning farmers as "price negotiators" rather than "price takers" (Berti et al., 2018); Knowledge sharing facilitating exchange across stakeholder groups (Routis et al., 2022); Rural development through digital technologies (Hrustek, 2020) and cooperatives functioning as both social and economic networks (Giagnocavo et al., 2017); Social capital formation through blockchain-enabled trust networks (Kumarathunga & Ginige, 2022); Gender inclusion with platforms helping female farmers achieve financial stability (Abdillah et al., 2023); Financial inclusion providing access to services in previously underserved rural areas (Abdillah et al., 2023).

This analysis demonstrates how the digital transition in agriculture enables multi-dimensional sustainability impacts, with the most substantial evidence for economic benefits, but significant potential across all sustainability dimensions.

3.3 Implementation Challenges

Implementing data-driven ecosystem business models for sustainability faces several challenges that can impede both digital transformation and sustainability goals: **Technical and Infrastructure Challenges** include: Interoperability issues between systems from different manufacturers (Huber & Markward, 2021; Routis et al., 2022); Digital infrastructure limitations affecting rural and developing regions, with Hansen et al. (2023, p. 526) noting that "poor network connectivity in Australia presents major barriers"; Data integration complexity when combining diverse agricultural and environmental data (Gebresenbet et al., 2023); System complexity in technical implementation and maintenance (Satya et al., 2021); Environmental conditions affecting hardware durability in agricultural settings with "moisture, dust, ammonia, and pests" (Neethirajan & Kemp, 2021); and lack of standardized vocabularies for data integration (Holden et al., 2018). **Organizational and Adoption Challenges** include: Significant skills and knowledge gaps among farmers and agricultural professionals (Kagabo et al., 2025; Di Vaio et al., 2020); Technology-driven rather than needs-driven innovation, with businesses "driven by technological advancements rather than providing tailor-made solutions to farmers" (Mahdad et al., 2022, p. 1865); Stakeholder diversity creating tension in developing shared value propositions; Business model transition difficulties as manufacturers struggle to shift from product-centric to solution-oriented approaches (Kampker et al., 2019); Fragmented implementation of technologies without comprehensive management systems; and Organizational support structures inadequate for maintaining complex digital systems (Cook et al., 2021). **Data Governance and Power Dynamic Challenges** include: Data ownership concerns with platform orchestrators collecting farming data while farmers have limited control (Grabher, 2020); increasing farmer dependency on technology; Trust issues regarding data use by competitors or governments (Rijswijk et al., 2019); Power asymmetries affecting farmer autonomy; Aftermarket lock-in where manufacturers maintain exclusive control over equipment-collected data (Atik & Martens, 2021); and fears about consumer misinterpretation of agricultural data (Kosior & Mlodawska, 2024). **Policy and Institutional Challenges** encompass: Regulatory gaps with insufficient

governance frameworks, as "Australia lags in providing appropriate regulation and governance frameworks for the sector" (Hansen et al., 2023, p. 530); Inadequate public sector support for digital food value frameworks (Cook et al., 2021); Administrative burdens from redundant reporting requirements (Kosior & Młodawska, 2024); Inter-firm collaboration barriers around economic incentives and strategic alignment (Zheng et al., 2009); Policy uncertainty affecting business models relying on subsidies (Hu et al., 2022); and weak quality standards limiting ecosystem functioning (Abdillah et al., 2023).

4 Discussion

This review highlights a fundamental shift from supply chain to ecosystem thinking in agricultural business models—reflecting how digital transformation enables sustainability transformation. As Mahdad et al. (2022, p. 1859) note, "the interdependencies among agri-food actors call for bringing in the innovation ecosystems perspective to replace the static supply chains perspective." This evolution aligns with Adner's (2017) Ecosystem Theory of Value Creation, which emphasizes coordinated alignment among complementary actors rather than merely optimizing individual supply chain elements. This evolution moves from precision farming and platform models toward integrated digital ecosystems delivering tailored solutions for farmers. Successful implementation requires balancing farmers' needs with market demands, as "the space for developing a collaborative and open business model is prepared" only when these align (Mahdad et al., 2022, p. 1857). This supports Stakeholder Theory principles (Freeman, 1984), demonstrating how technology-driven approaches that neglect key stakeholders lead to adoption reluctance. Trust is a foundational requirement for data-driven ecosystem business models. Fairness, accountability, and transparency (FAT) are essential principles for agricultural data ecosystems, with blockchain technologies showing potential to address power imbalances that have historically disadvantaged smallholder farmers—directly addressing the data governance challenges highlighted in the analysis. These observations resonate with Trust-Based Collaboration Theory (Bachmann & Inkpen, 2011), particularly in how blockchain-enabled mechanisms can facilitate collaboration where interpersonal trust may be limited. The matrix framework reveals distinct sustainability profiles: Technology-Focused Models excel at resource optimization, Value Chain Integration Models show balanced contributions, and Data & Governance Models emphasize social sustainability

through trust-building. This suggests comprehensive sustainability requires complementary business models, aligning with Triple Bottom Line principles (Elkington, 1997). While this study reflects literature-based patterns, empirical validation is still needed. Future research should apply Value Sensitive Design (Friedman, 1996; Van de Poel, 2020) to embed normative values like trust and fairness into ecosystem model development.

5 Conclusion

This systematic review examined data-driven ecosystem business models in agriculture with a focus on sustainability, identifying 32 distinct models across three categories: Technology-Focused, Value Chain Integration, and Data & Governance Models. These approaches demonstrate diverse applications of digital technologies that create agricultural value while advancing sustainability goals—exemplifying the twin transition of digitalization and sustainability. The models contribute to sustainability through multiple pathways, including resource efficiency, improved financing, environmental management, equitable value chain participation, and rural community development. However, implementation challenges persist—particularly around power asymmetries, data ownership, farmer autonomy, and equitable value distribution. Economic value remains the dominant adoption driver across studies. Recommendations for policymakers and practitioners include: (1) developing user-centric value propositions; (2) creating robust governance frameworks that promote equitable data use and trust-building; and (3) fostering public–private partnerships to address infrastructure gaps. Future research should: (1) build on this study’s framework to develop integrated sustainability assessment tools that are supported by interoperable data frameworks and applicable in real-world agricultural ecosystems; (2) explore user-centered design in relation to Global North–South power dynamics; and (3) empirically validate the conceptual patterns presented. The matrix framework (Figure 5) offers a practical tool for researchers and practitioners to evaluate and design business models addressing specific sustainability dimensions. Its relevance extends beyond agriculture to other sectors including healthcare, smart cities, energy transition, and financial services. Despite methodological limitations in identifying emerging models, this paper provides a foundation for understanding and implementing digital technologies for a more sustainable agricultural future, contributing to the broader goal of achieving twin digital and sustainability transitions across economic sectors.

Acknowledgements

This work was funded by the European Commission under the Doctoral Networks Programme (MSCA-DN-101073381-EnTrust) within the Horizon Europe (HORIZON) Marie Skłodowska-Curie Actions.

References

- Abdillah, A. F., Kusuma, C. S. D., Astuty, S., & Abdillah, M. L. W. (2023). The role and challenges of the food and agriculture digital platform ecosystem as driver for the creation of sustainable national food security. <https://doi.org/10.1063/5.0142824>
- Adner, R. (2017). Ecosystem as structure: An actionable construct for strategy. *Journal of Management*, 43(1), 39-58.
- Alves, R., Ascensão, J., Camelo, D., & Matos, P. (2021). eWeightSmart - A smart approach to beef production management. In S. Boumerdassi, M. Ghogho, & É. Renault (Eds.), *Smart and sustainable agriculture* (pp. 57-70). Springer International Publishing. https://doi.org/10.1007/978-3-030-88259-4_5
- Atik, C., & Martens, B. (n.d.). Competition problems and governance of non-personal agricultural machine data: Comparing voluntary initiatives in the US and EU.
- Azkan, C., Möller, F., Ebel, M., Iqbal, T., Otto, B., & Poepplbuss, J. (2022, December). Hunting the treasure: Modeling data ecosystem value co-creation. ICIS.
- Bachmann, R., & Inkpen, A. C. (2011). Understanding institutional-based trust building processes in inter-organizational relationships. *Organization Studies*, 32(2), 281-301.
- Balanescu, M., Badicu, A., Suci, G., Poenaru, C., Pasat, A., Vulpe, A., & Vochin, M. (2020). Decision support platform for intelligent and sustainable farming. In 2020 IEEE 26th International Symposium for Design and Technology in Electronic Packaging (SIITME) (pp. 89-93). IEEE. <https://doi.org/10.1109/SIITME50350.2020.9292196>
- Berti, G., Mulligan, C., & Yap, H. (2017). Digital food hubs as disruptive business models based on coopeition and 'shared value' for sustainability in the agri-food sector. In S. Sindakis & P. Theodorou (Eds.), *Global opportunities for entrepreneurial growth: Cooperation and knowledge dynamics within and across firms* (pp. 415-438). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-78714-501-620171023>
- Braun, A.-T., Colangelo, E., & Steckel, T. (2018). Farming in the era of Industrie 4.0. *Procedia CIRP*, 72, 979-984. <https://doi.org/10.1016/j.procir.2018.03.176>
- Brewster, C., Roussaki, I., Kalatzis, N., Doolin, K., & Ellis, K. (2017). IoT in agriculture: Designing a Europe-wide large-scale pilot. *IEEE Communications Magazine*, 55(9), 26-33. <https://doi.org/10.1109/MCOM.2017.1600528>
- Cavazza, A., Dal Mas, F., Paoloni, P., & Manzo, M. (2023). Artificial intelligence and new business models in agriculture: A structured literature review and future research agenda. *British Food Journal*, 125(13), 436-461. <https://doi.org/10.1108/BFJ-02-2023-0132>
- Cook, S., Jackson, E. L., Fisher, M. J., Baker, D., & Diepeveen, D. (2022). Embedding digital agriculture into sustainable Australian food systems: Pathways and pitfalls to value creation. *International Journal of Agricultural Sustainability*, 20(3), 346-367. <https://doi.org/10.1080/14735903.2021.1937881>
- Daum, T., Ravichandran, T., Kariuki, J., Chagunda, M., & Birner, R. (2022). Connected cows and cyber chickens? Stocktaking and case studies of digital livestock tools in Kenya and India. *Agricultural Systems*, 196, 103353. <https://doi.org/10.1016/j.agry.2021.103353>
- De Oliveira, F. J. B., Ferson, S., & Dyer, R. (2021). A collaborative decision support system framework for vertical farming business developments. *International Journal of Decision Support System Technology*, 13(1), 1-33. <https://doi.org/10.4018/IJDSST.2021010103>

- Di Vaio, A., Boccia, F., Landriani, L., & Palladino, R. (2020). Artificial intelligence in the agri-food system: Rethinking sustainable business models in the COVID-19 scenario. *Sustainability*, 12(12), 4851. <https://doi.org/10.3390/su12124851>
- Dwi Satya, R. R., Marimin, Eriyatno, & Ismayana, A. (2021). A digital business modelling for green supplier selection of potato chips agroindustry. *IOP Conference Series: Earth and Environmental Science*, 709(1), Article 012084. <https://doi.org/10.1088/1755-1315/709/1/012084>
- Elkington, J. (1997). *Cannibals with forks: The triple bottom line of 21st century business*. Capstone Publishing.
- Freeman, R. E. (2020). The stakeholder approach revisited. *Wirtschafts-und Unternehmensethik*, 657-671.
- Friedman, B. (1996). Value-sensitive design. *interactions*, 3(6), 16-23.
- Gebresenbet, G., Bosona, T., Patterson, D., Persson, H., Fischer, B., Mandaluniz, N., Chirici, G., Abebe, A., Andersson, C., Arlt, H.-J., Berg, S., Bernet, D., Bottek, S., Cazzato, E., Duffy, C., Gallardo, M., Ghirmai, W., Grace, P., Grosse-Lochtman, J., ... Vieri, M. (2023). A concept for application of integrated digital technologies to enhance future smart agricultural systems. *Smart Agricultural Technology*, 5, 100255. <https://doi.org/10.1016/j.atech.2023.100255>
- Giagnocavo, C., Bienvenido, F., Ming, L., Yurong, Z., Sanchez-Molina, J. A., & Xinting, Y. (2017). Agricultural cooperatives and the role of organisational models in new intelligent traceability systems and big data analysis. *International Journal of Agricultural and Biological Engineering*, 10(5), 115-125. <https://doi.org/10.25165/j.ijabe.20171005.3089>
- Gowri, D. P., & Ramachander, A. (2024). Digital agricultural ecosystem: Revolutionary advancements in agriculture (Chapter 8). In K. Singh & P. Kolar (Eds.), *Digital agriculture: Transforming farming practices and food systems for a sustainable future*. Wiley. <https://doi.org/10.1002/9781394242962.ch8>
- Grabher, G. (2021). Enclosure 4.0: Seizing data, selling predictions, scaling platforms. *Sociologica*, 241-265. <https://doi.org/10.6092/ISSN.1971-8853/12107>
- Graf-Drasch, V., Kauffeld, L., Kempf, L., Oberländer, A. M., & Teuchert, A. (2023, May). Driving twin transformation-the interplay of digital transformation and sustainability transformation. In *European Conference on Information Systems 2023*.
- Hansen, B. D., Leonard, E., Mitchell, M. C., Easton, J., Shariati, N., Mortlock, M. Y., Schaefer, M., & Lamb, D. W. (2022). Current status of and future opportunities for digital agriculture in Australia. *Crop & Pasture Science*, 74(6), 524-537. <https://doi.org/10.1071/CP21594>
- Holden, N. M., White, E. P., Lange, M. C., & Oldfield, T. L. (2018). Review of the sustainability of food systems and transition using the Internet of Food. *Npj Science of Food*, 2(1), 18. <https://doi.org/10.1038/s41538-018-0027-3>
- Hrustek, L. (2020). Sustainability driven by agriculture through digital transformation. *Sustainability*, 12(20), Article 8596. <https://doi.org/10.3390/su12208596>
- Hu, B., Zhou, P., & Zhang, L. P. (2022). A digital business model for accelerating distributed renewable energy expansion in rural China. *Applied Energy*, 316, 119084. <https://doi.org/10.1016/j.apenergy.2022.119084>
- Huber, F., & Markward, D. (2021). *AgriCircle: Innovating agricultural ecosystems*. In O. Gassmann & F. Ferrandina (Eds.), *Connected business* (pp. 325-331). Springer International Publishing. https://doi.org/10.1007/978-3-030-76897-3_21
- Issa, A. A., Majed, S., Ameer, A., & Al-Jawahry, H. M. (2024). IoT and AI in livestock management: A game changer for farmers. *E3S Web of Conferences*, 491, 02015. <https://doi.org/10.1051/e3sconf/202449102015>
- Kagabo, D. M., Byandaga, L., Gatsinzi, P., Mvuyibwami, P., Munyangeri, Y. U., Ntwari, N., & Ouedraogo, M. (2025). Scaling climate information services and climate smart agriculture through bundled business models. *Climate Services*, 37, 100526. <https://doi.org/10.1016/j.cliser.2024.100526>
- Kampker, A., Jussen, P., & Moser, B. (2018). Industrial smart services: Types of smart service business models in the digitalized agriculture. In *2018 IEEE International Conference on*

- Industrial Engineering and Engineering Management (IEEM) (pp. 1081-1085). IEEE.
<https://doi.org/10.1109/IEEM.2018.8607270>
- Kampker, A., Stich, V., Jussen, P., Moser, B., & Kuntz, J. (2019). Business models for industrial smart services – The example of a digital twin for a product-service-system for potato harvesting. *Procedia CIRP*, 83, 534-540. <https://doi.org/10.1016/j.procir.2019.04.114>
- Kealley, M. J., Murdoch, E. M., & Shannon, T. J. (n.d.). Emerging opportunities for Australian sugar using blockchain technology.
- Klug, L., & Prinz, W. (2023). Fair prices for sustainability in agriculture and food. Requirements and design options for a data-based transparency system. In *Proceedings of the 24th Annual International Conference on Digital Government Research* (pp. 496-507). ACM.
<https://doi.org/10.1145/3598469.3598525>
- Kosior, K., & Mlodawska, P. (2024). Understanding market actors' perspectives on agri-food data sharing: Insights from the digital food passports pilot in Poland. *Agriculture*, 14(12), 2340.
<https://doi.org/10.3390/agriculture14122340>
- Kumarathunga, M., & Ginige, A. (2022). Business model transformation with technology-enabled social capital in agriculture domain.
- Mahdad, M., Hasanov, M., Isakhanyan, G., & Dolfmsa, W. (2022). A smart web of firms, farms and Internet of Things (IoT): Enabling collaboration-based business models in the agri-food industry. *British Food Journal*, 124(6), 1857-1874. <https://doi.org/10.1108/BFJ-07-2021-0756>
- Malisic, B., Misic, N., Krco, S., Martinovic, A., Tinaj, S., & Popovic, T. (2023). Blockchain adoption in the wine supply chain: A systematic literature review. *Sustainability*, 15(19), 14408.
<https://doi.org/10.3390/su151914408>
- Neethirajan, S., & Kemp, B. (2021). Digital livestock farming. *Sensing and Bio-Sensing Research*, 32, 100408. <https://doi.org/10.1016/j.sbsr.2021.100408>
- Poppe, K., Vrolijk, H., & Van Dijk, R. (2021). Design of a system for information transfer to reduce administrative burdens in the agrifood sector. *International Journal on Food System Dynamics*, 12, 301-313. <https://doi.org/10.18461/IJFSD.V12I4.92>
- Rao, N. H. (2007). A framework for implementing information and communication technologies in agricultural development in India. *Technological Forecasting and Social Change*, 74(4), 491-518. <https://doi.org/10.1016/j.techfore.2006.02.002>
- Rijanto, A. (2021). Business financing and blockchain technology adoption in agroindustry. *Journal of Science and Technology Policy Management*, 12(2), 215-235.
<https://doi.org/10.1108/JSTPM-03-2020-0065>
- Rijswijk, K., Klerkx, L., & Turner, J. A. (2019). Digitalisation in the New Zealand agricultural knowledge and innovation system: Initial understandings and emerging organisational responses to digital agriculture. *NJAS: Wageningen Journal of Life Sciences*, 90-91(1), 1-14.
<https://doi.org/10.1016/j.njas.2019.100313>
- Routis, G., Paraskevopoulos, M., Vetsikas, I. A., Roussaki, I., Stavrakoudis, D., & Katsantonis, D. (2022). Data-driven and interoperable smart agriculture: An IoT-based use-case for arable crops. In *2022 IEEE International Conference on Omni-Layer Intelligent Systems (COINS)* (pp. 1-8). IEEE. <https://doi.org/10.1109/COINS54846.2022.9855001>
- Stitzlein, C., Fielke, S., Waldner, F., & Sanderson, T. (2021). Reputational risk associated with big data research and development: An interdisciplinary perspective. *Sustainability*, 13(16), 9280.
<https://doi.org/10.3390/su13169280>
- Tasic, I., & Cano, M.-D. (2024). An orchestrated IoT-based blockchain system to foster innovation in agritech. *IET Collaborative Intelligent Manufacturing*, 6(2), e12109.
<https://doi.org/10.1049/cim2.12109>
- Van de Poel, I. (2020). Embedding values in artificial intelligence (AI) systems. *Minds and machines*, 30(3), 385-409.

- Vial, F., & Tedder, A. (2017). Tapping the vast potential of the data deluge in small-scale food-animal production businesses: Challenges to near real-time data analysis and interpretation. *Frontiers in Veterinary Science*, 4, 120. <https://doi.org/10.3389/fvets.2017.00120>
- Vlachopoulou, M., Ziakis, C., Vergidis, K., & Madas, M. (2021). Analyzing AgriFood-Tech e-Business models. *Sustainability*, 13(10), 5516. <https://doi.org/10.3390/su13105516>
- World Commission on Environment and Development. (1987). *Our common future*. Oxford University Press.
- Zheng, X., Wu, C., Tian, D., & Zhang, X. (2009). B2B e-marketplace adoption in agriculture. *Journal of Software*, 4(3), 232-239. <https://doi.org/10.4304/jsw.4.3.232-239>

ADVANCEMENTS IN NEWS RECOMMENDATION SYSTEMS: THE ROLE AND IMPACT OF ARTIFICIAL INTELLIGENCE AND LARGE LANGUAGE MODELS

CEYHUN KAZANÇ, NAZİM TAŞKIN

Boğaziçi University, Department of Management Information Systems, Istanbul,
Türkiye
ceyhun.kazanc@std.bogazici.edu.tr, nazim.taskin@bogazici.edu.tr

The rapid evolution of artificial intelligence (AI) and large language models (LLMs) significantly advanced the news recommendation systems (NRS). However, comprehensive analyses evaluating recent advancements and practical implications of integrating AI and LLMs into NRS remain scarce in existing literature. This study systematically examines AI and LLMs' effects and usage methods on NRS and analyzes 42 studies using the Prisma methodology. It emphasizes the features of collaborative filtering (CF), content-based filtering (CB), hybrid systems, AI based systems including LLM-based models like BERT. While these technologies offer advanced semantic analysis and real-time adaptability opportunities, they are still partially affected by traditional challenges, such as cold-start and data sparseness, though less so than traditional methods. This study emphasizes innovations in AI-driven NRS, focusing on hybrid approaches, session-based and multi-interest models and efficient use of LLM. The findings provide actionable insights for researchers and practitioners seeking to optimize NRSs in an increasingly dynamic digital landscape.

DOI
[https://doi.org/
10.18690/um.fov.4.2025.42](https://doi.org/10.18690/um.fov.4.2025.42)

ISBN
978-961-286-998-4

Keywords:
news recommendation
systems,
artificial intelligence (AI),
large language models
(LLMs),
personalization,
hybrid systems



University of Maribor Press

1 Introduction

In an age of overloading information, personalized news recommendation systems (NRSs) play a critical role in providing content compatible with user preferences and behaviors. NRSs are algorithms that provide personalized news content based on users' interests and reading habits. These systems work using a variety of techniques, including content-based filtering CB recommending items based on the features of the items themselves and a user's past preferences (Symeonidis, Chaltsev, et al., 2021), collaborative filtering CF, recommending items based on the preferences of similar users or the similarity between items (Han, 2024), and deep learning (Karimi et al., 2018). In recent years, deep learning-based models have made great progress in NRS (Mahesh et al., 2022). Additionally, integrating editorial feedback into recommendation systems is an important step toward improving the accuracy and quality of news (Mahmood et al., 2024).

The rapid evolution of AI technologies, especially LLMs, provides features that can provide significant breakthroughs in this field by making systems better understand user behavior, news content semantics, and context (Amir et al., 2023). LLMs are AI models trained on large amounts of text data to perform natural language processing (NLP) tasks. These models excel at tasks, such as text generation, summarization, translation, and question-answering (Pahune & Chandrasekharan, 2023). Developments in LLMs and NLPs have significantly enhanced the accuracy and efficiency of digital news platforms in providing personalized news suggestions.

In spite of this progress, new problems can arise with new technologies while some challenge in the field continues. Previous approaches, such as CF and CB often struggled with problems, such as cold-start problem, which refers to a difficulty that occurs in recommendation systems when there is not enough data about new users or new products, preventing the system from making accurate predictions and reducing the quality of recommendations (Moghaddam & Elahi, 2019). Another issue is the risk of filter bubbles where personalized algorithms showing content aligned with users' interests. This limits exposure to different perspectives and reduces information diversity (Flaxman et al., 2016). Although hybrid systems offer improvements regarding some of these problems, the dynamic structure of user interest and the complexity of news data requires more sophisticated methods. The latest developments in LLMs, such as Bert and GPT models offer new opportunities

by improving their semantic understanding and contextual suggestion skills (Zhang et al., 2021). However, calculation costs and sustainability concerns show that more research is needed.

The motivation for this study stems from the gap in understanding the full scope of AI on NRS, especially the role of LLMs. While the existing investigations provide valuable insights (Raza & Ding, 2022) on previous methods, such as CB, CF, hybrid approaches, and deep learning-based models, the effects of the latest LLM and Generative AI progress are omitted (Amir et al., 2023) with the rapid development of these technologies. While some previous studies (Hou et al., 2023) have addressed the impact of LLMs and Generative AI, a study conducted after the development of these technologies will better explain and compare current methods. Although previous studies (Zhang et al., 2021) examined the impact of LLMs and generative AI, further research conducted following recent advancements in these technologies is needed to better understand and comparatively evaluate current methods. This study addresses this gap by systematically examining the integration of AI and LLMs into NRS, focusing on system performance, user personalization and contribution to contextual understanding.

To guide this exploration, the study addresses two central research questions:

1. What are the approaches used in news recommendation systems (NRS) in the literature after the development of large language models (LLMs), and what are their advantages, limitations, and the role of artificial intelligence in these approaches?
2. How have LLMs influenced the development and performance of NRS?

By systematically analyzing the 42 studies published after 2021, this research aims to provide a comprehensive understanding of progress, difficulties and future aspects in the field and to provide valuable insights for researchers and practitioners.

2 Methodology

This study uses the PRISMA guidelines to systematically review the existing literature on the integration of NRSs with AI and LLM (Moher et al., 2009). The PRISMA guidelines provide rigorous and transparent framework for identifying,

screening, and selecting studies relevant to the research questions, ensuring the reliability and reproducibility of the review process.

To collect relevant studies, we conducted a comprehensive search of two academic databases, Web of Science and Scopus, due to their worldwide importance (Mongeon & Paul-Hus, 2016). In the search, we used the following query strings to capture a wide range of research in the field:

"news recommendation" and "artificial intelligence", "news personalization" and "artificial intelligence", "news recommendation" and "large language models", "news personalization" and "large language models".

This search yielded a total of 129 studies. Of these, 23 were eliminated due to duplicate records, leaving 106 studies for further evaluation. We excluded 42 studies conducted before 2021, since the release of ChatGPT-3.5 that year has irrevocably changed the landscape of AI-based recommendation systems (Myers et al., 2024). The rationale for this decision was to be able to investigate the impact of LLMs and AI on NRS methods with these revolutionary developments in AI technology. We then screened each study based on the availability of title, abstract, and full text to determine its relevance to the research questions. Studies that were unrelated to the role of AI or LLMs in NRS or did not have full-text access were excluded, resulting in the elimination of 22 studies. Details can be examined on the flow diagram in Figure 1 (see Appendix A).

Considering that content-based, collaborative and hybrid filtering are definitions that have been used long without AI technologies (Aboutorab et al., 2023), we have explained the advantages and limitations of models that are based on AI and can be used with AI in Tables 1, 3, 5, and 6. Although the studies on the subject are limited for now, we meticulously compiled the information about the use of LLMs in NRSs in the existing literature. The last part included the results by interpreting the impact of AI on NRSs, the comparison of AI approaches, the impact of LLMs and their future according to the inferences we made from the literature.

The final selection included 42 studies that met the inclusion criteria. We analyzed these studies to address the research questions and focused on the impact of AI on NRS and the impact of LLMs on their development and performance. Through this

process, we aimed to provide a comprehensive and unbiased review of the latest developments in the field.

3 Common Approaches in News Recommendation Systems

3.1 Collaborative Filtering

Collaborative filtering (CF) is a recommendation system that analyzes user behavior and makes predictions about possible future behaviors of users (Shichang, 2021). CF methods are based on the idea that users who have made similar choices in the past will make similar choices in the future (Feng & Lv(U), 2022).

We can examine CF models under two headings: User-Based CF identifies users with similar tastes and recommends items that these users have liked in the past to others (Ludwig et al., 2023). For example, if user 1 and user 2 both showed interest in news x and y, the system predicts that these users' preferences are similar. If user 1 shows interest in news z, the system recommends news z to user 2 (X. Liu, 2022). Item-Based CF focuses on similarities between items and recommends items that are similar to items the user has liked before (Ludwig et al., 2023). For example, if a user is reading a news article, the recommendation system shows related articles in a news feed component (Lucas & de Figueiredo, 2023).

Although CF is one of the most prominent methods used by sources, such as Google News among NRSs, it inherently suffers from a cold-start problem when it comes to new users and new items (Lv et al., 2024). To overcome such difficulties, hybrid approaches are preferred by using models such as CB (Zhu et al., 2022) and session-based (Pourashraf & Mobasher, 2023).

3.2 Content-Based Filtering

Content-based (CB) NRSs use the content of news articles to make recommendations (Fieiras-Ceide et al., 2023). These systems analyze features, such as news topics or content and suggest news similar to what the user has previously been interested in (Ludwig et al., 2023). CB methods are widely used in news recommendations because users usually do not have long-term profiles on news sites (Alam et al., 2022).

CB NRSs represent news articles using feature vectors (Ludwig et al., 2023), which can include information like topics discussed in an article, the news outlet that published it, or other relevant metadata (Manh Nguyen et al., 2022). Similarity calculation in CB calculates the similarity between news articles based on their feature vectors (Alam et al., 2022). Term frequency-inverse document frequency (TF-IDF) is the most widely used method for numerically representing text. This technique converts the similarities between news items into values between 0 and 1 using cosine similarity, allowing for accurate recommendations (Ludwig et al., 2023).

CB methods build user profiles based on their past interactions with news articles. These profiles represent a user's interests and preferences (Lv et al., 2024). They do not rely on the behavior of other users (Ludwig et al., 2023). They expand user profiles using personal information and address the cold-start problem (J. Yang et al., 2023). With these features, they are useful at utilizing the rich textual content of news articles (Fu, 2023).

Since CB systems tend to show users news that is similar to news they have seen before, there is a possibility that they will present the same or very similar news to the reader again (Ludwig et al., 2023). These systems can increase the risk of locking users into “filter bubbles” by focusing on similar content, where users are constantly presented with news that is similar to their preferences or to news they have previously engaged with, potentially isolating users from different perspectives (Sun et al., 2021; Vrijenhoek, 2023). Some CF algorithms use only news headlines as input, and headlines often carry a significant bias (Ruan et al., 2023). In order overcome the problems in these systems or to reduce their effects, CB systems are used together with systems, such as CF to form Hybrid Systems (Q. Zhao et al., 2024).

3.3 Hybrid Systems

Hybrid NRSs combine different approaches to improve the accuracy and effectiveness of recommendations (Manh Nguyen et al., 2022). Besides demographic and knowledge-based filtering models, which are beyond the scope of this article as they do not rely on AI (Ludwig et al., 2023), hybrid systems combining models like CB and CF aim to overcome limitations of single-model approaches (Song et al., 2021).

Common hybrid approaches and techniques include CB and CF combining user profiles, news content, and user behavior gives recommendations based on user's interests and past actions (Manh Nguyen et al., 2022). Hybrid Deep Learning Models utilize neural networks to integrate content and user data. For example, ACCM (2018) uses Cold-Sampling strategy that can recommend items even if the user or item is new to the system (Shi et al., 2021). Multi-faceted Hybrid Systems combine content, CF and knowledge graphs (Xu & Gu, 2022). Similarly, Multi-Objective Approaches consider multiple objectives, such as accuracy and diversity, in their recommendations (M. Zhao et al., 2023).

Hybrid systems overcome the cold-start problem, increase diversity (Ludwig et al., 2023), accuracy (X. Liu, 2022) and stability (Song et al., 2021). Additionally, they provide a better user experience (Cui et al., 2021), reduce information overload (X. Liu, 2022), and create more robust systems (Shi et al., 2021).

4 Ai Driven Techniques in News Recommendation Systems

4.1 Session-Based Systems

Session-based NRSs focus on providing recommendations based on a user's current, short-term interactions with the news platform, rather than relying on long-term user profiles (Symeonidis, Chaltsev, et al., 2021). These systems are particularly effective in the news domain, as users' interests vary on a session-by-session basis and most users are anonymous (Lucas & de Figueiredo, 2023).

Session-based systems analyze the sequence of articles users interact with, in a single session to understand their immediate interests (Sheu et al., 2022). A session is defined as a series of interactions within a short time (Symeonidis, Kirjackaja, et al., 2021). These systems aim to capture short-term user preferences, as opposed to long-term interests (Sheu et al., 2022), as user interests can change quickly (Pourashraf & Mobasher, 2023).

Session-based methods can effectively deal with anonymous users since they do not require past user history or profiles (Lucas & de Figueiredo, 2023). The sequential order of articles viewed within a session plays a crucial role, as click order provides valuable insights into users' evolving interests within a given session (Sheu et al.,

2022). Some advanced models consider both intra-session element transitions and inter-session transitions to address the problem of insufficient information at the beginning of the session (Symeonidis, Chaltsev, et al., 2021). Table 1 briefly overviews session-based systems, highlighting their strength in capturing short-term behavior and limitation in reflecting long-term preferences(see Appendix B) .

4.2 Graph-Based Systems

Graph-based systems for news recommendation leverage the relationships between different entities, such as users, news articles, topics, to improve the accuracy and relevance of recommendations (Li et al., 2023). In other words, these systems can work with other recommendation systems that keep users, content, products(items), etc. in the form of graphics and control their relationships with each other (Sheu et al., 2022).

Table 2 provides a concise overview of various graph-based approaches currently employed in NRS, emphasizing how each leverages AI to capture and analyze relationships among entities. It highlights key characteristics and research insights, illustrating both their contributions and limitations within the domain (see Appendix C).

Table 3 provides a summary of the advantages and limitations of graph-based systems in news recommendation, highlighting their effectiveness in modeling relationships and their key challenges, such as computational complexity and scalability (Li et al., 2023) (see Appendix D).

4.3 Deep Learning Based Systems

Deep learning (DL) algorithms work in a way similar to the way the human brain works and can automatically be trained with large amounts of data to produce high-quality semantic inferences. Thanks to their ability to be trained continuously, they can optimize their recommendations up to date (Han, 2024). While DL-based methods have shown their effectiveness in different domains, such as image recognition and NLP, extensive experiments have been conducted in NRSs (Lim et al., 2022).

The approaches summarized in Table 4 represent significant advancements in applying Deep Neural Networks (DNNs) within NRS. These methods utilize AI's capability to handle complex user interactions, semantic contexts, and sequential data, directly addressing challenges such as data sparsity and personalization (see Appendix E). This aligns with the research questions by highlighting how various DNN architectures, including CNNs, RNNs, Transformers, and attention mechanisms, influence the performance and development of NRS, particularly in the context of recent advancements in AI and LLMs.

Table 5 summarizes the advantages and limitations of deep learning-based systems in NRS, emphasizing their strengths in modeling complex user interactions, semantic understanding, and addressing sparsity, but also highlighting their challenges, such as computational complexity and data requirements (see Appendix F).

4.4 Multi-Interest Based Systems

Instead of single-interest models in NRSs, multi-interest based models that aim to capture diverse and varied interests of users, overcome the limitations of single-interest based models (P. Zhao et al., 2023). These systems observe that users generally have multiple areas of interest that can change over time (Hou et al., 2023). Instead of fitting a user's interests into a single mold, multi-interest systems attempt to create several different interest profiles for each user (P. Zhao et al., 2023).

Multi-interest systems learn multiple representation vectors for each user, with each vector representing a different interest (Hou et al., 2023). Some models create interest prototypes that serve as the basis for learning multiple user representations. These prototypes may reflect different topics or interests (S. Wang et al., 2023). Multi-head self-attention models can be used to detect potential interest areas in each clicked news. Each title vector represents an interest area (R. Wang et al., 2022).

Graph Neural Networks (GNNs) are used to model user interests using information collected from user neighbors in a user-news graph. Some models leverage personalized graphs to capture more nuanced relationships between user behaviors (Wu et al., 2021). GNNs can also be used to improve user representations by

evaluating the relationships between users and news on graphs containing different types of nodes and relationships (S. Wang et al., 2023).

Multi-channel information fusion involves modeling user interests and capturing diversity by combining news content from different sources, at different levels of detail, and using different types of information, such as text and assets (Z. Yang et al., 2023).

As shown in Table 6, multi-interest models capture diverse, multi-grained user preferences, from low-level keywords to high-level topics, and track both short- and long-term dynamics to enhance recommendation accuracy and adaptability. However, as detailed in Table 6, they may misinterpret curiosity-driven clicks, introduce noise and redundancy when the number of interests grows, and oversimplify news by overlooking key events and stylistic nuances (see Appendix G).

4.5 LLMs In NRSs

Despite the limited research on the topic, important insights have emerged from the use of LLMs particularly models like BERT and other PLMs in NRSs.

BERT (Bidirectional Encoder Representations from Transformers) is a pre-trained language model (PLM) that can capture contextual information and semantic meaning of words. BERT-based models improve the understanding of news content by learning richer semantic information (Hou et al., 2023). It is used to initialize the word-level module of some news recommendation models and enables the development of textual representations. It further helps to alleviate the cold-start problem by taking advantage of rich grammar. Pre-trained models like BERT are used to create contextual word embeddings. These embeddings help models capture the meaning of words based on how they are used in a sentence (Zhang et al., 2021).

Using BERT or other PLMs such as RoBERTa and GPT often leads to significant performance improvements in NRSs (S. Wang et al., 2023). For example, one study showed that a model using BERT outperformed other models, with a 2.81% improvement in AUC (area under the curve) on a large dataset (Zhang et al., 2021).

Although PLMs are effective, they often require high computational resources. Traditional end-to-end training paradigms also face efficiency challenges, particularly due to the repeated encoding of the same news content. To address this, Q. Liu et al. (2024) introduced the OLEO (Only-Encode-Once) paradigm, which involves pre-training the content encoder before integrating it into the model. This approach has shown comparable accuracy while significantly improving sustainability metrics by up to 2992% compared to state-of-the-art end-to-end methods. (Q. Liu et al., 2024). Beyond BERT, transformer-based models like ELECTRA are also being used for pre-training text encoders as discriminators rather than generators (J. Yang et al., 2023). This shift in pre-training strategy offers a more sample-efficient alternative, which could further reduce computational costs while maintaining or even improving model performance in NRSs.

5 Discussion

This research aims to examine the impact and applications of AI, and especially LLMs, on NRSs. The findings show that AI-driven approaches provide improved results in personalization, accuracy, and efficiency compared to traditional methods. However, challenges such as bias, filter bubbles, and computational costs remain for some methods.

5.1 The Impact of AI

With the advancement in AI, significant advances have been made in NRSs by developing personalized experiences tailored to individual user preferences. CF, which we also encounter in traditional approaches, effectively uses user behavior patterns to predict future interests, although it has a cold-start problem. CB significantly prevents the cold-start problem by analyzing news content, but it carries the risk of creating filter bubbles because it constantly recommends similar content. Similarly, Flaxman et al. (2016) note that CB often narrows exposure by reinforcing user biases, contributing to the filter bubble effect. Hybrid systems combining CF and CB significantly eliminate these limitations by taking advantage of both user behavior and content similarity. X. Liu (2022) also supports that by emphasizing how hybrid models outperform single approaches in terms of accuracy and diversity. Session-based systems, which focus on short-term interactions, make effective recommendations to anonymous users and dynamic interests, and demonstrate the

ability to adapt to real-time user behavior thanks to AI-supported systems. Symeonidis et al. (2023) also highlight the value of session-based models in addressing the transient nature of user interests in news domains. Graph-based systems and deep learning-based systems allow complex understanding of user preferences and news content connections by using complex relationships between entities together with traditional methods. Li et al. (2023) and W. Liu et al. (2024) similarly show that graph-based methods capture higher-order interactions, improving recommendation precision.

5.2 Comparing AI-driven Techniques

Each AI-based NRS we mentioned has its own strengths and weaknesses. While CF is very successful in capturing the behavior of users by grouping them, it struggles when dealing with new users and items. While CB overcomes these problems with its content-focused approach, it fails to direct users to new interests, constantly recommending similar content. Hybrid systems aim to overcome these problems by using the strengths of both. While session-based systems address the dynamic structure of the news domain, graph-based methods exploit complex relationships to provide more sophisticated recommendations. Deep learning models, such as CNNs, RNNs, and attention networks are powerful tools to analyze complex data and provide more accurate and personalized recommendations when used in conjunction with other methods. Multi-interest-based systems increase the performance in personalization by capturing different interests of users. The selection of the most appropriate method depends on the requirements and data characteristics.

5.3 The Influence of LLMs

Although there are few studies on the use of LLMs in NRSs, we can say that LLM models, such as BERT have started a new era with the performance increase, they bring in news recommender systems as well as in many other areas. When we look at CB models before LLMs, LLMs that capture contextual information and meanings provide a more detailed understanding of news content compared to models that measure the frequency of words used. In this way, the increase in recommendation accuracy has been proven numerically by the studies (Zhang et al., 2021). However, when it comes to cost and sustainability, the use of LLMs in NRSs

can be a source of concern. LLMs have become more competitive in terms of cost and sustainability with innovative approaches such as the OLEO paradigm. Recent frameworks like OLEO aim to address these concerns, aligning with Hou et al. (2023) who suggest lightweight multi-head attention as a viable cost-efficient alternative.

5.4 The Future of LLMs in News Recommendation

The potential of LLMs in developing NRSs cannot be underestimated. Future research should explore more efficient training paradigms by focusing on reducing costs to increase sustainability. Hybrid use of LLMs with other recommender system methods should also be tried to optimize costs. Combining LLMs with session-based models can help capture the dynamic nature of news and improve accuracy, while using them with multi-interest models may help personalize content and introduce users to new topics. In addition to CB methods, using LLMs together with CF methods, especially with item-based CF methods, is a promising method for developing recommendation systems with semantic similarities of news, and reducing costs.

5.5 Limitations

The study is not without limitations. First of all, since this review focused on AI-driven methods, it ignored studies that examined the possible contributions of other approaches. Although the focus was on recent studies, the rapid development of the field under investigation means that new techniques can emerge rapidly, and potentially some findings can quickly become outdated. Additionally, the inclusion criteria limited the scope to studies published after 2021, which may have excluded foundational works or earlier influential models still relevant today. Another limitation is the variation in evaluation metrics and datasets used across studies, which makes it difficult to directly compare the performance of different models. Finally, the review emphasizes technical and performance-based aspects, while socio-ethical implications, such as user trust, transparency, and fairness in AI-driven news recommendation received limited attention.

6 Conclusion

This study highlights the role of AI and LLMs in the development of NRSs. AI has been effective in improving personalization, scalability, and contextual understanding with hybrid, graph-based, and deep-learning approaches that address limitations such as cold-start problems and data sparsity. Although LLMs such as BERT offer new paradigms for efficiency and sustainability by further improving semantic analysis and recommendation accuracy, more observations are needed as these developments are still very recent.

Despite these developments, challenges remain in terms of computational efficiency, ethical issues, and maintaining diversity in recommendations. Future efforts should focus on energy-efficient systems, cost reduction, sustainability, and integrating multimodal data such as textual, visual, and behavioral.

References

- Aboutorab, H., Hussain, O. K., Saberi, M., Hussain, F. K., & Prior, D. (2023). Reinforcement learning-based news recommendation system. *IEEE Transactions on Services Computing*, 16(6), 4493-4502.
- Alam, M., Iana, A., Grote, A., Ludwig, K., Müller, P., & Paulheim, H. (2022). Towards Analyzing the Bias of News Recommender Systems Using Sentiment and Stance Detection. *Companion Proceedings of the Web Conference 2022*, 448-457.
- Amir, N., Jabeen, F., Ali, Z., Ullah, I., Jan, A. U., & Kefalas, P. (2023). On the current state of deep learning for news recommendation. *Artificial Intelligence Review*, 56(2), 1101-1144.
- Cui, C., Wang, C., Wen, C., Fang, X., & Deng, Y. (2021). Intelligent News Recommendation System for Automobile Based on Bio-Sensing and Physiological Computing. *Proceedings of the 2021 4th International Conference on Signal Processing and Machine Learning*, 107-111.
- Feng, Y., & Lv, X. (2022). Frontier application and development trend of artificial intelligence in new media in the AI era. In *The 2021 International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy: SPIoT-2021 Volume 1* (pp. 58-64). Springer International Publishing.
- Fieiras-Ceide, C., Vaz-Álvarez, M., & Túniz-López, M. (2023). Designing personalisation of European public service media (PSM): trends on algorithms and artificial intelligence for content distribution. *Profesional de la información*, 32(3).
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1), 298-320.
- Fu, W. (2023). AI-news personalization system combining complete content characterization and full term interest portrayal in the big data era. *IEEE Access*, 11, 85086-85096.
- Han, L. (2024). Research on the Application of Fusion Media News Push Technology Based on Deep Learning. *2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)*, 1-4.
- Hou, Y., Ouyang, Y., Liu, Z., Han, F., Rong, W., & Xiong, Z. (2023, August). Multi-level and multi-interest user interest modeling for news recommendation. In *International Conference on*

- Knowledge Science, Engineering and Management (pp. 200-212). Cham: Springer Nature Switzerland.
- Karimi, M., Jannach, D., & Jugovac, M. (2018). News recommender systems—Survey and roads ahead. *Information Processing & Management*, 54(6), 1203-1227..
- Li, X., Li, R., Peng, Q., & Ma, H. (2023, August). Candidate-aware attention enhanced graph neural network for news recommendation. In *International Conference on Knowledge Science, Engineering and Management* (pp. 244-255). Cham: Springer Nature Switzerland..
- Lim, H., Lee, Y. C., Lee, J. S., Han, S., Kim, S., Jeong, Y., ... & Kim, S. W. (2022, May). AiRS: a large-scale recommender system at naver news. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)* (pp. 3386-3398). IEEE.
- Liu, Q., Zhu, J., Dai, Q., & Wu, X.-M. (2024). Benchmarking News Recommendation in the Era of Green AI. *Companion Proceedings of the ACM Web Conference 2024*, 971–974.
- Liu, W., Zhang, Z., & Wang, B. (2024). Dual-view hypergraph attention network for news recommendation. *Engineering Applications of Artificial Intelligence*, 133, 108256.
- Liu, X. (2022). Impact of Personalized Recommendation on 'Today's News Communication through Algorithmic Mechanism in the New Media Era. *Advances in Multimedia*, 2022(1), 1284071.
- Lucas, J. P., & de Figueiredo, L. F. (2023). Challenges and Solutions in Large-Scale News Recommenders. In D. Durães, A. González-Briones, M. Lujak, A. El Bolock, & J. Carneiro (Eds.), *Highlights in Practical Applications of Agents, Multi-Agent Systems, and Cognitive Mimetics. The PAAMS Collection* (pp. 170–181). Springer Nature Switzerland.
- Ludwig, K., Grote, A., Iana, A., Alam, M., Paulheim, H., Sack, H., Weinhardt, C., & Müller, P. (2023). Divided by the Algorithm? The (Limited) Effects of Content- and Sentiment-Based News Recommendation on Affective, Ideological, and Perceived Polarization. *Social Science Computer Review*, 41(6), 2188–2210.
- Lv, P., Zhang, Q., Shi, L., Guan, Z., Fan, Y., Li, J., Zhong, K., & Deveci, M. (2024). Exploring on role of location in intelligent news recommendation from data analysis perspective. *Information Sciences*, 662, 120213.
- Mahesh, N., Periwai, N., Kaur, N., Jayaram, N. M. P., Asharani, K. P., & Gowrishankar, S. (2022). A Study on Deep Learning based News Recommender Systems. 675–682. Scopus.
- Mahmood, B., Elahi, M., Touileb, S., Steskal, L., & Trattner, C. (2024). Incorporating Editorial Feedback in the Evaluation of News Recommender Systems. *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, 148–153.
- Manh Nguyen, C., Xuan Bach, N., & Minh Phuong, T. (2022). Combining User Specific and Global News Features for Neural News Recommendation. In N. T. Nguyen, T. K. Tran, U. Tukayev, T.-P. Hong, B. Trawiński, & E. Szczerbicki (Eds.), *Intelligent Information and Database Systems* (pp. 648–660). Springer International Publishing.
- Moghaddam, F. B., & Elahi, M. (2019). Cold-start solutions for recommendation systems. *Big Data Recommender Systems - Volume 2: Application Paradigms*.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., Götzsche, P. C., ... Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7). Scopus.
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. Scopus.
- Myers, D., Mohawesh, R., Chellaboina, V. I., Sathvik, A. L., Venkatesh, P., Ho, Y.-H., Henshaw, H., Alhawawreh, M., Berdik, D., & Jararweh, Y. (2024). Foundation and large language models: Fundamentals, challenges, opportunities, and social impacts. *Cluster Computing*, 27(1), 1–26.
- Pahune, S., & Chandrasekharan, M. (2023). Several categories of Large Language Models (LLMs): A Short Survey. *International Journal for Research in Applied Science and Engineering Technology*, 11(7), 615–633.

- Pourashraf, P., & Mobasher, B. (2023). Modeling Users' Localized Preferences for More Effective News Recommendation. In H. Degen & S. Ntoa (Eds.), *Artificial Intelligence in HCI* (pp. 366–382). Springer Nature Switzerland.
- Raza, S., & Ding, C. (2022). News recommender system: A review of recent progress, challenges, and opportunities. *Artificial Intelligence Review*, 55(1), 749–800.
- Ruan, Q., Mac Namee, B., & Dong, R. (2023). Unveiling the Relationship Between News Recommendation Algorithms and Media Bias: A Simulation-Based Analysis of the Evolution of Bias Prevalence. In M. Bramer & F. Stahl (Eds.), *Artificial Intelligence XL* (pp. 210–215). Springer Nature Switzerland.
- Sheu, H.-S., Chu, Z., Qi, D., & Li, S. (2022). Knowledge-Guided Article Embedding Refinement for Session-Based News Recommendation. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12), 7921–7927. *IEEE Transactions on Neural Networks and Learning Systems*.
- Shi, S., Ma, W., Wang, Z., Zhang, M., Fang, K., Xu, J., Liu, Y., & Ma, S. (2021). WG4Rec: Modeling Textual Content with Word Graph for News Recommendation. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 1651–1660.
- Shichang, Z. (2021). Research on Recommendation Algorithm Based on Collaborative Filtering. 2021 2nd International Conference on Artificial Intelligence and Information Systems, 1–4.
- Song, Z., Zhang, D., Shi, X., Li, W., Ma, C., & Wu, L. (2021). DEN-DQL: Quick Convergent Deep Q-Learning with Double Exploration Networks for News Recommendation. 2021 International Joint Conference on Neural Networks (IJCNN), 1–8.
- Sun, Y., Kong, Q., Huangfu, L., & Pan, J. (2021). Domain-oriented News Recommendation in Security Applications. 2021 IEEE International Conference on Intelligence and Security Informatics (ISI), 1–6.
- Symeonidis, P., Chaltsev, D., Zanker, M., & Manolopoulos, Y. (2021). News Recommendations by Combining Intra-session with Inter-session and Content-Based Probabilistic Modelling. In N. T. Nguyen, L. Iliadis, I. Maglogiannis, & B. Trawiński (Eds.), *Computational Collective Intelligence* (pp. 154–166). Springer International Publishing.
- Symeonidis, P., Kirjackaja, L., & Zanker, M. (2021). Session-based news recommendations using SimRank on multi-modal graphs. *Expert Systems with Applications*, 180, 115028.
- Vrijenhoek, S. (2023). Do You MIND? Reflections on the MIND Dataset for Research on Diversity in News Recommendations. In L. Boratto, S. Faralli, M. Marras, & G. Stilo (Eds.), *Advances in Bias and Fairness in Information Retrieval* (pp. 147–154). Springer Nature Switzerland.
- Wang, R., Wang, S., Lu, W., & Peng, X. (2022). News Recommendation Via Multi-Interest News Sequence Modelling. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 7942–7946.
- Wang, S., Guo, S., Wang, L., Liu, T., & Xu, H. (2023). Multi-interest Extraction Joint with Contrastive Learning for News Recommendation. In M.-R. Amini, S. Canu, A. Fischer, T. Guns, P. Kralj Novak, & G. Tsoumakas (Eds.), *Machine Learning and Knowledge Discovery in Databases* (pp. 606–621). Springer International Publishing.
- Wu, C., Wu, F., Huang, Y., & Xie, X. (2021). User-as-Graph: User Modeling with Heterogeneous Graph Pooling for News Recommendation. In Z. H. Zhou (Ed.), *PROCEEDINGS OF THE THIRTIETH INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, IJCAI 2021* (pp. 1624–1630). Ijcai-Int Joint Conf Artif Intell.
- Xu, Y., & Gu, W. (2022). Research on the Impact of Embedded Intelligent Robots on English News Dissemination. *Wireless Communications and Mobile Computing*, 2022(1), 2127411.
- Yang, J., Lin, J., Gui, F., Meng, H., Chen, H., & An, N. (2023). StoryLens: Personalizing News Recommendations for Older Adults with Their Life Stories. In Q. Gao, J. Zhou, V. G. Duffy, M. Antona, & C. Stephanidis (Eds.), *HCI International 2023 – Late Breaking Papers* (pp. 246–263). Springer Nature Switzerland.
- Yang, Z., Cui, L., Wang, X., Zhang, T., & Cheng, Y. (2023). MIAR: Interest-Activated News Recommendation by Fusing Multichannel Information. *IEEE Transactions on*

- Computational Social Systems, 10(6), 3433–3443. IEEE Transactions on Computational Social Systems.
- Zhang, Q., Li, J., Jia, Q., Wang, C., Zhu, J., Wang, Z., & He, X. (2021). UNBERT: User-News Matching BERT for News Recommendation. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, 3356–3362.
- Zhao, M., Liu, Z., Yu, M., Zhang, W., Zhao, Y., Yang, M., & Yu, J. (2023). Hypergraph Enhanced Contrastive Learning for News Recommendation. In Z. Jin, Y. Jiang, R. A. Buchmann, Y. Bi, A.-M. Ghiran, & W. Ma (Eds.), Knowledge Science, Engineering and Management (pp. 136–147). Springer Nature Switzerland.
- Zhao, P., Wang, S., Lu, W., Peng, X., Zhang, W., Zheng, C., & Huang, Y. (2023). News Recommendation via Jointly Modeling Event Matching and Style Matching. In D. Koutra, C. Plant, M. Gomez Rodriguez, E. Baralis, & F. Bonchi (Eds.), Machine Learning and Knowledge Discovery in Databases: Research Track (pp. 404–419). Springer Nature Switzerland.
- Zhao, Q., Chen, X., Zhang, H., & Li, X. (2024). Dynamic Hierarchical Attention Network for news recommendation. Expert Systems with Applications, 255, 124667.
- Zhu, P., Cheng, D., Luo, S., Yang, F., Luo, Y., Qian, W., & Zhou, A. (2022). SI-News: Integrating social information for news recommendation with attention-based graph convolutional network. Neurocomputing, 494, 33–42.

Appendix

Appendix A

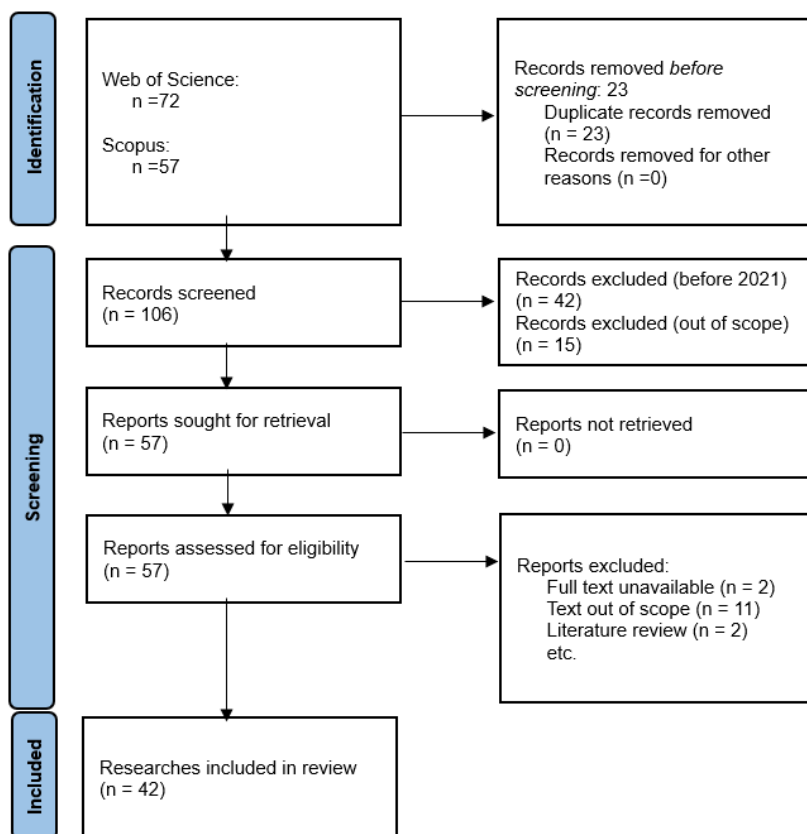


Figure 1: PRISMA Flow Diagram

Appendix B

Table 1: Advantages and Limitations of Session-Based Systems

Session Based Systems			
Advantages		Limitations	
Advantage	Paper	Limitation	Paper
<i>Short-Term Preference Modeling:</i> Session-based systems excel at capturing users' short-term interests and intentions by analyzing their interactions within a single session. This is crucial in the news domain, where user interests can change rapidly.	Sheu et al., 2022	<i>Limited user history:</i> Session-based systems primarily focus on the current session and often neglect long-term user preferences they may struggle to model or incorporate long-term user interests that span multiple sessions	Pourashraf & Mobasher, 2023
	Symeonidis, Chaltsev, et al., 2021		Symeonidis, Chaltsev, et al., 2021
	Pourashraf & Mobasher, 2023		
<i>Handling Anonymity:</i> Many users browse news websites without logging in. Session-based systems don't require user profiles. They rely on the sequence of items a user interacts with during a session, making them suitable for anonymous users.	Sheu et al., 2022	<i>Potential for Over-Personalization:</i> By focusing heavily on the immediate session, these systems may lead to recommendations that are too narrow or repetitive, potentially creating a "filter bubble" effect	Alam et al., 2022
	Alam et al., 2022		
	Lucas & de Figueiredo, 2023		
	Symeonidis, Chaltsev, et al., 2021		
<i>Timeliness:</i> Session-based methods can quickly recommend relevant news stories, reducing the time lag between breaking news and recommendations. These systems are effective at recommending news items due to the importance of recency and freshness.	Lucas & de Figueiredo, 2023	<i>Data Sparsity:</i> While session-based systems are good at handling anonymous users and new items, they can still face data sparsity issues, especially with very short sessions or when dealing with new content that has not been viewed by many users	Lucas & de Figueiredo, 2023
	Pourashraf & Mobasher, 2023		Sheu et al., 2022
<i>Reduced Cold-Start Problems:</i> Session-based methods can make recommendations based on a single session.	Lucas & de Figueiredo, 2023	<i>Contextual Limitations:</i> While session-based systems consider the immediate context of a user's session, they may not fully account for external factors or broader contexts, such as location, time of day, or current events, which can influence user preferences	Lucas & de Figueiredo, 2023

Session Based Systems			
Advantages		Limitations	
<i>Advantage</i>	<i>Paper</i>	<i>Limitation</i>	<i>Paper</i>
<i>Adaptability:</i> Session-based methods adapt to changes in user preferences as they occur, continuously updated with every new user click, making them sensitive to changes in preferences	Symeonidis, Chaltsev, et al., 2021	<i>Lack of Semantic Understanding:</i> Some session-based models may not fully exploit the semantic-level structure information of news article. They might focus on the sequence of interactions without capturing deeper semantic relationships between articles.	Sheu et al., 2022
<i>Multi-Interest Capture:</i> Some session-based systems are capable of capturing multiple interests within a single session	R. Wang et al., 2022		
<i>Use of Meta-Paths:</i> Session-based systems can use meta-paths to define relationships between items such as "article-session-article". These paths reveal semantic context and enhance the accuracy and explainability of recommendations	Symeonidis et al., 2023		

Appendix C

Table 2: Summary of Graph-Based Approaches in NRS

Approach	Description	Role of AI	References
Knowledge Graphs	Capture semantic context but face computational complexity.	Semantic modeling, preference propagation	Y. Wang et al. (2022), Alam et al. (2022)
User-News Graphs	Model detailed user-news interactions; effective yet complex.	User/news embedding, detailed behavior modeling	Wu et al. (2021), Li et al. (2023)
Hypergraphs	Represent complex, higher-order interactions; computationally intensive.	Complex user-news-topic interaction modeling	W. Liu et al. (2024), M. Zhao et al. (2023)
Session Graphs	Track short-term user behaviors dynamically; limited for long-term preferences.	Dynamic behavior analysis, temporal sensitivity	Symeonidis et al. (2021, 2023)
Graph Convolutional Networks	Embed users and news effectively but with high computational requirements.	Node embedding, interest representation	M. Zhao et al. (2023), Zhu et al. (2022)

Appendix D

Table 3: Advantages and Limitations of Graph-Based Systems

Graph Based Systems			
Advantages		Limitations	
<i>Advantage</i>	<i>Paper</i>	<i>Limitation</i>	<i>Paper</i>
<i>Improved User and News Representation:</i> GNNs can learn better representations of both users and news by propagating information through the graph structure. Graph structures help to prepare better hybrid systems refining article embeddings and user embeddings.	Li et al., 2023	<i>Data Sparsity and Cold Start:</i> While graph structures help alleviate data sparsity issues, they do not completely solve the problem. While GNNs used as CF model, the model can suffer from cold-start problem.	Li et al., 2023
	W. Liu et al., 2024		X. Liu, 2022
	Sheu et al., 2022		Song et al., 2021
	S. Wang et al., 2023		Han, 2024
	Y. Wang et al., 2022		
	Wu et al., 2021		
	Shi et al., 2021		
<i>Capturing High-Order Relationships:</i> Some graph structures like Hypergraph can model relationships beyond simple pairwise connections between users and news, including triadic, tetradic, or higher-order relations	W. Liu et al., 2024	<i>Over-smoothing:</i> In graph neural networks (GNNs), increasing the number of propagation layers can lead to over-smoothing, where the node embeddings become too similar and difficult to distinguish.	M. Zhao et al., 2023
<i>Encoding Complex Relationships:</i> Graph-based methods can incorporate various types of relationships, such as semantic similarity, co-occurrence, co-click.	Shi et al., 2021	<i>Computational Complexity:</i> Some models, like implicit semantic and matrix decomposition models, suffer from decreased computational performance when the amount of data is large	Lucas & de Figueiredo, 2023
			X. Liu, 2022
<i>Addressing Data Sparsity:</i> By modeling relationships in a graph, GNNs can alleviate the sparsity of user-item interactions. This is particularly helpful for addressing the cold start problem.	Li et al., 2023	<i>Hyperparameter Sensitivity:</i> The performance of some graph based systems can be sensitive to the choice of hyper parameters, for example increasing the number of layers in GNNs can improve performance up to a point, after that performance may decrease.	Li et al., 2023
	W. Liu et al., 2024		Y. Wang et al., 2022
	Ludwig et al., 2023		

Graph Based Systems			
Advantages		Limitations	
Advantage	Paper	Limitation	Paper
<i>Utilizing Heterogeneous Information:</i> Graphs can easily incorporate different types of nodes and edges, representing different aspects of news and user behavior like clicked news, topics and entities.	Wu et al., 2021	<i>Difficulty in capturing complex semantics:</i> Some methods have difficulty using hypergraphs to capture the potential relationships between news due to the rich textual semantics.	W. Liu et al., 2024
<i>Improved Diversity:</i> Graph-based approaches have shown improvements in the diversity of recommendations, wich is important to avoid filter bubbles.	Symeonidis, Kirjackaja, et al., 2021		
	Fieiras-Ceide et al., 2023		
<i>Flexibility:</i> Graph based systems are flexible and can combine information from different sources, including cross-platform data and user interactions	Shi et al., 2021		

Appendix E

Table 4: DNN Approaches and their Roles in NRS

Approach	Description	Role of AI	References
Convolutional Neural Networks (CNN)	Extract features from textual sequences (e.g., news headlines) to improve recommendations.	Semantic feature extraction, sequence modeling	Lu et al. (2022), Manh Nguyen et al. (2022), Pourashraf & Mobasher (2023)
Recurrent Neural Networks (RNN)	Model sequential user interactions; capture dynamic preferences (includes LSTM, GRU).	Sequential modeling, personalization	Manh Nguyen et al. (2022), Lu et al. (2022), Pourashraf & Mobasher (2023)
Deep Matrix Factorization	Uses deep layers to process user-item interactions, enhancing rating prediction.	Non-linear user-item interaction modeling	Li et al. (2023), Zhu et al. (2022)
Transformers	Employ encoder-decoder structures to model contextual relationships in data.	Contextual embedding, semantic understanding	Zhang et al. (2021)

Approach	Description	Role of AI	References
Attention Mechanisms (Word, News, User Level)	Assign weights to important words/news/user interactions, enhancing relevance.	Context-sensitive modeling, personalization	Manh Nguyen et al. (2022), Z. Yang et al. (2023), Fu (2023), Li et al. (2023)
Multi-Head Attention Networks	Attend multiple input parts simultaneously, capturing diverse contextual aspects.	Multi-aspect representation learning	Hou et al. (2023), Iana et al. (2023), W. Liu et al. (2024)
Deep Reinforcement Learning (DRL)	Optimizes recommendations based on long-term user engagement through interactions.	Adaptive long-term recommendation optimization	Aboutorab et al. (2023), Song et al. (2021)
Deep Knowledge-Aware Networks (DKN)	Integrate semantic and knowledge-level information to enhance recommendation accuracy.	Semantic and knowledge integration	Lim et al. (2022), Lu et al. (2022), Y. Wang et al. (2022), Feng & Lv(U) (2022)
Deep Structured Semantic Models (DSSM)	Rank documents based on semantic similarity using deep neural architectures.	Semantic similarity modeling, ranking	Zhu et al. (2022)

Table 5: Advantages and Limitations of Deep Learning Based Systems

Deep Learning Based			
Advantages		Limitations	
Advantage	Paper	Limitation	Paper
Powerful learning ability: Deep learning algorithms can automatically learn features from large datasets and extract high-level semantic expressions.This allows the system to understand complex relationships in data without manual feature engineering.	Han, 2024	Timeliness and Dynamic Nature of News: The rapidly changing nature of news articles, with new articles constantly published and older ones becoming outdated, presents a challenge for deep learning systems.The models need to adapt to the short shelf-life of news and the dynamic nature of user interests. News is highly time-sensitive, requiring models to capture real-time user interests	Lucas & de Figueiredo, 2023
	Z. Yang et al., 2023		Fu, 2023
	Zhu et al., 2022		Zhu et al., 2022
Representation Learning: Deep learning models can produce good representations for news metadata, which is a key component in content-based news recommendation. Traditional methods often rely on categorical features (e.g. news IDs, news categories) or bag-of-words (tokens or n-grams) which do not capture the complexities of language as well as deep learning approaches.	Manh Nguyen et al., 2022		
	Zhang et al., 2021		
Processing multiple data types: Deep learning models can handle various types of input data, such as text, images, audio, and video.	Han, 2024		
Fine-Grained Information: Deep learning methods are capable of capturing fine-grained aspect-level information, leading to more accurate recommendations	Lu et al., 2022		

Deep Learning Based			
Advantages		Limitations	
Advantage	Paper	Limitation	Paper
Non-linear relations: Deep learning models are capable of extracting non-linear relationships and dependencies between news articles more effectively than traditional methods	Zhu et al., 2022		
Mitigating Cold-Start and Data Sparsity: Deep learning models can extract valuable features from news and user data to improve user and item profiles, thus alleviating the cold-start and data sparsity issues often encountered in recommender systems	Manh Nguyen et al., 2022		
Improved Accuracy: Deep learning algorithms can enhance the precision and personalization of recommendations by analyzing user behavior and creating user models, making it easier for users to find relevant news	Zhu et al., 2022		
	Han, 2024		

Appendix G

Table 6: Advantages and Limitations of Multi Interest Based Systems

Multi Interest Based			
Advantages		Limitations	
<i>Advantage</i>	<i>Paper</i>	<i>Limitation</i>	<i>Paper</i>
<i>Improved User Representation:</i> Multi-interest models capture the diverse and multi-grained nature of user interests. Unlike methods that learn a single interest vector for each user, multi-interest models can represent a user's interest in multiple topics such as sports, movies, and finance.	R. Wang et al., 2022	Relying solely on sequential behavior may not be accurate because users sometimes click on incorrect or unrelated news out of curiosity	Fu, 2023
	Li et al., 2023		
	Hou et al., 2023		
<i>Better Handling of User Dynamics:</i> These models can capture both short-term and long-term user interests, accommodating the evolving nature of user preferences over time	Z. Yang et al., 2023	There is a possibility of introducing noise and redundancy; when the number of interests (K) becomes too large, a lot of interest-level noise information will be introduced while aggregated into the final user representations, which is likely to be detrimental to the recommendation performance	Hou et al., 2023
	Fu, 2023		S. Wang et al., 2023
<i>Enhanced Recommendation Performance:</i> By modeling user interests at multiple granularities (e.g., word-level, news-level, and higher-levels), multi-interest models can more accurately represent diverse user interests, leading to improved recommendation performance.	Hou et al., 2023	They aggregate relevant information of news (e.g., title, content, topic) into a simple and unified news presentation and therefore miss the influence of some important factors such as news event and news style, which are important for guiding users' preferences toward news.	P. Zhao et al., 2023
<i>Capture of Low-Level and High-Level Interests:</i> Multi-interest models can capture both low-level (e.g., specific words or phrases) and high-level user interests, allowing for a more comprehensive understanding of user preferences	Hou et al., 2023		

RESEARCH IN PROGRESS

EXPLORING THE RELATIONSHIP BETWEEN PERSONAL VALUES AND PERCEIVED OBSOLESCENCE OF HIGH-TECH PRODUCTS

MEHMET CEM BÖLEN,¹ BURAK BORULU,¹
FULYA ACIKGOZ²

¹ Ataturk University, Faculty of Open and Distance Education, Erzurum, Turkey
mehmetcem.bolen@atauni.edu.tr, burak.borulu@atauni.edu.tr

² University of Sussex, Marketing, Brighton, United Kingdom of Great Britain and
Northern Ireland
f.acikgoz@sussex.ac.uk

This study explores the relationship between personal values (biospheric, altruistic, egoistic, and hedonic) and perceived obsolescence of high-tech consumer products. Obsolescence, characterized by the loss of product value, functionality, or desirability, is a significant concern in the consumer technology sector, influencing product life cycles and consumer demand. The study proposes that the degree of perceived obsolescence may differ based on individuals' personal values. For example, individuals who strongly endorse biospheric values may be less likely to replace their smartphones, while those who strongly endorse egoistic values may be more likely to do so. The study employs "Multivariate Analysis of Variance" to examine the relationship between personal values and the perceived obsolescence types (economic, functional, technological, and aesthetic). A representative sample of individuals who have voluntarily replaced their working smartphones will be recruited, and personal values will be measured using a validated value scale. The study expects to find significant differences in the perceived importance of obsolescence types among individuals with varying personal values. The findings may help manufacturers adapt their strategies to effectively meet consumer demands and foster long-term relationships with their target audience while addressing the environmental concerns associated with rapid product obsolescence.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.43](https://doi.org/10.18690/um.fov.4.2025.43)

ISBN

978-961-286-998-4

Keywords:

obsolescence,
high-tech products,
personal values,
manova,
product life cycle



University of Maribor Press

1 Introduction

In an era dominated by rapid technological advancement, the notion of obsolescence looms significantly, casting a shadow over the lifecycle of high-tech products. While technological progress brings forth innovation and convenience, it also fosters a sense of transience, whereby products once deemed cutting-edge swiftly become outdated. Within this dynamic landscape, understanding the phenomenon of perceived obsolescence has emerged as a critical pursuit that shapes consumer behaviors, market dynamics, and product development strategies (Guillard et al., 2023). Perceived obsolescence refers to the psychological aspect of a product becoming outdated, leading consumers to feel the need to upgrade to newer versions (Khatriwip & First, 2012). This perception drives consumers to constantly reevaluate product qualities, devalue existing products, and contribute to the disposal of functioning items (Spinney et al., 2012). Consumers often base their disposal decisions on the perceived obsolescence of a product, which is influenced by psychographic variables, technology trends, and market characteristics (Khatriwip & First, 2012). In conclusion, perceived obsolescence plays a significant role in shaping consumer behavior by influencing attitudes, purchasing decisions, and product disposal practices.

This study examines how individuals' personal values shape perceived obsolescence of high-tech consumer products. Specifically, we present new theorizing on how biospheric, altruistic, egoistic and hedonic values and perceived obsolescence are interrelated in the context of smartphones. Accordingly, we propose that the degree of perceived obsolescence regarding smartphones may differ based on the individuals' biospheric, altruistic, egoistic and hedonic values. For example, individuals who strongly endorse biospheric values are typically less likely to replace their smartphones due to aesthetics related obsolescence reasons, whereas individuals who strongly endorse egoistic values are typically more likely to replace their smartphones due to aesthetics related obsolescence reasons. By exploring the impact of personal values on perceptions regarding obsolescence of smartphones, manufacturers can adapt their strategies to effectively meet consumer demands and foster long-term relationships with their target audience.

Accordingly, this study seeks to answer the following research question:

- How do personal values (biospheric, altruistic, egoistic, and hedonic) influence individuals' perceptions of different types of obsolescence (economic, functional, technological, and aesthetic) in high-tech consumer products such as smartphones?

2 Theoretical Framework

2.1 Defining Obsolescence: A Conceptual Overview

Obsolescence refers to the state of a product or technology that is no longer useful or functional, often characterized by the loss of product value as social and symbolic currency, unfavorable cost-benefit expenditure to extend product life, or loss of product functionality. According to Packard (1960), a product becomes obsolete when it is still available for use, but it becomes "worn out" in the mind and starts to seem less desirable. This decline in utility occurs even if the product has a good shape and a good working order. According to Sandborn (2007), obsolescence is the inevitable end of manufacturing and is characterized as a weakness. He argued that eventually, the production of each product's spare parts would cease due to a lack of demand, the manufacturer's decision not to produce them anymore, or the inability to access raw materials. For instance, VHS tape is a classic example of a product that has become both functionally and technologically obsolete. Before the emergence of DVDs and online streaming services, VHS tapes were the primary means of watching movies at home. However, as technological advances and better alternatives emerged, the VHS format became increasingly outdated and less relevant. Today, it is rare to find VHS tapes in stores, and most people have switched to digital streaming or DVD/Blu-ray discs for their home entertainment needs, making this a prime example of obsolescence. This phenomenon occurs when a product or technology becomes outdated and less relevant due to the emergence of newer and better alternatives (Amankwah-Amoah, 2017).

Voluntary or involuntary actions can lead to obsolescence in several ways (Amankwah-Amoah, 2017; Mellal, 2020). In the literature, several types have been identified, which are summarized in Table 1. Although there are different opinions on the diversity of obsolescence types in the studies on this subject, it is stated in

(Acikgoz et al., 2024) study that only four types are more effective in high-tech products. For this reason, in this study, these four obsolescence factors will be discussed since only cell phones will be examined.

Table 1: Typology of Obsolescence

Dimension	Definition	Reference
Technological obsolescence	occurs when advancements in technology make older product less desirable or older product is overshadowed by newer, more advanced technology.	(Barros & Dimla, 2021)
Economic obsolescence	happens when external factors such as changes in market conditions or regulations make a product less financially advantageous.	(Trabelsi et al., 2021)
Functional obsolescence	occurs when a product is still operational but no longer meets current needs or standards. It may also occur when a product or system no longer performs its intended function effectively, even if it's still operational.	(Bradley & Dawson, 1998; Mellal, 2020)
Environmental obsolescence	occurs when environmental factors, such as increased awareness of sustainability and eco-friendly practices, can lead to obsolescence of products with high energy consumption, limited recyclability, or environmentally unfriendly components.	(Trabelsi et al., 2021)
Cultural/social obsolescence	occurs when products become obsolete due to shifts in cultural or societal norms. Changes in lifestyle trends, values, or social perceptions can impact the demand for certain high-tech products, especially those associated with specific cultural or social contexts.	(Guillard et al., 2023; Wilson et al., 2017)
Aesthetical obsolescence	occurs when products become outdated owing to changes in design trends or aesthetics. Consumer preferences for specific styles or visual appearances can impact on the desirability of products, leading to a decline in the demand for older models.	(Guillard et al., 2023); (Lilley et al., 2016)
Logistical obsolescence	refers to the state in which a product or technology becomes less desirable or useful owing to changes or limitations in its logistical support system. This can include challenges in sourcing replacement parts, obtaining technical support, and accessing compatible software or infrastructures.	(Trabelsi et al., 2021)
Psychological obsolescence	occurs when a product is perceived as outdated or no longer desirable by consumers, not because it lacks functionality or has been surpassed by newer technology, but simply because of changing trends, aesthetics, or marketing influences. This revolves around the idea that consumers are constantly enticed by the latest trends and innovations, creating a sense of dissatisfaction or inadequacy with their current possessions. Marketers often harness this phenomenon to instill a desire for the newest products and create a perception of obsolescence in the minds of consumers, prompting them to seek constant upgrades or replacements.	(Cooper, 2004; Hagedorn et al., 2018)

2.2 Values and Perceived Obsolescence

Personal values are stable, universally desired goals that transcend specific situations, with individuals sharing the same basic values but differing in the intensity and priority they assign to each (Bauman et al., 2021). Four value types—biospheric, altruistic, egoistic, and hedonic—are particularly critical in the environmental domain (De Groot & Steg, 2007, Schultz & Zelezny, 1999; Steg et al., 2014; Stern, Dietz, & Guagnano, 1998). The interplay of biospheric, altruistic, egoistic, and hedonic values significantly shape consumer perceptions of obsolescence in consumer tech products. Each value type contributes to a complex decision-making process that influences whether consumers feel compelled to upgrade or resist the pressures of obsolescence related reasons. For example, hedonic motivations are found to have a significant effect on consumer satisfaction and purchase intentions since people try to select products that enhance their enjoyment and lifestyle (Gan & Wang, 2017). Therefore, the enjoyment in the use of the latest gadgets may create the perception of obsolescence as consumers may not feel that the older models are as satisfying or exciting (Ozturk et al., 2023). Furthermore, consumers tend to assess products based on the consumers' utilization of the product and the consumers' assessment of the worth of the benefits received. Egoistic motivations are known to include a perceived loss of value in the current product which can result in the desire to upgrade to a new product when a new product is introduced (Levinthal & Purohit, 1989). This is especially the case in technology markets where people may feel that they must upgrade to the latest devices to keep up with the social status or simply to increase their personal satisfaction (Widarmanti et al., 2024).

On the other hand, individuals with strong biospheric orientations may have more concerns about products that are made to be used for a short period of time because they think about the negative impact on the environment through waste and natural resource depletion. The literature also indicates that those who have biospheric values as their norm are more likely to participate in environmentally friendly actions like refusing products that may harm the environment through the planned obsolescence strategy (De Groot and Steg, 2006). This indicates that biospheric values can lead to a heightened awareness of product lifespan and sustainability, affecting consumer choices.

In the consumer technology sector, obsolescence risk is a significant concern that impacts product life cycles and consumer demand (Jennings et al., 2016). It influences consumers' decisions on whether to replace or repair products, as it indicates the actual or perceived loss of value of the products (Sonego et al., 2022). Previous studies have noted that high-tech industries face rapid obsolescence due to factors such as increased competition, changing customer needs, and short product life cycles (Goktan & Miles, 2011; Pangburn & Sundaresan, 2009).

3 Methodology

This study will employ Multivariate Analysis of Variance (MANOVA) as the primary analytical method to examine the relationship between personal values (biospheric, altruistic, egoistic, and hedonic) and the perceived importance of different types of obsolescence (economic, functional, technological, and aesthetic). MANOVA is particularly suitable for this research as it allows for the simultaneous analysis of multiple dependent variables, helping to identify whether individuals with different personal value orientations prioritize specific obsolescence types differently. MANOVA is an appropriate assessment to investigate “groups of subjects on several dependent variables simultaneously; focusing on cases where the variables are correlated and share a common conceptual meaning” (Stevens, 2002, p. 173). Since participants could rate the importance of multiple obsolescence reasons on a Likert scale, MANOVA enables the detection of significant differences across personal value groups while accounting for potential interdependencies between obsolescence perceptions. This approach provides a comprehensive understanding of how personal values influence perceptions of smartphone obsolescence beyond what univariate analyses can offer. Accordingly, we propose the following main hypothesis:

H1: There is a significant difference in the perceived importance of obsolescence types (economic, functional, technological, aesthetic) based on personal values (biospheric, altruistic, egoistic, hedonic).

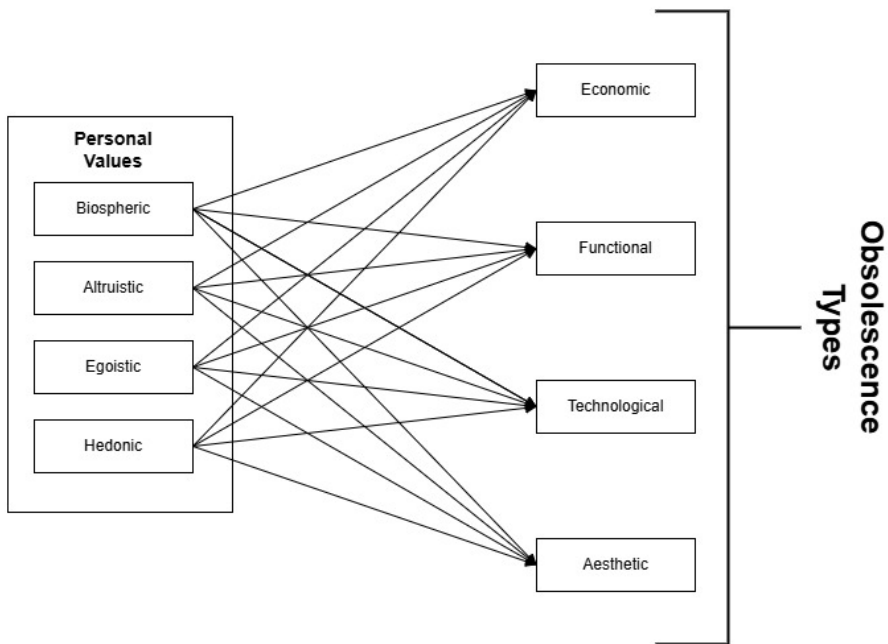


Figure 1: Theoretical model

Table 2: Value Scale

Hedonic values
1. Pleasure
2. Enjoying life
3. Gratification for oneself
Egoistic values
4. Social power
5. Wealth
6. Authority
7. Influential
8. Ambitious
Altruistic values
9. Equality
10. A world at peace
11. Social justice
12. Helpful
Biospheric values
13. Respecting the earth
14. Unity with nature
15. Protecting the environment
16. Preventing pollution

The study will recruit a representative sample of individuals who are currently using a smartphone. These individuals must have decided to replace their smartphones voluntarily, even though they are still in working condition. The decision to replace the product should also not be based on products that are not of the person's choice, such as gifts or rewards. Interviewees are also expected to be over 18 years of age and have sufficient economic freedom to replace their old device. People who meet these criteria will be allowed to see the questions of the online survey. Otherwise, the survey will be terminated for those who do not meet the criteria. In addition, convenience sampling will be conducted in order to collect data with a distribution similar to Turkey's demographic data (age, gender, educational status and income) distribution.

Table 3: Obsolescence Scale

<i>Technological Obsolescence</i>
The technology was outdated (Wilson et al.,2017).
The phone was not compatible anymore with the latest devices and software (Wieser & Tröger, 2018).
It is less efficient than the latest models (Guillard et al., 2023).
There are more modern products that I could use better (Guillard et al., 2023).
<i>Economic Obsolescence</i>
It cost too much money to repair (if broken) (Wilson et al.,2017).
I was offered a free/discounted upgrade in my current contract (Wilson et al.,2017).
I wanted a different contract with better cost value (Wilson et al.,2017).
<i>Functional Obsolescence</i>
It didn't have the specific functions that I wanted (Wilson et al.,2017).
I wanted a different contract with better features (Wilson et al.,2017).
<i>Aesthetic Obsolescence</i>
It was no longer novel, stylish or prestigious (Wilson et al.,2017).
It was no longer clean, shiny, or new (Wilson et al.,2017).
The phone made me look oldfashioned (Wieser & Tröger, 2018)
The phone had a lot of scratches (Wieser & Tröger, 2018)

Personal values were measured using a value scale (Table 2) developed and validated by Steg, Perlaviciute, van der Werff, and Lurvink (2012). Participants evaluated the significance of values "as a guiding principle in their lives" using a 9-point scale, ranging from -1 (opposed to my guiding principles), 0 (not important at all), to 7 (extremely important). The constructs related to obsolescence types (Table 3) adopted in this study will be adapted from Wilson et al. (2017), and a 9-point Likert scale ranging from -1 (opposed to my guiding principles), 0 (not important at all), to 7 (extremely important) will be employed to measure the related items. Values were measured in the same way in present studies. If the same values demonstrate

the expected relationship across studies, irrespective of the various methods for measuring the dependent variables, that would provide strong support that the relationship indeed exists and is not merely an artifact of a specific measurement used (Perlaviciute & Steg, 2015).

4 Preliminary/Expected Results

We believe that the findings of this study is likely to reveal significant differences in the perceived importance of different types of obsolescence (economic, functional, technological, aesthetic) among individuals with varying personal values (biospheric, altruistic, egoistic, hedonic). Consistent with consumer behavior theories and prior research on personal values and consumption patterns, we expect the following results:

- Egoistic individuals rate aesthetic obsolescence significantly higher than biospheric and altruistic individuals.
- Biospheric individuals rate functional obsolescence significantly higher than hedonic and egoistic individuals.
- Altruistic individuals rate economic obsolescence significantly higher than egoistic and hedonic individuals.
- Hedonic individuals rate technological obsolescence significantly higher than biospheric and altruistic individuals

References

- Acikgoz, F., Borulu, B. and Bölen, M.C. (2024), "How does obsolescence risk influence consumer resistance to smartwatches?", *Information Technology & People*, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/ITP-08-2023-0843>
- Amankwah-Amoah, J. (2017). An integrative process model of organizational failure. *Journal of Business Research*, 79, 338-350. <https://doi.org/10.1016/j.jbusres.2017.06.053>.
- Bauman, Z., Beck, U., Giddens, A., & Lash, S. (2021). *Liquid modernity*. Polity Press.
- Barros, M., & Dimla, E. (2021). From planned obsolescence to the circular economy in the smartphone industry: An evolution of strategies embodied in product features. *Proceedings of the Design Society*, 1(AUGUST), 1607–1616. <https://doi.org/10.1017/pds.2021.422>
- Bradley, M., & Dawson, R. J. (1998). An analysis of obsolescence risk in IT systems. *Software Quality Journal*, 7(2), 123–130. <https://doi.org/10.1023/A:1008808708860>
- Cooper, T. (2004). Inadequate Life? Evidence of Consumer Attitudes to Product Obsolescence. *Journal of Consumer Policy*, 421–449.

- De Groot, J. I. M., & Steg, L. (2006). The role of value orientations in evaluating quality of life consequences of a transport pricing policy. *Transportation Research Part D: Transport and Environment*, 11(2), 160-165.
- De Groot, J. I. M., & Steg, L. (2007). Value orientations and environmental beliefs in five countries: Validity of an instrument to measure egoistic, altruistic and biospheric value orientations. *Journal of Cross-Cultural Psychology*, 38(3), 318-332. <https://doi.org/10.1177/0022022107300276>
- Gan, C. and Wang, W. (2017). The influence of perceived value on purchase intention in social commerce context. *Internet Research*, 27(4), 772-785. <https://doi.org/10.1108/intr-06-2016-0164>
- Goktan, A. B., & Miles, G. (2011). Innovation speed and radicalness: Are they inversely related? *Management Decision*, 49(4), 533–537. <https://doi.org/10.1108/00251741111126477>
- Guillard et al., 2023. A typology of consumers regarding perceived obsolescence: The paradox of eco-conscious consumers. *Journal of Cleaner Production*, 412. <https://doi.org/10.1016/j.jclepro.2023.137202>
- Hagedorn, L., Buchert, T., & Stark, R. (2018). Empirical study on aesthetics as an influencing factor on sustainability. 2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017 - Proceedings, 2018-Janua, 776–783. <https://doi.org/10.1109/ICE.2017.8279963>
- Jennings, C., Wu, D., & Terpenney, J. (2016). Forecasting obsolescence risk and product life cycle with machine learning. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 6(9), 1428–1436. <https://doi.org/10.1109/TCPMT.2016.2589206>
- Khetriwal, D. S., & First, I. (2012). Enabling Closed Resource Loops in Electronics: Understanding Consumer Disposal Behaviour using insights from Diffusion models. *Economic Research-Ekonomska Istraživanja*, 25(sup2), 47–68. <https://doi.org/10.1080/1331677X.2012.11517573>
- Levinthal, D. A., & Purohit, D. (1989). Durable Goods and Product Obsolescence. *Marketing Science*, 8(1), 35–53. <https://doi.org/10.1287/mksc.8.1.35>
- Lilley, D., Smalley, G., Bridgens, B., Wilson, G. T., & Balasundaram, K. (2016). Cosmetic obsolescence? User perceptions of new and artificially aged materials. *Materials and Design*, 101, 355–365. <https://doi.org/10.1016/j.matdes.2016.04.012>
- Mellal, M. A. (2020). Obsolescence – A review of the literature. *Technology in Society*, 63(June), 101347. <https://doi.org/10.1016/j.techsoc.2020.101347>
- Ozturk, A., Pizam, A., Hacikara, A., An, Q., Chaulagain, S., Balderas-Cejudo, A., ... & State, O. (2023). Hotel customers' behavioral intentions toward service robots: the role of utilitarian and hedonic values. *Journal of Hospitality and Tourism Technology*, 14(5), 780-801. <https://doi.org/10.1108/jhtt-07-2022-0223>
- Packard, V. (1960). *The Waste Makers*. D. McKay Co.
- Pangburn, M. S., & Sundaresan, S. (2009). Capacity decisions for high-tech products with obsolescence. *European Journal of Operational Research*, 197(1), 102–112. <https://doi.org/10.1016/j.ejor.2008.05.026>
- Perlavičiute, G. & Steg, L., 2015. "The influence of values on evaluations of energy alternatives," *Renewable Energy*, Elsevier, vol. 77(C), pages 259-267.
- Sandborn, P. A. (2007). Software obsolescence - Complicating the part and technology obsolescence management problem. *IEEE Transactions on Components and Packaging Technologies*, 30(4), 886–888. <https://doi.org/10.1109/TCAPT.2007.910918>
- Schultz, P. W., & Zelezny, L. (1999). Values as predictors of environmental attitudes: Evidence for consistency across 14 countries. *Journal of environmental psychology*, 19(3), 255-265.
- Sonego, M., Echeveste, M. E. S., & Debarba, H. G. (2022). Repair of electronic products: Consumer practices and institutional initiatives. *Sustainable Production and Consumption*, 30(2010), 556–565. <https://doi.org/10.1016/j.spc.2021.12.031>

- Spinney, J., Burningham, K., Cooper, G., Green, N., & Uzzell, D. (2012). What I've found is that your related experiences tend to make you dissatisfied: Psychological obsolescence, consumer demand and the dynamics and environmental implications of de-stabilization in the laptop sector. *Journal of Consumer Culture*, 12(3), 347–358.
<https://doi.org/10.1177/1469540512456928>
- Steg, L., Bolderdijk, J. W., Keizer, K., & Perlaviciute, G. (2014). An integrated framework for encouraging pro-environmental behaviour: The role of values, situational factors and goals. *Journal of Environmental Psychology*, 38, 104–115.
<https://doi.org/10.1016/j.jenvp.2014.01.002>
- Steg, L., Perlaviciute, G., Van der Werff, E., & Lurvink, J. (2014). The significance of hedonic values for environmentally relevant attitudes, preferences, and actions. *Environment and behavior*, 46(2), 163–192.
- Stern, P. C., Dietz, T., & Guagnano, G. A. (1998). A brief inventory of values. *Educational and Psychological Measurement*, 58(6), 984–1001.
<https://doi.org/10.1177/0013164498058006008>
- Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (Vol. 4). Mahwah, NJ: Lawrence Erlbaum Associates.
- Trabelsi, I., Zolghadri, M., Zeddini, B., Barkallah, M., & Haddar, M. (2021). Prediction of obsolescence degree as a function of time: A mathematical formulation. *Computers in Industry*, 129, 103470. <https://doi.org/10.1016/j.compind.2021.103470>
- Widarmanti, T., Ramantoko, G., PILLAI, S., & Rachmawati, I. (2024). Towards a unified model of planned obsolescence and innovation adoption in consumer behavior: a literature review and conceptual proposition using the stimulus-organism-response framework. *Management and Production Engineering Review*, 15 (2), 25–41. <https://doi.org/10.24425/mper.2024.151128>
- Wilson, G. T., Smalley, G., Suckling, J. R., Lilley, D., Lee, J., & Mawle, R. (2017). The hibernating mobile phone: Dead storage as a barrier to efficient electronic waste recovery. *Waste Management*, 60, 521–533. <https://doi.org/10.1016/j.wasman.2016.12.023>

RESEARCH IN PROGRESS

A SYSTEMATIC APPROACH TO DEFINE DIGITAL WELL-BEING AT WORK

NADIA BIJ DE VAATE, DIAN DE VRIES,
MACHIEL BOUWMANS

HU Utrecht University of Applied Sciences, Utrecht, the Netherlands
nadia.bijdevaate@hu.nl, dian.devries@hu.nl, machiel.bouwmans@hu.nl

The advancements of digital technologies continuously reshape how, where, and when employees work. While such digitalization can bring benefits for employees, it may also impose a significant burden on employee well-being. Research on how digitalization in the workplace impacts employee well-being is highly fragmented due to the myriad of digital technologies studied, as well as the vast number of well-being indicators examined. The wide variety of conceptualizations and operationalizations for both digitalization and well-being indicators hinders the ability to draw concrete conclusions. With the rapid development of novel technologies, it is crucial to find consensus about what digital well-being at work entails. The aim of this review is to unravel the existing academic landscape on employees' digital well-being and work towards an overarching framework of employee digital well-being.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.44](https://doi.org/10.18690/um.fov.4.2025.44)

ISBN

978-961-286-998-4

Keywords:

digitalization,
well-being,
employee,
review,
novel technologies



University of Maribor Press

1 Introduction

Technological advancements continuously impact how, when, and where employees work, presenting both challenges and opportunities for organizations and their employees. Many organizations have made a large scale shift towards an increasingly digitalized workplace, largely fueled by the COVID-19 pandemic. Moving forward, organizations are expected to undergo profound transformations due to technological advancements, with AI playing a crucial role in reshaping jobs and required skills (World Economic Forum, 2025).

As digitalization processes increase, so do the academic endeavors to unravel the impact of digital technologies on employee well-being. Despite – or perhaps due to – the rapid acceleration of academic endeavors, findings on the impact of digitalization on employees' well-being remain fragmented (cf. Cramarenco et al., 2023; Lunde et al., 2022). This fragmentation not only stems from the diverse range of digitalization processes studied, such as mobile connectivity and AI adoption (Büchler et al., 2020; Kong et al., 2021), but also from the variety of well-being outcomes under investigation, including life satisfaction, stress, and burnout (e.g., Lunde et al., 2022; Song & Gao, 2020; Tan et al., 2024).

Several meta-analyses have attempted to untangle the complex relationship between digitalization and employee well-being (e.g., Charalampous et al., 2019; Cramarenco et al., 2023; Johnson et al., 2020). Such overview studies often focus on specific forms of digital technologies, such as telework (Hill et al., 2024), or the implementation of AI (Cramarenco et al., 2023). Even though these overviews offer valuable insights into how particular forms of digital technologies influence employee well-being, they do not provide a comprehensive framework for synthesizing the broader concept of employees' digital well-being. With the rapid pace of emerging novel technologies, we face a research field in which it becomes important to find consensus about what digital well-being at work entails.

Employee well-being is a highly complex, multi-dimensional construct (Khalid & Syed, 2024). Collapsing diverse findings from a single well-being indicator into an overarching concept of well-being oversimplifies the complexity of the construct. For example, the spatial distance created between workers in remote work has been found to increase positive affect (for an overview see: Hill et al., 2024). Yet, based

on such a single indicator of well-being it cannot be concluded that remote work helps employee well-being. In fact, remote work is also associated with negative outcomes for well-being, such as burnout, emotional exhaustion, and loneliness (Hill et al., 2024). Given that digitalization impacts well-being in different ways, there is a need to focus on individual well-being indicators to accurately tap into the complexity of well-being.

This study aims to map the academic landscape of employees' digital well-being and develop an overarching framework. We conduct a qualitative systematic literature review (e.g., Paré et al., 2015; Schryen et al., 2021) focused on how digital technologies affect employee well-being. Beyond synthesizing existing research, the review integrates domain-specific insights to extend current theories and deepen our understanding of digital well-being (Schryen et al., 2021).

2 Theoretical Framework

2.1 Defining (employee) well-being

Employee well-being has been studied for decades, with anxiety, positive and negative affect, and stress as important outcome variables (Khalid & Syed, 2024). The variety of different indicators of well-being, indicates that well-being, as an overarching concept, oversimplifies the complex structure of the different dimensions of well-being that are impacted by digital technologies. To be able to scrutinize how digitalization at work affects employee well-being, it is required to differentiate indicators of well-being in more detail.

When a person experiences well-being, it is often described as 'feeling well' or 'doing well' (Martela & Sheldon, 2019). Feeling well refers to the hedonic part of well-being, whereas doing well reflects eudaimonic well-being. Hedonic well-being, often used interchangeably with subjective well-being, is reflected by subjective experiences of pleasure and includes aspects of life satisfaction, happiness, and positive affect (Kjell & Diener, 2021; Ryff et al., 2021). Eudaimonic well-being reflects the degree to which one is fully living up to one's potential (Ryan & Deci, 2001). Eudaimonic indicators of well-being include vitality, meaningfulness, and purpose in life (Huta & Waterman, 2014; Ryan & Deci, 2001). In the work context, an example of a eudaimonic indicator is work engagement (Hill et al., 2024).

Well-being is a subcategory of mental health. An individual's mental health constitutes of an individual's well-being and ill-being, which is known as the two-continua model of mental health (Greenspoon & Saklofske, 2001; Keyes, 2002). A person experiences ill-being in instances of disturbance or personal distress, and is sometimes also referred to as psychopathology (Lahey et al., 2017). Whereas ill-being and well-being have been studied on one continuum, more academic attention has recently been drawn to separating well-being from ill-being. Having low levels of stress (i.e., low indicators of ill-being), does not necessarily mean that someone is happy (i.e., high levels of well-being).

Events, daily experiences, or moment-to-moment occurrences may induce ups and downs in an individual's well-being and ill-being. When mental health indicators are assessed at a given time, within a specific period, or under specific circumstances they provide insights into the fleeting, so-called state components of mental health (Bakker, 2015; Huta, 2017). These ups and downs fluctuate around one's general level of mental health, which is an average degree of mental health that remains relatively stable over time. This is referred to as trait level well-being and ill-being (cf. Huta, 2017). Over time, the trait level component of mental health can gradually evolve (cf. Granic, 2005; Smith & Thelen, 2003).

The current study aims to delineate the domain specific aspect of digital well-being in the workplace. The study utilizes the current trends in distinguishing ill-being from well-being and examines both the more fleeting and more stable aspects of ill- and well-being. Ultimately, this would lead to a general overarching conceptualization of digital well-being and ill-being at work, which ought to be seen as a subcategory of the digital well-being framework at large.

2.2 Digital well-being (at work)

The rapid advancements in technologies and their impact on society have led to the emergence of the concept of digital well-being. Digital well-being, often used to describe how digital technologies shape overall well-being (Büchi, 2024), has become a societal buzzword and an emerging focus of academic research. The concept of digital well-being is generally built around daily uses of digital media such as the 24/7 mobile connectivity, screentime, and social media (Büchi, 2024; Twenge et al., 2018; Vanden Abeele, 2021). Whereas, a vibrant scholarly community is working towards

a framework for digital well-being in daily life (e.g., Büchi, 2024; Burr et al., 2020; Vanden Abeele, 2021), no such framework exists for digital well-being at work.

To a certain extent, similarities between technology use in daily life as described in the digital well-being framework and the work context exist. For example, the 24/7 mobile connectivity is very prevalent in daily life, but is also a significant concern in the work context for disturbing the work-life balance (Johnson et al., 2020). However, the introduction of innovative technologies at work also brings along unique job-related challenges and opportunities. That is, worries about how advancements in new technologies might render one's job obsolete are work-specific concerns, and can have a far-reaching impact on well-being (Ali et al., 2024). Hence, the (unique) concerns and opportunities that arise from digitalization at work, which can be seen as a subpart of digital well-being, needs to be studied on their own.

Understanding employee well-being often draws on job strain models, such as the Job Demands-Resources model, which unravels the impact on employee well-being as a result of job demand characteristics (Bakker, 2015; Bakker & Demerouti, 2018). While the mechanisms identified in digital well-being and job strain frameworks offer valuable insights into the mechanisms influencing (employee) well-being—both positively and negatively—these frameworks fall short in accounting for the evolving dynamics associated with the implementation and sustained use of digital technologies. Specifically, they tend to overlook the temporal dimension as a core aspect of digital well-being, thereby failing to capture the varying impacts that emerge across different stages of technology adoption.

To develop a comprehensive framework for employee digital well-being, we examine two key questions: (1) What are the associations between technology use and employee well-being and ill-being? and (2) What are the mechanisms that explain the relationship between technology use and well-being and ill-being outcomes? In identifying these mechanisms, we place special emphasis on the evolving nature of technology use and its implications for employee well-being, including the different stages of technology adoption and the temporal stability—or instability—of its impact over time.

3 Method

This systematic review will be conducted in accordance with the Preferred Reporting Items for Systematic Reviews, and Meta-Analysis (PRISMA) guidelines (Page et al., 2021). To identify the relevant studies, the review includes an identification, screening, and eligibility procedure.

3.1 Identification Procedure

To identify all potentially relevant publications, an extensive academic database search was conducted in April 2025. The searched databases include (a) Web of Science and (b) PsychINFO. The title, abstract and keywords of publications published in peer-reviewed journals written in English are searched for (a) mentions of digital technologies (e.g., artificial intelligence), (b) mentions of well-being indicators (e.g., well-being), and (c) words related to the work environment (e.g., employee). See Appendix I for a full overview of the advanced search strategy. Duplicate records were removed for the next phase. During the final phase, a backward search will be performed to scan for additional relevant publications.

3.2 Screening Procedure

The remaining abstracts from the identification procedure are further screened for relevance. This screening process is partially automated via ASreview. ASreview uses machine learning techniques to assist researchers in their effort to determine relevant publications (Van De Schoot et al., 2021). We adhere to the SAFE procedure when using ASreview (Boetje & Van De Schoot, 2024), including four phases. First, a random set of abstracts is screened and labeled by the researcher, used for the training data (1% of the publications from the identification procedure).

Second, the active learning process continuously improves predictions of the unlabeled publications. Based on the training set from phase one, patterns of relevant records are identified, and the researcher indicates whether or not these records are indeed relevant. The first default setup (TF_IDF with Naive Bayes) to identify patterns in the abstracts that the researcher codes as (ir)relevant has a fast and excellent overall performance (Van De Schoot et al., 2021). The active learning phase will be stopped when (1) key papers have been marked as relevant; (2) a

minimum of 10% of the total dataset has been screened; and (3) no relevant records have been identified in the last 50 records.

Third, to ensure that no relevant records are missed, more advanced and computationally complex models (such as doc2vec with a neural network) are trained to check if the fast model missed relevant records. In the last (fourth) phase it is essential to evaluate the accuracy to ensure the review is comprehensive. A simple learning model (TF_IDF with Naive Bayes) will be used to screen through records that have been identified as irrelevant during the previous phases. To train this final model, the 10 highest ranked and lowest ranked records from phase 3 are used. The researcher will stop screening when no extra relevant records have been identified in the last 50 records. After identifying all relevant records, a subset of the articles including both relevant and irrelevant records are evaluated by an independent coder, after which the interrater reliability will be established.

3.2 Eligibility procedure

The final step is to assess full-text analysis to determine if the record should be included in the review. Records should report on the link between a form of digitalization and indicators of employee well-being. Digitalization variables should be the independent variable, and employee well-being should be the dependent variable. In cross-sectional designs, variables can be analyzed in different directions, as there is no clear temporal ordering. For cross-sectional designs we will solely rely on correlational analyses to examine the relationships between variables.

References

- Ali, T., Hussain, I., Hassan, S., & Anwer, S. (2024). Examine how the rise of AI and automation affects job security, stress levels, and mental health in the workplace. *Bulletin of Business and Economics (BBE)*, 13(2), 1180–1186. <https://doi.org/10.61506/01.00506>
- Bakker, A. B. (2015). Towards a multilevel approach of employee well-being. *European Journal of Work and Organizational Psychology*, 24(6), 839–843. <https://doi.org/10.1080/1359432X.2015.1071423>
- Bakker, A. B., & Demerouti, E. (2018). Multiple levels in job demands–resources theory: Implications for employee well-being and performance. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being*. DEF Publishers. nobascholar.com
- Boetjé, J., & Van De Schoot, R. (2024). The SAFE procedure: A practical stopping heuristic for active learning-based screening in systematic reviews and meta-analyses. *Systematic Reviews*, 13(1), 81. <https://doi.org/10.1186/s13643-024-02502-7>

- Büchi, M. (2024). Digital well-being theory and research. *New Media & Society*, 26(1), 172–189. <https://doi.org/10.1177/14614448211056851>
- Büchler, N., Ter Hoeven, C. L., & Van Zoonen, W. (2020). Understanding constant connectivity to work: How and for whom is constant connectivity related to employee well-being? *Information and Organization*, 30(3), 100302. <https://doi.org/10.1016/j.infoandorg.2020.100302>
- Burr, C., Taddeo, M., & Floridi, L. (2020). The ethics of digital well-being: A thematic review. *Science and Engineering Ethics*, 26(4), 2313–2343. <https://doi.org/10.1007/s11948-020-00175-8>
- Charalampous, M., Grant, C. A., Tramontano, C., & Michailidis, E. (2019). Systematically reviewing remote e-workers' well-being at work: A multidimensional approach. *European Journal of Work and Organizational Psychology*, 28(1), 51–73. <https://doi.org/10.1080/1359432X.2018.1541886>
- Cramarencu, R. E., Burcă-Voicu, M. I., & Dabija, D. C. (2023). The impact of artificial intelligence (AI) on employees' skills and well-being in global labor markets: A systematic review. *Oeconomia Copernicana*, 14(3), 731–767. <https://doi.org/10.24136/oc.2023.022>
- Granic, I. (2005). Timing is everything: Developmental psychopathology from a dynamic systems perspective. *Developmental Review*, 25(3–4), 386–407. <https://doi.org/10.1016/j.dr.2005.10.005>
- Greenspoon, P. J., & Saklofske, D. H. (2001). Toward an integration of subjective well-being and psychopathology. *Social Indicators Research*, 54(1), 81–108. <https://doi.org/10.1023/A:1007219227883>
- Hill, N. S., Axtell, C., Raghuram, S., & Nurmi, N. (2024). Unpacking virtual work's dual effects on employee well-being: An integrative review and future research agenda. *Journal of Management*, 50(2), 752–792. <https://doi.org/10.1177/01492063221131535>
- Huta, V. (2017). An overview of hedonic and eudaimonic well-being concepts. In L. Reinecke & M. B. Oliver (Eds.), *The Routledge handbook of media use and well-being: International perspectives on theory and research on positive media effects* (pp. 14–33). Routledge/Taylor & Francis Group.
- Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies*, 15(6), 1425–1456. <https://doi.org/10.1007/s10902-013-9485-0>
- Johnson, A., Dey, S., Nguyen, H., Groth, M., Joyce, S., Tan, L., Glozier, N., & Harvey, S. B. (2020). A review and agenda for examining how technology-driven changes at work will impact workplace mental health and employee well-being. *Australian Journal of Management*, 45(3), 402–424. <https://doi.org/10.1177/0312896220922292>
- Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, 43(2), 207. <https://doi.org/10.2307/3090197>
- Khalid, A., & Syed, J. (2024). Mental health and well-being at work: A systematic review of literature and directions for future research. *Human Resource Management Review*, 34(1), 100998. <https://doi.org/10.1016/j.hrmr.2023.100998>
- Kjell, O. N. E., & Diener, E. (2021). Abbreviated three-item versions of the satisfaction with life scale and the harmony in life scale yield as strong psychometric properties as the original scales. *Journal of Personality Assessment*, 103(2), 183–194. <https://doi.org/10.1080/00223891.2020.1737093>
- Kong, H., Yuan, Y., Baruch, Y., Bu, N., Jiang, X., & Wang, K. (2021). Influences of artificial intelligence (AI) awareness on career competency and job burnout. *International Journal of Contemporary Hospitality Management*, 33(2), 717–734. <https://doi.org/10.1108/IJCHM-07-2020-0789>
- Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin*, 143(2), 142–186. <https://doi.org/10.1037/bul0000069>
- Lunde, L.-K., Fløvik, L., Christensen, J. O., Johannessen, H. A., Finne, L. B., Jørgensen, I. L., Mohr, B., & Vleeshouwers, J. (2022). The relationship between telework from home and employee health: A systematic review. *BMC Public Health*, 22(1), 47. <https://doi.org/10.1186/s12889-021-12481-2>

- Martela, F., & Sheldon, K. M. (2019). Clarifying the concept of well-being: Psychological need satisfaction as the common core connecting eudaimonic and subjective well-being. *Review of General Psychology*, 23(4), 458–474. <https://doi.org/10.1177/1089268019880886>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52(1), 141–166. <https://doi.org/10.1146/annurev.psych.52.1.141>
- Ryff, C. D., Boylan, J. M., & Kirsch, J. A. (2021). Eudaimonic and hedonic well-being: An integrative perspective with linkages to sociodemographic factors and health. In M. T. Lee, L. D. Kubzansky, & T. J. VanderWeele (Eds.), *Measuring Well-Being* (1st ed., pp. 92–135). Oxford University Press New York. <https://doi.org/10.1093/oso/9780197512531.003.0005>
- Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, 7(8), 343–348. [https://doi.org/10.1016/S1364-6613\(03\)00156-6](https://doi.org/10.1016/S1364-6613(03)00156-6)
- Song, Y., & Gao, J. (2020). Does rework stress employees out? A Study on working at home and subjective well-being for wage/salary workers. *Journal of Happiness Studies*, 21(7), 2649–2668. <https://doi.org/10.1007/s10902-019-00196-6>
- Tan, J., Lim-Soh, J., & Tan, P. L. (2024). The impact of teleworking on women’s work–life balance and life satisfaction: A longitudinal study from Singapore. *Applied Research in Quality of Life*, 19(5), 2595–2615. <https://doi.org/10.1007/s11482-024-10340-x>
- Twenge, J. M., Martin, G. N., & Campbell, W. K. (2018). Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology. *Emotion*, 18(6), 765–780. <https://doi.org/10.1037/emo0000403>
- Van De Schoot, R., De Bruin, J., Schram, R., Zahedi, P., De Boer, J., Weijdemans, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, 3(2), 125–133. <https://doi.org/10.1038/s42256-020-00287-7>
- Vanden Abeele, M. M. P. (2021). Digital wellbeing as a dynamic construct. *Communication Theory*, 31(4), 932–955. <https://doi.org/10.1093/ct/ctaa024>
- World Economic Forum. (2025). *Future of jobs report*. World Economic Forum. <https://www.weforum.org/publications/the-future-of-jobs-report-2025/>

Appendix I

Search string indicators

Well-Being indicators
well-being, wellbeing, mental health, job satisfaction, work engagement, happiness at work, flourishing, work-life balance, ill-being, psychopathology, burnout, emotional exhaustion, cynicism, depersonalization, negative professional self-evaluation, bore-out, job dissatisfaction, work stress, job stress, work strain, job strain, work disengagement, job anxiety, workaholism, techno-stress, technostress, techno-eustress
Digitalization indicators
digital transformation, digitali*ation, technolog*, industry 4.0, industry 5.0, artificial intelligence, machine learning, automation, algorithms, cloud computing, robot*, Internet of Things, smartphone*, metaverse, augmented reality, virtual reality, extended reality, big data, ICT, machine, wearables, e-communication*, video conferenc*, video call*, e-mail, email, social media, instant messag*
Workplace indicators
Employee*, Worker*, Labo*r, workplace NOT students NOT adolescents

RESEARCH IN PROGRESS

GREEN IS – EXTERNAL PRESSURES AND IS EXECUTIVES

TOBIAS HASSMANN

Munich Business School, Department of Management, Munich, Germany
tobias.hassmann@munich-business-school.de

Addressing the urgent need for sustainable solutions aligned with the United Nations' (UN) Sustainable Development Goals (SDGs) requires collective action across all societal levels, including the information systems (IS) discipline. One strand of IS research focuses on understanding how IS executives' actions and beliefs contribute to sustainability outcomes through Green IS practices. This research proposes to investigate how external pressures influence IS executives' perceptions of the salience of sustainability issues, and how these perceptions, in turn, shape their Green IS-related actions and underlying beliefs. Adopting a qualitative research design, the proposed research study spans two phases: first, expert interviews with IS executives experienced in Green IS will be conducted to explore their perspectives. Second, a case study will validate and deepen these findings by identifying the factors that influence executives' Green IS actions and beliefs, and examining how and why these factors exert their influence. The proposed research aims to generate both theoretical contributions to Green IS and practical insights for organizations seeking to enhance sustainability through IS leadership.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.45](https://doi.org/10.18690/um.fov.4.2025.45)

ISBN

978-961-286-998-4

Keywords:

sustainability,
green IS,

IS executives,

CIO,

CDO,

qualitative research



University of Maribor Press

1 Introduction

As the global community copes with pressing sustainability challenges, the urgency for innovative solutions has become pronounced (Dwivedi et al., 2022; Lennerfors et al., 2015; Pan et al., 2022). The United Nations' (UN) Sustainable Development Goals (SDGs) offer a detailed roadmap towards achieving sustainability on the economic, environmental, and societal level (Watson et al., 2021). This ambitious agenda requires transformative actions across all sectors (Butler & Hackney, 2021). The discipline of information systems (IS) is uniquely positioned to contribute significantly to this endeavor by developing and implementing IS-based solutions that foster sustainability (Dwivedi et al., 2022; Melville, 2010; Seidel et al., 2017).

Sustainability is a broad and terminologically fuzzy concept (Guandalini, 2022). Thus, we distinguish sustainability outcomes from practices geared towards achieving those (Chen & Roberts, 2024). Sustainability outcomes can be classified by intention, scope of impact – direct, or indirect –, and alignment with the UNs' SDGs (Hassmann & Westner, 2024b; Schoormann et al., 2025). Relevant sustainability practices in context of IS are Green IS, Green IT, and digital sustainability (DS). Green IS describes practices that are future-oriented, leveraging technology-driven scenarios and use cases to achieve sustainability outcomes through strategic use of IS (Chen & Roberts, 2024; Guandalini, 2022; Loeser et al., 2017). Green IT practices are restricted to the IT function, aiming to efficiently operate existing IT systems (Kotlarsky et al., 2023; Loeser et al., 2017; Pan et al., 2022). Lastly, DS practices combine the realm of digital transformation (DT) with sustainability, focusing on enhancing environmental, social, and economic well-being and ensuring that digital technologies drive sustainable innovation and minimize their own environmental impact (Guandalini, 2022; Kotlarsky et al., 2023; Schoormann et al., 2025).

Numerous IS studies have emphasized the need for a more profound understanding of sustainability within the IS discipline (Dwivedi et al., 2022; Gholami et al., 2016; Melville, 2010; Seidel et al., 2017; Watson et al., 2010). Nonetheless, considerable gaps persist in IS literature. One example is the dearth of research exploring the role of IS executives in championing sustainability outcomes within organizations (Hassmann & Westner, 2024b) – despite being deemed an important topic (Gholami et al., 2013; Loeser et al., 2017; Melville, 2010). Building on this, we propose to

explore how IS executives, specifically Chief Information Officers (CIOs) and Chief Digital Officers (CDOs), contribute to achieving sustainability outcomes through Green IS practices. CIOs and CDOs are the highest-ranking IS leaders within organizations, who influence corporate strategy (Karahanna & Preston, 2013; Karahanna & Watson, 2006) and drive topics of DT (Fitzgerald et al., 2014; Weill & Woerner, 2013). Thus, their role is pivotal for integrating Green IS practices into an organization's DT and digital strategy (Menz, 2012; Tumbas et al., 2018).

This paper is structured as follows: Chapter 2 introduces the foundational concepts underlying the proposed research project: institutional theory, the Input-Mediator-Outcome (IMO) framework, and the connection between external pressures and salient sustainability issues. Chapter 3 outlines the proposed research model and the research objectives (ROs). Chapter 4 details the suggested methodological approach.

2 Background and theoretical underpinning

My proposed research endeavor is theoretically underpinned by institutional theory (Butler & Hackney, 2021; Campbell, 2007; Chen & Roberts, 2024; DiMaggio & Powell, 1983; Scott, 2014) and the IMO framework (Ilgen et al., 2005; Klotz et al., 2014; Mathieu et al., 2008).

Existing IS literature shows a gap in explaining how *external pressures* impact IS executives' actions and beliefs towards Green IS practices (Hassmann & Westner, 2024b). This is a noteworthy finding: external pressures are closely linked to institutional theory (Butler & Hackney, 2021; Campbell, 2007; Chen & Roberts, 2024; Scott, 2014), which is a frequently employed theoretical model to explain organizational and individual Green IS practices (Butler & Hackney, 2021). Institutional theory explores how organizations are influenced by the norms, rules, values, and cultural expectations of the environments in which they operate (Scott, 2014). It considers institutions to be systems of established norms and rules that guide behavior. For organizations to survive and thrive, they must align with institutional expectations to gain legitimacy (Butler & Hackney, 2021; DiMaggio & Powell, 1983; Scott, 2014). Thus, they strive for isomorphism to obtain institutional legitimacy and consequently become similar over time (Liang et al., 2007; Scott, 2014). The trends towards institutional isomorphism are propelled by *coercive*, *mimetic*, and *normative* pressures that shape organizational behavior, structures, and practices,

often leading to conformity and homogeneity within industries or fields (Butler & Hackney, 2021; Campbell, 2007; Scott, 2014). *Coercive* pressures arise from formal and informal pressures exerted by other organizations or authorities, such as governments, regulators, and powerful stakeholders, which demand compliance with specific rules or norms (Butler & Hackney, 2021; DiMaggio & Powell, 1983). *Mimetic* pressures describe how, in situations of uncertainty, organizations often imitate successful or legitimate peers, adopting similar practices to reduce risk and uncertainty (DiMaggio & Powell, 1983; Scott, 2014). Finally, *normative* pressures stem from professional standards and norms established by networks of professionals, trade associations, or industries, emphasizing conformity to accepted practices and values (Butler & Hackney, 2021; Campbell, 2007; Chen & Roberts, 2024; Scott, 2014). Relatedly, external pressures also impact organizations' strategic actions: they impact stakeholders and consumers through salient issues (Chen & Roberts, 2024). These are topics of matter for stakeholders and consumers, which can become part of organizations' *instrumental* and *expressive* logics, influencing managers' perception and prioritization of issues (Chen & Roberts, 2024). *Instrumental* logic pertains to the extent to which an issue supports the achievement of the organization's strategic goals. *Expressive* logic refers to the alignment of a salient issue with an organization's core values and beliefs, reflecting how a particular stakeholder concern resonates with the organization's identity and mission (Chen & Roberts, 2024).

Finally, the IMO framework suggests that IS executives receive different input, for example, from the Top Management Team. Following institutional theory (Butler & Hackney, 2021; DiMaggio & Powell, 1983; Scott, 2014), I consider external pressures one important input factor for IS executives that influence their actions, which, in turn, are mediated by team processes (Ilgen et al., 2005; Mathieu et al., 2008). These are processes such as team formation, for example, trust-building or planning, team functioning, for instance, collaboration or collective learning, and emergent cognitive and affective states (Ilgen et al., 2005). The IMO framework operates through episodic cycles: each outcome feeds back into the system as new input, reinforcing continuous development (Ilgen et al., 2005; Klotz et al., 2014).

3 Research model and propositions

I posit that external pressures influence IS executives' Green IS actions and beliefs (figure 1). In fact, IS research has shown that *coercive*, *mimetic*, and *normative* pressures directly affect human behaviors and beliefs, including those of the top management.

(Butler & Hackney, 2021; Campbell, 2007; Li et al., 2023; Liang et al., 2007). RO1 systematically explores how IS executives perceive external pressures and examines both the nature and mechanisms through which these pressures impact IS executives in their Green IS beliefs and corresponding actions (C'). I also hypothesize that external pressures influence organizations' stakeholders and consumers so that salient sustainability issues are fostered to which organizations align their instrumental and expressive logics (Chen & Roberts, 2024). Accordingly, I seek to examine (RO2) which sustainability issues arise from external pressures (a_1) and how IS executives consider these to be salient within their organization (b_1). Relatedly, recent IS research on social corporate responsibility showed that organizational factors, for example, organization size (Liang et al., 2007) or ideology (Gupta et al., 2017), impact how salient issues are perceived by organizations. Thus, I posit that organizational contexts condition the impact of external pressures by amplifying or filtering the salience of sustainability issues, which I plan to explore as part of RO3. Finally, drawing on the IMO framework, I argue that IS executives' Green IS beliefs and actions generate sustainability outcomes through mediating team processes and emergent states (Ilgen et al., 2005; Klotz et al., 2014; Mathieu et al., 2008). I will exclusively examine team formation and functioning processes to analyze how these translate IS executives' actions into tangible sustainability outcomes (a_2), which then serve as new inputs for IS executives in an iterative feedback loop (b_2).

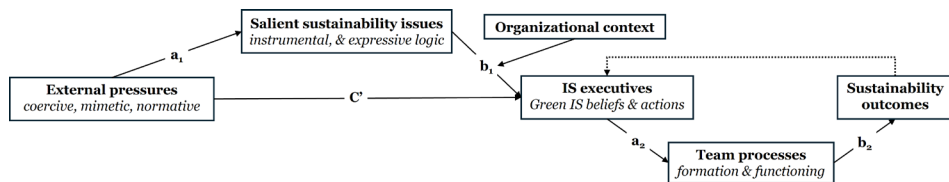


Figure 1: Research propositions and proposed theoretical model

4 Methodological approach

The overarching goal of this research is to generate evidence supporting the proposed theoretical model and research propositions, enabling theory abduction and offering practical guidance for IS executives to achieve sustainability outcomes through Green IS. I plan to use a mixed-method, qualitative research approach that follows an exploratory sequential research design, rooted in a pragmatist viewpoint (Creswell & Creswell, 2023). A pragmatist approach is appropriate since it allows to focus on a realworld challenge and actionable solutions (Johnson & Duberley, 2000; Simpson & Den Hond, 2022; Visser, 2019). Also, it supports abductive reasoning, which aligns with studying dynamically evolving and complex phenomena such as sustainability (Creswell & Creswell, 2023; Hassmann & Westner, 2024b).

Table 1 summarizes the proposed research approach, which hinges on two research phases: first, I plan to employ systematizing expert interviews, focusing on collecting factual-technical, process, and context expert knowledge (Bogner et al., 2009; Hassmann & Westner, 2024a). For interviewing, I plan to purposefully select experts who are IS executives with Green IS experience and working in different organizational settings (Hassmann & Westner, 2024a). I acknowledge the challenges in gaining access to IS executives for interviewing and will therefore apply different mitigation strategies such as lowering the participation barriers by conducting interviews online (Hassmann & Westner, 2024a; Robinson, 2021). The interview will probe how IS executives perceive and prioritize external pressures for their Green IS actions (RO1). The second aim is to identify external pressures that shape sustainability issues such that IS executives believe them to be salient (RO2). By interviewing IS executives from organizations that have different characteristics, I can trace patterns and contrasts in how these differences filter or amplify salient sustainability issues (RO3). Second, a subsequent case study will allow me to corroborate and triangulate the findings from the expert interviews, tracing the identified themes in realworld organizational context (Eisenhardt, 1989; Yin, 2018). By immersing in an organizational setting, I can observe how contextual factors influence the salience of sustainability issues (RO3) and how team collaboration processes translate IS executives' Green IS requirements into sustainability actions that generate outcomes, which become new inputs (RO4). To synthesize insights, I will apply Gioia's inductive coding approach (Gioia et al., 2013), identifying first-order and aggregated concepts related to external pressures and team processes.

Sustainability outcomes will be classified by scope of impact, intent (Schoormann et al., 2025), and alignment with UNs' SDGs (Hassmann & Westner, 2024b).

Table 1: Research objectives and planned research methodologies

# RO	Research objective	Methodology
RO1	Systematically explore <i>how</i> IS executives <i>perceive, respond to, and prioritize external pressures</i> in their Green IS actions and beliefs	Expert interview
RO2	Identify <i>salient sustainability issues</i> that arise from external pressures that shape IS executives' Green IS beliefs and actions	Expert interview, Case study
RO3	Explore <i>how organizational features</i> enable or constrain the perception of salient sustainability issues by Green IS executives	Expert interview Case study
RO4	Reveal <i>which and how team formation and functioning processes</i> translate IS executives' Green IS actions and beliefs into sustainability outcomes	Case study

5 Conclusion

My proposed research project pursues two overarching objectives. First, it seeks to advance theoretical understanding by refining existing models that explain the formation of IS executives' Green IS actions and beliefs – particularly in response to external pressures – and how beliefs and actions subsequently contribute to sustainability outcomes aligned with the global UNs' SDGs. Second, the proposed study aims to generate actionable recommendations for practitioners, providing guidance on effectively implementing and sustaining Green IS practices within their organizations to achieve measurable sustainability outcomes.

References

- Bogner, A., Littig, B., & Menz, W. (Eds.). (2009). Interviewing experts. Palgrave Macmillan UK.
<https://doi.org/10.1057/9780230244276>
- Butler, T., & Hackney, R. (2021). The role of informational mechanisms in the adoption of Green IS to achieve eco-sustainability in municipalities. *Information & Management*, 58(3), 103320.
<https://doi.org/10.1016/j.im.2020.103320>
- Campbell, J. L. (2007). Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. *Academy of Management Review*, 32(3), 946–967.
<https://doi.org/10.5465/amr.2007.25275684>
- Chen, A., & Roberts, N. (2024). Joint drivers of different shades of green IT/IS practices: A strategic cognition perspective. *Information Technology & People*, 37(3), 1103–1125.
<https://doi.org/10.1108/ITP-07-2022-0562>

- Creswell, J. W., & Creswell, J. D. (2023). *Research design* (Sixth edition). SAGE.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160. <https://doi.org/10.2307/2095101>
- Dwivedi, Y. K., Hughes, L., Kar, A. K., Baabdullah, A. M., Grover, P., Abbas, R., Andreini, D., Abumoghli, I., Barlette, Y., Bunker, D., Chandra Kruse, L., Constantiou, I., Davison, R. M., De', R., Dubey, R., Fenby-Taylor, H., Gupta, B., He, W., Kodama, M., ... Wade, M. (2022). Climate change and COP26: Are digital technologies and information management part of the problem or the solution? An editorial reflection and call to action. *International Journal of Information Management*, 63, 102456. <https://doi.org/10.1016/j.ijinfomgt.2021.102456>
- Eisenhardt, K. M. (1989). Building theories from case study research. *The Academy of Management Review*, 14(4), 532. <https://doi.org/10.2307/258557>
- Fitzgerald, M., Kruschwitz, N., Bonnet, D., & Welch, M. (2014). Embracing digital technology: A new strategic imperative. 55(2), 1–12.
- Gholami, R., Sulaiman, A. B., Ramayah, T., & Molla, A. (2013). Senior managers' perception on green information systems (IS) adoption and environmental performance: Results from a field survey. *Information & Management*, 50(7), 431–438. <https://doi.org/10.1016/j.im.2013.01.004>
- Gholami, R., Watson, R., Hasan, H., Molla, A., & Bjorn-Andersen, N. (2016). Information systems solutions for environmental sustainability: How can we do more? *Journal of the Association for Information Systems*, 17(8), 521–536. <https://doi.org/10.17705/1jais.00435>
- 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>
- Guandalini, I. (2022). Sustainability through digital transformation: A systematic literature review for research guidance. *Journal of Business Research*, 148, 456–471. <https://doi.org/10.1016/j.jbusres.2022.05.003>
- Gupta, A., Briscoe, F., & Hambrick, D. C. (2017). Red, blue, and purple firms: Organizational political ideology and corporate social responsibility. *Strategic Management Journal*, 38(5), 1018–1040. <https://doi.org/10.1002/smj.2550>
- Hassmann, T., & Westner, M. (2024a). Enhancing expert interviews: Insights from IS and digital transformation research. *Harnessing Opportunities: Reshaping ISD in the Post-COVID-19 and Generative AI Era (ISD2024 Proceedings)*. 32nd International Conference on Information Systems Development, Gdańsk, Poland. <https://doi.org/10.62036/ISD.2024.48>
- Hassmann, T., & Westner, M. (2024b). IS executives and sustainability: A literature review. 2024 26th International Conference on Business Informatics (CBI), 60–69. <https://doi.org/10.1109/CBI62504.2024.00017>
- Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, 56(1), 517–543. <https://doi.org/10.1146/annurev.psych.56.091103.070250>
- Johnson, P., & Duberley, J. (2000). *Understanding management research*. SAGE Publications Ltd. <https://doi.org/10.4135/9780857020185>
- Karahanna, E., & Preston, D. S. (2013). The effect of social capital of the relationship between the CIO and top management team on firm performance. *Journal of Management Information Systems*, 30(1), 15–56. <https://doi.org/10.2753/MIS0742-1222300101>
- Karahanna, E., & Watson, R. T. (2006). Information systems leadership. *IEEE Transactions on Engineering Management*, 53(2), 171–176. <https://doi.org/10.1109/TEM.2006.872247>
- Klotz, A. C., Hmieleski, K. M., Bradley, B. H., & Busenitz, L. W. (2014). New venture teams: A review of the literature and roadmap for future research. *Journal of Management*, 40(1), 226–255. <https://doi.org/10.1177/0149206313493325>
- 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. <https://doi.org/10.17705/1jais.00825>

- Lennerfors, T. T., Fors, P., & Van Rooijen, J. (2015). ICT and environmental sustainability in a changing society: The view of ecological world systems theory. *Information Technology & People*, 28(4), 758–774. <https://doi.org/10.1108/ITP-09-2014-0219>
- Li, X., Wang, N., Jiang, B., & Jia, T. (2023). Institutional pressures and proactive environmental strategy: The mediating effect of top managerial environment attitude and the moderating effect of new media pressure. *Business Strategy and the Environment*, 32(8), 6106–6123. <https://doi.org/10.1002/bse.3475>
- Liang, Saraf, Hu, & Xue. (2007). Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management. *MIS Quarterly*, 31(1), 59–87. <https://doi.org/10.2307/25148781>
- Loeser, F., Recker, J., Brocke, J. V., Molla, A., & Zarnekow, R. (2017). How IT executives create organizational benefits by translating environmental strategies into Green IS initiatives. *Information Systems Journal*, 27(4), 503–553. <https://doi.org/10.1111/isj.12136>
- Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, 34(3), 410–476. <https://doi.org/10.1177/0149206308316061>
- Melville, N. P. (2010). Information systems innovation for environmental sustainability. *MIS Quarterly*, 34(1), 1–21. <https://doi.org/10.2307/20721412>
- Menz, M. (2012). Functional top management team members: A review, synthesis, and research agenda. *Journal of Management*, 38(1), 45–80. <https://doi.org/10.1177/0149206311421830>
- Pan, S. L., Carter, L., Tim, Y., & Sandeep, M. S. (2022). Digital sustainability, climate change, and information systems solutions: Opportunities for future research. *International Journal of Information Management*, 63, 102444. <https://doi.org/10.1016/j.ijinfomgt.2021.102444>
- Robinson, G. (2021). Capturing a moving target: Interviewing fintech experts via LinkedIn. *Area*, 53(4), 671–678. <https://doi.org/10.1111/area.12726>
- Schoormann, T., Möller, F., Hoppe, C., & Vom Brocke, J. (2025). Digital Sustainability: Understanding and managing tensions. *Business & Information Systems Engineering*. <https://doi.org/10.1007/s12599-025-00937-3>
- Scott, W. R. (2014). *Institutions and organizations ideas, interests, and identities* (Fourth edition). SAGE.
- Seidel, S., Bharati, P., Fridgen, G., Watson, R. T., Albizri, A., Boudreau, M.-C. (Maric), Butler, T., Chandra Kruse, L., Guzman, I., Karsten, H., Lee, H., Melville, N., Rush, D., Toland, J., & Watts, S. (2017). The sustainability imperative in information systems research. *Communications of the Association for Information Systems*, 40, 40–52. <https://doi.org/10.17705/1CAIS.04003>
- Simpson, B., & Den Hond, F. (2022). The contemporary resonances of classical pragmatism for studying organization and organizing. *Organization Studies*, 43(1), 127–146. <https://doi.org/10.1177/0170840621991689>
- Tumbas, S., Berente, N., & Brocke, J. V. (2018). Digital innovation and institutional entrepreneurship: Chief digital officer perspectives of their emerging role. *Journal of Information Technology*, 33(3), 188–202. <https://doi.org/10.1057/s41265-018-0055-0>
- Visser, M. (2019). Pragmatism, critical theory and business ethics: Converging lines. *Journal of Business Ethics*, 156(1), 45–57. <https://doi.org/10.1007/s10551-017-3564-9>
- Watson, R. T., Boudreau, M.-C., & Chen, A. J. (2010). Information systems and environmentally sustainable development: Energy informatics and new directions for the IS community. *MIS Quarterly*, 34(1), 23–38. <https://doi.org/10.2307/20721413>
- Watson, R. T., Elliot, S., Corbett, J., Farkas, D., Feizabadi, A., Gupta, A., Iyer, L., Sen, S., Sharda, R., Shin, N., Thapa, D., & Webster, J. (2021). How the AIS can Improve its Contributions to the UN's Sustainability Development Goals: Towards A Framework for Scaling Collaborations and Evaluating Impact. *Communications of the Association for Information Systems*, 48(1), 476–502. <https://doi.org/10.17705/1CAIS.04841>
- Weill, P., & Woerner, S. L. (2013). The future of the CIO in a digital economy. 12(2), 65–75.

Yin, R. K. (2018). Case study research and applications: Design and methods (Sixth edition). SAGE.

RESEARCH IN PROGRESS

TOWARDS UNDERSTANDING COGNITIVE BIASES IN CYBERSECURITY GOVERNANCE

GULET BARRE, TIM HUYGH, DINH KHOI NGUYEN,
ARNO NUIJTEN

Open University of the Netherlands, Faculty of Science, Limburg, the Netherlands
gulet.barre@ou.nl, tim.huygh@ou.nl, khoi.nguyen@ou.nl, arno.nuijten@ou.nl

Cognitive biases can influence the decision-making of board members and CISOs responsible for managing cyber risks. However, limited attention has been given to understanding how these biases affect cybersecurity governance, specifically in the communication of risks between CISOs and boards. This paper aims to address this gap by identifying cognitive biases and proposing how these biases influence communication and strategic decision-making in cybersecurity governance. By further examining their impact, we strive to uncover the mechanisms that contribute to underestimations or distortions in risk perception, which can compromise an organization's ability to respond effectively to cyber threats. This short paper provides three exemplary biases expected to influence communication and decision-making in cybersecurity governance. Following the initial results, we propose a series of interviews with CISOs to reveal the challenges they face when communicating cyber risks to boards, focusing on how biases influence the decisions regarding cybersecurity risks.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.46](https://doi.org/10.18690/um.fov.4.2025.46)

ISBN

978-961-286-998-4

Keywords:

cybersecurity governance,
board decision-making,
cognitive biases,
risk communication,
cyber threats



University of Maribor Press

1 Introduction

Despite cybersecurity being recognized as a critical component of corporate governance and something that should be on the radar of the board of directors (De Haes et al., 2020), many boards remain ill-equipped to fulfil their strategic responsibilities in managing cyber risks (Valentine, 2016). To be able to take responsibility for cyber risk in the boardroom and ask the right questions to the CISOs within the organization, thereby holding them accountable, adequate governance measures should be in place. This pertains to proper information about cyber risks coming from the organization and board composition and expertise to be able to make an appropriate assessment of cyber risks (Smaili et al., 2022).

Moreover, decision-making in the boardroom is inherently complex. Directors rely on mental shortcuts, known as heuristics, to simplify information processing. Although heuristics can expedite decision-making, they can frequently lead to systematic errors called cognitive biases (Berthet, 2022). Such biases can significantly influence the decision-making process, particularly in the context of cybersecurity governance (Vedadi & Warkentin 2020). In other words, any decision-making is subject to biases in how we process information and estimate risks.

While we identify six biases through a systematic literature review, this short paper discusses three exemplary cognitive biases that affect board-level decision-making in the context of cybersecurity governance. We subsequently aim to address the following research question: What specific biases influence boards and CISOs communication and interactions regarding cybersecurity? Particularly, we validate and expand these biases by conducting interviews with CISOs. By exploring how cognitive biases manifest in decision-making in cybersecurity governance, we aim to pinpoint which biases are most relevant in practice and under which scenarios the biases emerge.

Our intermediate results indicate that there is a fragmented understanding of cognitive biases within the domain of Information Systems (IS), and no literature to date investigates relevant biases in the context of cybersecurity governance. Therefore, our study contributes in the following ways. We outline a literature review of cognitive biases within leading IS literature, and rationalize the relevance of these biases in the context of cybersecurity governance. As such, we demonstrate cognitive

biases as a meaningful theoretical lens to understand board-level communication and decision-making around cybersecurity. Toward this contribution, this study validates the results with biases identified in broader board-level decision-making literature. We will also empirically investigate these findings through interviews with CISOs.

2 Scope of Literature Search and Procedure

Our search methodology aligns with the evolving academic discourse surrounding biases and their implications for risk assessment and decision-making in cybersecurity. We initiated our investigation by exploring core terms that encapsulate the essence of our study: 'bias' and 'risk.' These terms were selected for their broad applicability and relevance across a spectrum of studies pertaining to decision-making and judgment in organizational contexts. To ensure a comprehensive and academically robust foundation, we utilized the Senior Scholars' List of Premier Journals as a basis, combined with snowballing.

To ensure a focused and efficient selection process, we established a set of criteria for identifying papers that would be relevant to our study. Specifically, we sought publications where the terms related to biases appeared prominently in the title, keywords, or abstract. This step was essential to exclude articles that only tangentially mentioned bias without exploring it as a primary topic of investigation. The application of these criteria enabled us to refine our search and identify papers that specifically addressed the types of biases pertinent to our research. We looked at papers published between January 1992 and September 2023. The initial screening phase resulted in 120 papers. We then applied a snowballing technique (both backward and forward), using the same inclusion/exclusion criteria to identify additional relevant papers. Although there was an overlap among papers due to similar research focus, this overlap helped confirm the significance of key studies and increased the reliability of the examined papers. The papers were selected based on their apparent alignment with our research focus, demonstrating a range of biases. Each of these papers was subsequently subjected to a manual review process, wherein we examined their content to verify their relevance and depth of analysis on bias-related issues.

During the manual review, papers that did not sufficiently address or identify bias as a primary subject were excluded from further consideration. This step was crucial in ensuring that only papers with substantial discussions on bias, whether through theoretical analysis, empirical investigations, or methodological studies, were included in the final list for deeper examination. The review process facilitated the exclusion of papers that, while potentially informative, did not contribute directly to our objective of understanding biases within decision-making frameworks in the IS context.

Ultimately, the combination of database search, keyword filtering, and manual review enabled us to curate a robust and relevant set of 52 papers. We identify six biases, and provide discussions on three exemplary biases below.

3 Exemplary biases

3.1 Optimism - Pessimism bias

Literature from the field of IS typically characterizes optimism bias as an individual's general tendency to underestimate the probability of unfavorable outcomes (Legoux et al., 2014). The opposite of optimism bias is pessimism bias, which occurs when a manager reports a project is in a worse state than it actually is (Snow et al., 2007).

The primary mechanism that influences optimism bias is overconfidence. This phenomenon manifests in two primary ways: overreaction and underreaction in the market (Daniel et al., 1998). Overreaction occurs when investors attribute excessive significance to their information, causing stock prices to rise or fall excessively. Conversely, underreaction happens when investors underestimate or ignore new public information, resulting in delayed price adjustments. Overconfidence amplifies both behaviors, causing prices to deviate from the actual fundamentals of the market. According to Daniel et al. (1998), the trigger that elicits this reaction is private information signals. Specifically, investors receive new private information, such as analyst reports, which serves as a trigger for investor overconfidence (Hilary & Menzly, 2006). Consequently, this leads to stock prices being displaced from their intrinsic value based on inaccurate assessments (Odean, 1998). In the context of cybersecurity, this can manifest when individuals receive new information, for example, a report indicating that their systems are due for new security audits. This

trigger may give decision-makers a misleading sense of security, making boards believe their defenses are stronger than they actually are. As a result, boards may overlook potential risks and believe that they are fully protected, which can leave their organization vulnerable to cybersecurity threats.

A second mechanism is risk perception, which is defined as the subjective evaluation of the probability and potential severity of a risk among individuals and groups. Risk perception in cybersecurity mostly stems from subjective evaluations of the likelihood of a cyberattack (Eling et al., 2021). An example is, after hearing of a significant ransomware attack in the industry, board members may overestimate the probability of a similar attack within their organization.

3.2 Herding bias

The IS literature defines herding as an individual's propensity to conform to the behavior of preceding peers. This mechanism typically emerges in environments characterized by uncertainty. When individuals lack confidence in their knowledge, they are more inclined to follow others (Baddeley, 2013). Observing the actions of peers, they assume that the majority possesses superior information (Baddeley, 2013). The rationale is that if a larger number of people believe something, it may be perceived as more accurate (Bikhchandani et al., 1992). This trigger creates a process known as informational cascades, which occur when an individual observes the actions of predecessors and adopts their decision without considering their own judgment (Wang & Greiner, 2010).

Herding is observable in managerial decision-making. According to Kaufman and Li (2003), IT managers are known to follow crowds when making decisions regarding IT investments. This tendency to herd shows that managers may prioritize conformity over independent risk assessments, as they believe others' decisions are based on relevant information (Zhou & Lai, 2009). Therefore, managers may focus on supporting their choices with perceived consent rather than making decisions based purely on their own risk preferences (Vedadi & Warkentin, 2020). This behavior is also evident at the board level, where boards may often make decisions about cyber risk in response to external pressures, rather than basing it on their own risk appetite (Benaroch & Chernobai, 2017). For example, boards may respond to high-profile incidents such as data breaches or ransomware attacks by investing in

new security tools, primarily to emulate peer organizations, rather than selecting measures aligned with their own specific cyber risk profile (Kwon & Johnson, 2014).

As individuals' experiences can offer significant insights, it is vital for IT managers to create an environment that values individual insights, thus providing room for personal experience in decision-making. Neglecting such an environment increases the risk of errors due to herding. In the cybersecurity context, past research shows that managers upscaled their organization's security by the information they received about the security behavior of others (Barlow et al., 2018; Vedadi & Warkentin, 2020).

4 Discussion and Prospective Research Pathway

This study ultimately seeks to advance our understanding of how cognitive biases can influence the decision-making process in the context of cybersecurity governance. Our preliminary findings highlight a fragmented understanding of cognitive biases within the IS domain, with no prior research specifically addressing cognitive biases in the context of cybersecurity governance. This study bridges this gap in the following ways. First, we present a comprehensive literature review of cognitive biases within leading IS literature and rationalize their relevance to cybersecurity governance. Second, we demonstrate cognitive biases as a meaningful theoretical lens to understand board-level communication and decision-making on cybersecurity issues.

To these ends, we aim to integrate research on cognitive biases and board-level decision-making in the context of cybersecurity. By focusing on the specific impact of these biases on critical cybersecurity areas, this study lays the foundation for developing a comprehensive bias-aware approach to improve decision-making in cybersecurity governance. The overall goal of this research is to provide a set of theory-driven guidelines for board members and CISOs to make better decisions by recognizing and mitigating these biases.

The preliminary results presented in this paper contribute to the understanding of how cognitive biases affect cybersecurity governance decision-making. These insights foresee an important first step toward the ambition of this research, which

is to improve the decision-making of board members and CISOs, by integrating awareness of psychological influences into the strategic management of cyber risks.

The next phase of this study involves the extension and empirical validation of our preliminary results. Particularly, we refine these results through cross-referencing biases identified in broader (non-IS) board-level decision-making literature and subsequently empirically validate them via semi-structured interviews with CISOs. Regarding the interviews, we aim at indirectly asking CISOs about challenges and pinpointing these challenges to different identified biases. This strategy ensures that CISOs do not simply deny certain biases, which could potentially hinder obtaining interesting insights (Merendino et al., 2018). This validation phase is essential to fulfilling our overall objective: to help boards and CISOs make better-informed decisions by considering psychological factors (biases) that may influence their judgment in the context of cybersecurity.

References

- Baddeley, M. (2013). Herding, social influence and expert opinion. *Journal of Economic Methodology*, 20(1), 35-44.
- Barlow, J. B., Warkentin, M., Ormond, D., & Dennis, A. (2018). Don't even think about it! The effects of antineutralization, informational, and normative communication on information security compliance. *Journal of the Association for Information Systems*, 19(8), 3.
- Benaroch, M., & Chernobai, A. (2017). Operational IT failures, IT value destruction, and board-level IT governance changes. *MIS Quarterly*, 41(3), 729-A6.
- Berthet, V. (2022). The impact of cognitive biases on professionals' decision-making: A review of four occupational areas. *Frontiers in Psychology*, 12, 802439.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5), 992-1026.
- Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *The Journal of Finance*, 53(6), 1839-1885.
- De Haes, S., Van Grembergen, W., Joshi, A., Huygh, T., De Haes, S., Van Grembergen, W., ... & Huygh, T. (2020). Enterprise Governance of IT, alignment, and value. *Enterprise Governance of Information Technology: Achieving Alignment and Value in Digital Organizations*, 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.
- French, A. M., Storey, V. C., & Wallace, L. (2023). The impact of cognitive biases on the believability of fake news. *European Journal of Information Systems*, 1-22.
- Hilary, G., & Menzly, L. (2006). Does past success lead analysts to become overconfident?. *Management Science*, 52(4), 489-500.
- Jiang, Y., Ho, Y. C., Yan, X., & Tan, Y. (2022). What's in a "username"? The effect of perceived anonymity on herding in crowdfunding. *Information Systems Research*, 33(1), 1-17.
- Kaufmann, C., & Kock, A. (2023). The performance effects of optimistic and pessimistic project status reporting behavior. *International Journal of Project Management*, 41(7), 102514.

- Kwon, J., & Johnson, M. E. (2014). Proactive versus reactive security investments in the healthcare sector. *MIS Quarterly*, 38(2), 451-A3.
- 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), 1.
- Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L., & Canhoto, A. (2018). Big data, big decisions: The impact of big data on board level decision-making. *Journal of Business Research*, 93, 67-78.
- Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *The Journal of Finance*, 53(6), 1887-1934.
- Smaili, N., Radu, C., & Khalili, A. (2023). Board effectiveness and cybersecurity disclosure. *Journal of Management and Governance*, 27(4), 1049-1071.
- 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.
- Valentine, E. L. (2016). *Enterprise technology governance: New information and technology core competencies for boards of directors* (Doctoral dissertation, Queensland University of Technology).
- Vedadi, A., & Warkentin, M. (2020). Can secure behaviors be contagious? A two-stage investigation of the influence of herd behavior on security decisions. *Journal of the Association for Information Systems*, 21(2), 3.
- Vlek, C. A. (1996). A multi-level, multi-stage and multi-attribute perspective on risk assessment, decision-making and risk control. *Risk Decision and Policy*, 1(1), 9-31.
- Walden, E. A., & Browne, G. J. (2009). Sequential adoption theory: a theory for understanding herding behavior in early adoption of novel technologies. *Journal of the Association for Information Systems*, 10(1), 1.
- Wang, H., & Greiner, M. (2010). Herding behavior in the stock market: Evidence from stock trading volume data of Chinese stock market. *Journal of Empirical Finance*, 17(3), 483-494.
- Zhou, R. T., & Lai, R. N. (2009). Herding and information based trading. *Journal of Empirical Finance*, 16(3), 388-393.
- Zou, H., Sun, H., & Fang, Y. (2023). Satisfaction to stay, regret to switch: understanding post-adoption regret in choosing competing technologies when herding. *Information Systems Research*, 34(4), 1455-1475.

RESEARCH IN PROGRESS

EXPLORING THE FEASIBILITY OF GENERATIVE AI IN ENHANCING THE IDENTIFICATION OF SPATIAL AND REGULATORY OPPORTUNITIES USING URBAN DIGITAL TWINS

ROB PETERS,¹ FABIAN KOK,² JOHN VAN MEERTEN,³
KOEN SMIT³

¹ Province of Utrecht, Utrecht, the Netherlands
rob.peters@provincie-utrecht.nl

² HU University of Applied Sciences Utrecht, Data Science Department, Utrecht,
the Netherlands
fabian.kok@hu.nl

³ HU University of Applied Sciences Utrecht, Digital Ethics, Utrecht, the Netherlands
john.vanmeerten@hu.nl, koen.smit@hu.nl

This study explores the feasibility of using generative AI to enhance spatial and legislative opportunity finding within urban digital twins. Digital Twins (DTs) integrate real-time and historical data to provide comprehensive views of built environments, potentially aiding different disciplines involved in spatial planning practices. However, the complexity of legal frameworks and their visualization in DTs remains a challenge. Leveraging advancements in large language models (LLMs), this research investigates how multimodal AI can interpret complex legislative data to improve spatial planning. The study employs a Design Science Research (DSR) methodology, focusing on tuning existing LLMs with spatial content. Key findings include the successful generation of Geography Markup Language (GML) code, enhancing interoperability with spatial planning tools, and the iterative design process that improved the model's performance. Preliminary results indicate that a multimodal approach, including text, images, and GML code, significantly enhances the model's capability. Future research will focus on improving data quality, expanding multimodal capabilities, and evaluating real-world applications. This study contributes to the development of transparent, contestable, and explainable AI solutions for spatial planning.

Keywords:
generative AI,
LLMs,
spatial planning,
feasibility,
digital twins



University of Maribor Press

DOI
[https://doi.org/
10.18690/um.fov.4.2025.47](https://doi.org/10.18690/um.fov.4.2025.47)

ISBN
978-961-286-998-4

1 Introduction

The development of Digital Twins (DTs) for spatial planning offers potential advantages for policymakers, council members, GIS specialists, and other related disciplines. They enable the integration of real-time data, historical data, and predictive analytics to create a comprehensive view of built environments. Despite the fact that there is no consensus on a standardized or formal definition of a DT, most research views a DT as a cyber-physical system that shares the concepts of a physical entity, a virtual model and connections between both (Liu et al., 2022). In the context of this study, we adhere to the following definition of a DT (Boschert & Rosen, 2016): ‘The DT is not one complete model of the physical product, but a set of linked operation data artefacts and (simulation) models, which are of suitable granularity for their intended purpose and stakeholders and evolve throughout the product life-cycle.’ Yet, this definition does not fully comply with a true DT in the sense that it should support bi-directional communication (Fuller et al., 2020; Liu et al., 2022). Given this definition, there are still areas for improvement, one of which is the complexity of the underlying legal fundament and its visualization in a DT. Recent advancements in AI, especially the accessibility and potential multimodality of large language models (LLMs), present an opportunity to enhance spatial planning by interpreting complex legislative frameworks through various modalities. Multimodal AI models can process and integrate data from different sources, such as text, images, and spatial data, to provide more accurate and context-aware insights (Huang & Zhang, 2024; Ji & Gao, 2023; Sangeetha et al., 2024).

The research problem addressed in this paper is the necessity of explorative opportunity finding, particularly legislative opportunity finding, in urban spatial planning. This involves navigating numerous conflicting policy ambitions within limited space, such as parking, housing, mobility, recreation, and adhering to legislative norms like sound levels, heat stress, water safety, and emissions (Binnenlandsbestuur, 2011; Boer et al., 2007). The primary users and testers of this research are urban and provincial policy planners. They aim to identify specific areas for planning complex structures, such as windmills and solar panel fields, for which they currently use traditional GIS overlay methods to explore and pinpoint opportunity zones. To study the potential of LLM-technology in this context, we address the following research question in this paper: To what extent can a large

language model, tuned with spatial content, assist actors in answering the "where can I do what?" question in urban planning?

The where-can-I-do-what capability was expressed as an important need in 2004 to decrease the administrative burden for SME's seeking permission for business activities, which was later described again by the Ministry of Internal Affairs (Binnenlands Bestuur, 2011). It has a long history of rule-based capability research (Peters & van Engers, 2004) and we investigate the additional power of generative AI to enhance this capability. Given the aim to develop solutions for governmental legislative processes, transparency, contestability, and explainability are crucial (Dittmar, 2024; Liao & Vaughan, 2024). This paper outlines the initial configuration of our solution, preliminary findings, and the iterative design challenges encountered.

2 Background and related work

Large language models have the potential to aid spatial planning combined into Digital Twins. Recent research suggests that fine-tuned LLMs can support urban design by integrating city-specific data. For example, specialized models like CityGPT use the "CityInstruction" dataset, which incorporates geographical information, allowing them to tackle urban planning tasks effectively. These models can even perform at a level comparable to proprietary solutions on city-focused benchmarks (Bentley et al., 2024; Choi & Yoon, 2025; Feng et al., 2024). Another example is UrbanLLM, a customized model designed for city management. It breaks down complex planning questions into smaller tasks and coordinates specialized tools to handle them. This approach allows UrbanLLM to perform significantly better than general LLMs, like the GPT series and LLaMA, when tackling complex urban planning challenges (Jiang et al., 2024).

Multi-agent systems are also becoming more popular. In one study, researchers combined LLM 'agents' with a city's digital platform and knowledge base to answer questions about urban policies and services. The system efficiently directs queries to GIS data, legal documents, and AI models, significantly reducing planning time from days to just a few hours while maintaining high accuracy (Kalyuzhnaya et al., 2025). Furthermore, a study by Ji and Gao (2023) examined how LLMs, such as GPT-2 and BERT, process spatial data. By encoding geometries in the Well-Known Text

(WKT) format, these models were tested on identifying geometry types and spatial relations. Results showed they maintained geometry types and captured spatial relations (73% accuracy) but struggled with numerical estimates and object retrieval.

Taking these recent developments into account, we argue that further research into the application of domain-specific LLMs in geospatial tasks is needed, as the current knowledge base is rather immature and few studies focus on applying the fundamental knowledge on LLMs and DT's for spatial planning.

3 Research method

This study employs Design Science Research (DSR) methodology, involving multiple iterations of the model. DSR is a research paradigm that focuses on the creation and evaluation of artifacts designed to solve identified problems. This study, in its current phase, can best be positioned in the relevance cycle (Hevner et al., 2004, figure 2). The design choices in this study are guided by the FAIR principles (Findable, Accessible, Interoperable, and Reusable) and CODIO (Code Goed Openbaar Bestuur) implementation guidelines (Government of the Netherlands, 2022; Meijer, 2023). For tuning of the LLM model we used the content of all Dutch spatial plans 2012-2022 of all cities at www.ruimelijkeplannen.nl. To test the feasibility of the research approach and the capabilities of the LLM, we investigated three opportunity exploration maps (wind energy, solar power field and forest planting), see e.g., (Provincie Utrecht, 2024), together with their policy makers and technical designers to set the capability benchmark. This focus group of policy practitioners were familiar with the legalistic issues, the policy ambitions, the area constraints and the (end-)user needs such as from city planners. By directly including the GIS designers in the focus group who had built the traditional application we had access to the work process and content variables.

Key design decisions included:

- Using a large dataset of spatial content: The assumption was that a larger dataset would improve the model's ability to generalize and provide accurate insights. However, the downside was that the content was outdated, which posed challenges for data cleaning and preprocessing.

- Employing open-source LLM technology: For reasons of transparency, only open-source tools were used. This choice limited the processing power available, as commercial models and tools generally offer better performance.
- Focusing on tuning rather than training: Due to resource constraints, the study focused on tuning existing language models rather than training new ones from scratch. This approach was expected to limit the feasibility study's scope but was necessary given the available resources.
- Triangulation approach: To improve tuning efficiency, a triangulation approach was used, involving spatial, textual, and pictorial data. The assumption was that consistent data organization would enhance the tuning process. However, the quality of the pre-training data was less than expected, which posed additional challenges.

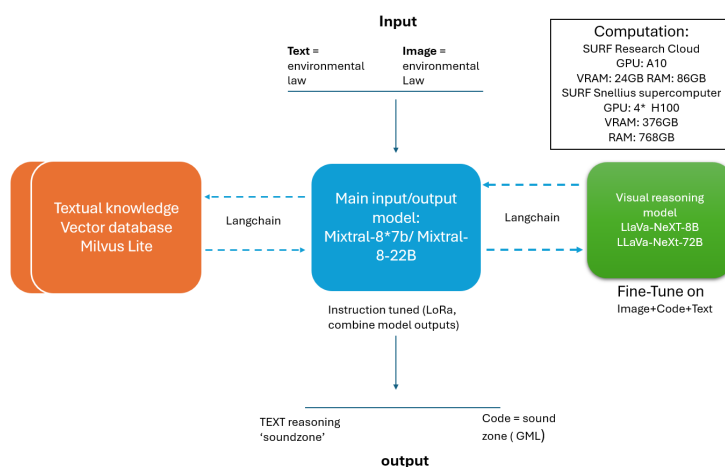


Figure 1: Architectural overview of the AI4Spatial solution

The feasibility study involved several steps, including:

- Testing picture recognition: Initial tests involved using LLaVa for picture recognition, such as identifying landmarks like the Utrecht Dom tower.
- Increasing content filtering efforts: Efforts were made to filter and clean the legacy GML-coded (Geography Markup Language) spatial plans to improve the quality of the input data.

- Optimizing configuration for multimodal cross-referencing: The configuration was optimized to handle cross-referencing between different types of data, such as text, images, and GML code.
- Upgrading configuration performance: The performance of the configuration was upgraded using the Dutch SURF Snellius supercomputer platform.
- Shifting from textual retrieval to GML code generation: The focus shifted from generating legal text to generating GML code, which is more useful for spatial planning tools and visualizations.
- Mixtral Model upgrade to Pixtral: The model was upgraded to Pixtral, eliminating the need for LLaVa and improving the handling of multimodal content.
- Testing with specific norms: Tests were conducted with specific norms, such as sound levels, to evaluate the model's ability to generate meaningful GML code.
- Prompt engineering: Prompt engineering techniques were used to filter the output and improve the relevance of the generated content. Examples were the inclusion of names of locations such as Utrecht or objects such as the Dom Tower in Utrecht in the prompt.

3 Preliminary results

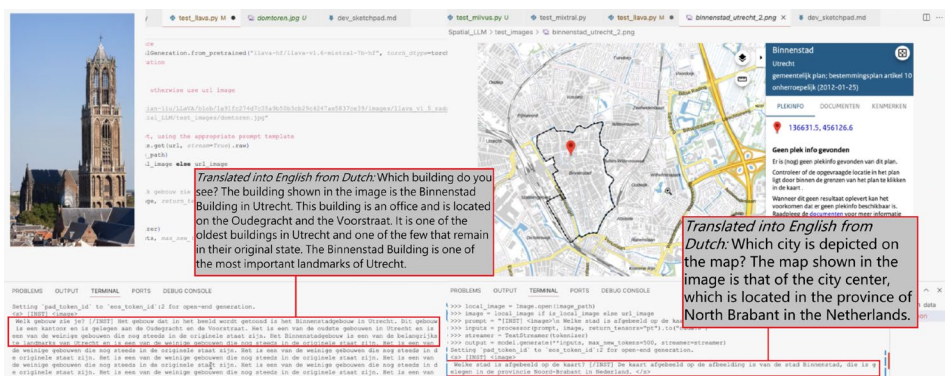


Figure 2: Erroneous results (hallucination) of the first iteration, where the computer does not recognize the right province (Brabant instead of Utrecht) and the model labels the object as 'binnenstad' (inner city), rather than the Dom Tower

The initial setup faced challenges, particularly with text-based retrieval depending heavily on direct content feeds. The model generated answers to prompts, but the quality of the output was limited by the quality of the input data. Subsequent iterations focused on generating meaningful GML code rather than generating legal text, enhancing interoperability with existing spatial planning tools.

Key Findings

- The first iteration generated answers to prompts but required significant content filtering. The reliance on direct pre-filtered content feeds in the prompts limits the model's ability to generalize and provide accurate insights.
- The second iteration demonstrated the feasibility of generating GML code, improving integration with spatial planning tools. The model was able to generate GML code that represented meaningful contours for specific norms, such as sound levels in housing and business areas, but it was yet unable to put those polygons into the context of a spatial plan.
- The third iteration (ongoing) involves semi-automated content cleaning and a method of image-binding, enhancing the model's capability to handle multimodal content. The use of image binding allowed for better cross-referencing between different types of data (code, image and juridical text), improving the overall performance of the model.

Technical Findings

The second iteration of the feasibility test provided proof of the concept of GML code generation. This outcome would enable better feeds of spatial planning content and better integration with spatial planning tools such as Tygron or ESRI and environmental permit processing tools. Figure 3 illustrates the GML coding capability as output of the large language model configuration.

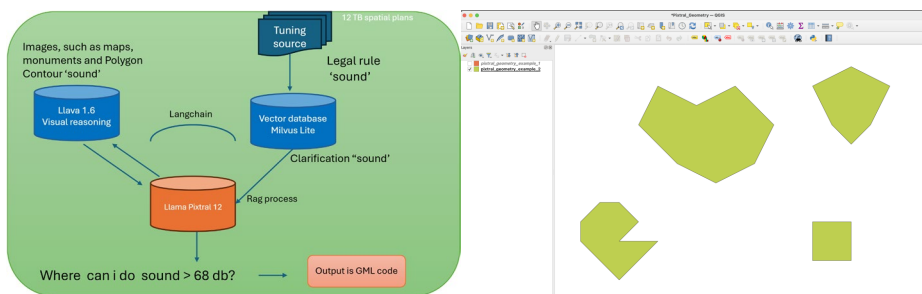


Figure 3: Example of the creation of the 'sound' polygon in GML code

During the testing of the second configuration, it was discovered that the Pixtral-based setup was not capable of handling tri-modal embedding of content triplets, such as satellite image/legal text/GML code or polygon/legal text/GML code, without significant additional scripting. Therefore, two additional techniques were applied: semi-automated content cleaning (iteration two) and image binding (iteration three).

4 Conclusions and Future Research

The feasibility study is ongoing, with promising results so far. However, several design challenges were encountered, such as the initial focus on text-based standards and the need for a multimodal approach. Future research will continue to refine the model and the use case process description such as 'finding windmill space' or 'solar panel farming area' or 'power net congestion optimization' with domain experts. We will also test with access to substantial geodata of the province of Utrecht (Open Geo Data, 2024) that was used to create the environmental legislative framework for this region (Omgevingsvisie Provincie Utrecht, 2024).

Key Conclusions

- Initial Setup Challenges: The initial setup failed to deliver the expected results due to the focus on text-based standards rather than geo-based standards. The shift to GML code generation improved the model's performance.
- Multimodal Approach: The second setup required a multimodal approach, which was not fully supported by the initial models. The upgrade to Pixtral

and the use of image binding improved the model's ability to handle multimodal content.

- **Iterative Design Process:** The iterative design process allowed for continuous improvement of the model, addressing the challenges encountered in each iteration. The focus on generating meaningful GML code rather than legal text enhanced the model's relevance for spatial planning applications.

Future Research Directions

- **Enhancing Data Quality:** Improving the quality of the input data through better content filtering and preprocessing techniques. This will involve collaboration with planning experts to ensure the relevance and accuracy of the data.
- **Expanding Multimodal Capabilities:** Further developing the model's ability to handle multimodal content, including the integration of additional data types such as satellite imagery and real-time sensor data.
- **Improving Explainability and Transparency:** Enhancing the explainability and transparency of the model's outputs to ensure trust and acceptance by policymakers and other stakeholders. This will involve developing methods for generating transparent justifications for the model's decisions.
- **Scaling Up the Model:** Addressing the computational constraints of open-source models by exploring ways to scale up the model's capabilities. This may involve leveraging cloud computing resources and optimizing the model's architecture for better performance.
- **Evaluating Real-World Applications:** Conducting pilot studies and real-world evaluations of the model's performance in various urban planning scenarios. This will provide valuable feedback for further refinement and validation of the model.

Acknowledgements

We would like to thank Luc de Horde and Bregtje Rijerman of the province of Utrecht for their help with this project. This project has been made possible by the Dutch Responsible Applied Artificial Intelligence programme.

References

- Bentley, P. J., Lim, S. L., Mathur, R., & Narang, S. (2024). Automated Real-World Sustainability Data Generation from Images of Buildings.
- Binnenlandsbestuur. (2011). Geozet viewer zet overheidsbekendmakingen op kaart. <https://www.binnenlandsbestuur.nl/digitaal/overheidsbekendmakingen-op-interactieve-kaart-door-geozet-viewer>
- Boer, A., van Engers, T., Peters, R., & Winkels, R. (2007). Separating law from Geography in GIS-based eGovernment services. *Artificial Intelligence and Law*, 15(1), 49–76. <https://doi.org/10.1007/s10506-007-9042-4>
- Boschert, S., & Rosen, R. (2016). Digital Twin—The Simulation Aspect. In *Mechatronic Futures* (pp. 59–74). Springer International Publishing. https://doi.org/10.1007/978-3-319-32156-1_5
- Choi, S., & Yoon, S. (2025). AI Agent-Based Intelligent Urban Digital Twin (I-UDT): Concept, Methodology, and Case Studies. *Smart Cities*, 8(1), 28. <https://doi.org/10.3390/smartsities8010028>
- Dittmar, L. (2024). The Explainability Challenge of Generative AI and LLMs - OCEG. <https://www.oceg.org/the-explainability-challenge-of-generative-ai-and-llms/>
- Feng, J., Du, Y., Liu, T., Guo, S., Lin, Y., & Li, Y. (2024). CityGPT: Empowering Urban Spatial Cognition of Large Language Models.
- Fuller, A., Fan, Z., Day, C., & Barlow C. (2020). Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access*, 8, 952–971.
- Kalyuzhnaya, A., Mityagin, S., Lutsenko, E., Getmanov, A., Aksenkin, Y., Fatkhiev, K., Fedorin, K., Nikitin, N. O., Chichkova, N., Vorona, V., & Boukhanovsky, A. (2025). LLM Agents for Smart City Management: Enhancing Decision Support Through Multi-Agent AI Systems. *Smart Cities 2025*, Vol. 8, Page 19, 8(1), 19. <https://doi.org/10.3390/SMARTCITIES8010019>
- Government of the Netherlands. (2022). Impact Assessment Fundamental Rights and Algorithms . <https://www.government.nl/documents/reports/2022/03/31/impact-assessment-fundamental-rights-and-algorithms>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science In Information Systems Research. *Design Science in IS Research MIS Quarterly*, 28(1), 75.
- Huang, J., & Zhang, J. (2024). A Survey on Evaluation of Multimodal Large Language Models.
- Ji, Y., & Gao, S. (2023). Evaluating the Effectiveness of Large Language Models in Representing Textual Descriptions of Geometry and Spatial Relations.
- Jiang, Y., Chao, Q., Chen, Y., Li, X., Liu, S., & Cong, G. (2024). UrbanLLM: Autonomous Urban Activity Planning and Management with Large Language Models. Findings of the Association for Computational Linguistics: EMNLP 2024, 1810–1825. <https://doi.org/10.18653/V1/2024.FINDINGS-EMNLP.98>
- Liao, Q. V., & Vaughan, J. W. (2024). AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. *Harvard Data Science Review*, (Special Issue 5). <https://doi.org/10.1162/99608F92.8036D03B>
- Liu, Y. K., Ong, S. K., & Nee, A. Y. C. (2022). State-of-the-art survey on digital twin implementations. *Advances in Manufacturing*, 10(1), 1–23.
- Meijer, A. (2023). Code Digitale Overheid kan vertrouwen in goed bestuur waarborgen - Nieuws - Universiteit Utrecht. <https://www.uu.nl/nieuws/code-digitale-overheid-kan-vertrouwen-in-goed-bestuur-waarborgen>
- Omgevingsvisie provincie Utrecht. (2024). Plannenvision Provincie Utrecht. <https://ruimtelijkeplannen.provincie-utrecht.nl/NL.IMRO.9926.2020OWVISIE-VA01>
- Open Geo Data. (2024). Geo-Point Utrecht. <https://geo-point.provincie-utrecht.nl/pages/open-data>
- Sangeetha, S. K. B., Mathivanan, S. K., Rajadurai, H., Cho, J., & Easwaramoorthy, S. V. (2024). A multi-modal geospatial-temporal LSTM based deep learning framework for predictive modeling of urban mobility patterns. *Scientific Reports*, 14(1), 31579. <https://doi.org/10.1038/s41598-024-74237-3>

RESEARCH IN PROGRESS

USING DIGITAL TWIN TECHNOLOGY FOR DEVELOPING A HEALTHY AND SUSTAINABLE LIVING ENVIRONMENT: A DUTCH CASE STUDY

SERHAT ÇALISKAN, SAM LEEWIS, KOEN SMIT

HU University of Applied Sciences Utrecht, Digital Ethics, Utrecht, the Netherlands
serhat.caliskan@hu.nl, sam.leewis@hu.nl, koen.smit@hu.nl

As urban environments face increasing challenges related to sustainability in development and planning, Digital Twin (DT) technology has emerged as a potential solution for enriching real-time and historical data into spatial planning processes. This study examines the potential and application of a DT in facilitating sustainable urban development, using the Utrecht Science Park (USP) as a case study to illustrate these possibilities. Specifically, the Healthy Heidelberglaan project, a collaborative initiative with different partners within the USP, demonstrates how this technology could support sustainable urban transformation. This research employs a methodological approach that includes sensor-based data acquisition, integration with external datasets and the implementation of scalable processes for historical data analysis. As this research is ongoing, subsequent phases will focus on development, translation of geographic data, analysing potential barriers and identifying key challenges and opportunities associated with DT technology and its application.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.48](https://doi.org/10.18690/um.fov.4.2025.48)

ISBN

978-961-286-998-4

Keywords:

digital twin,
urban development,
internet of things,
sustainability,
GIS



University of Maribor Press

1 Introduction

The rapid expansion of cities and the increasing complexity of urban environments have heightened the urgency of addressing environmental challenges. Innovative solutions are required to monitor and analyze the impacts of urbanization, and to develop strategies that enhance environmental sustainability and resilience. Urban areas, being significant contributors to environmental degradation, have the potential to implement changes that positively impact the environment and public health (Vardoulakis et al., 2016). Integrating green infrastructure, improving energy efficiency, and promoting sustainable transportation are examples of urban changes that can mitigate environmental challenges (Mazzetto, 2024).

A Digital Twin (DT) is a virtual representation of a physical object or process capable of collecting information from the real environment to represent, validate, and simulate the physical twin's present and future behavior (Botín-Sanabria et al., 2022). In the context of spatial planning, DT technology has emerged as potential tool that facilitates data-driven decision-making and supports spatial planning (El-Agamy et al., 2024). By creating a virtual representation of physical spaces, DTs enable (continuous) synchronization for an accurate representation of the physical landscape, offering decision-makers dynamic and context-specific insights for informed decision-making (Mongo & Daidj, 2025).

DT technology is increasingly being adopted by local and regional governmental organizations in the Netherlands to support spatial planning decision-making and to balance contradicting policy ambitions (Nochta et al., 2021). These DT solutions are becoming more complex as more perspectives, based on sophisticated calculations, are added to one integrated view of the problem space (Ranatunga et al., 2024). For example, integrating DT technologies with Geographic Information Systems (GIS) marks a transformative shift in geospatial analysis, enabling real-time monitoring and simulation of urban environments (Botín-Sanabria et al., 2022). This combination enhances the capability to analyze and visualize spatial data, which is an important capability for urban planning and environmental monitoring (Azadi et al., 2025; Botín-Sanabria et al., 2022).

Despite the potential of DT technology, there are still too few studies that approach this topic with case studies and empirical research, combining both the technical perspective and the challenges related, in the context of urban digital twins in the Netherlands (Mongu & Daidj, 2025; Oakes et al., 2024). Furthermore, our case is unique in a sense that it involves a large Dutch Science Park that serves a complex combination of zoning types in one shared area, i.e., living, working, traveling (e.g. by several options for transport in the form of bus, tram, bicycle, walking), recreation and study. To address this gap, we present a technical and empirical case study focused on the Utrecht Science Park (USP) in the Netherlands. The research question guiding this study is: *How can Digital Twin (DT) technology be utilized to support healthy and sustainable urban development, specifically in the context of the Utrecht Science Park (USP), and what are the key challenges and opportunities associated with its implementation?*

This paper is structured as follows: The next section provides an overview of the background and related work. Following this, we present the proposed research method and preliminary results of the case study. Finally, we discuss future avenues for research.

2 Background and related work

We identified two scientific contributions that have explored the application of digital twins in the context of sustainability and healthy living in relation to a science park. No relevant studies were identified that focus specifically on a Dutch case study. Fernandez et al. (2023) present the Smart DCU Digital Twin initiative, part of a broader Smart Dublin project aimed at developing intelligent, responsive, and adaptive urban environments through real-time monitoring and multi-stakeholder collaboration. It primarily addresses DTs as innovative tools for efficient infrastructure management, decision-making, and environmental monitoring in urban environments. The research integrates advanced 3D modeling, drone surveys (DJI Mavic 2 Pro), and Bentley's ecosystem (Context Capture, OpenCities Planner, iTwin Platform, Unreal Engine). IoT sensors used include occupancy, temperature, noise, illumination, and radar-based footfall counting sensors (WIA, CIVIC, HiData, Bigbelly). Bentley's 4D Analytics platform integrates and visualizes data, enabling predictive analytics. Challenges addressed include data management, interoperability, scalability, and stakeholder collaboration. This DT initiative effectively captures and analyzes campus environmental conditions and

infrastructure usage, offering immersive and intuitive interfaces for stakeholders through advanced visualization technologies. The methodology demonstrates significant potential for applying DTs to multifunctional urban areas like science parks, effectively corresponding with your research question on DT technology utilization and its associated challenges and opportunities in sustainable urban development.

The paper by Lu et al. (2020) addresses the development of a digital twin (DT) at the building and city levels, specifically applied to the West Cambridge Campus. The primary research goal was to create a systematic and comprehensive DT framework supporting operation and maintenance (O&M), decision-making, and asset management processes. The developed architecture includes a multitier system comprising data acquisition, transmission, digital modeling, data/model integration, and a service layer. Technologies and data involved in the study included BIM (Building Information Modeling), Geographic Information Systems (GIS), IoT sensor networks, QR code-based asset management, and real-time sensor data. The IoT-enabled Wireless Sensor Networks (WSNs) were deployed to monitor environmental and equipment conditions, such as indoor temperature, humidity, HVAC system status, vibration, and other asset-specific data. The integration layer utilized MySQL databases to incorporate data from building management systems (BMS), asset management systems (AMS), and space management systems (SMS), along with UAV-based photogrammetry and laser-scanned point cloud data for spatial information. This comprehensive DT framework successfully supported O&M activities by providing real-time monitoring, predictive analytics, visualization capabilities, and enhanced human interaction, all critical for sustainable urban and campus development.

This paper frames Digital Twin technology not just as a technical solution but as a sociotechnical system as stated by Nochta et al., (2021). In this framework, DTs support decision making processes that are influenced by social values, policy structures and stakeholder interactions. This aligns the ambition to foster sustainable urban development by integrating technological infrastructure with the focus on liveability.

3 Research method & Preliminary Results

This research adopts a Design Science Research (DSR) (Hevner et al., 2004) approach, supported by a narrative literature review (Paré et al., 2015). The narrative literature review provides a broad summary and synthesis of existing research on a specific topic, focusing on key concepts, theories, and findings. It is flexible in structure, allowing for the identification of trends and gaps in the literature, and is particularly useful for exploring new topics or providing background context for research.

Following DSR, an artifact (DT-prototype) is developed, assessed, and refined to support development of the digital twin for USP. The prototype is designed to address a practical business need, i.e., assisting with data-driven decision-making that supports spatial planning by visualizing spatial data in one integrated 2D/3D environment. The development of the prototype is informed by applicable technologies, frameworks and methods identified in the literature presented in the previous sections. In this section, we further present our preliminary findings based on the first prototype version that makes up the first design cycle. The activities within this study are aligned with the stages of the DSR framework. First, in the identification stage, we identified the lack of data-driven decision-making for sustainable development of the USP. In the objectives stage, together with stakeholders, we defined the objective to focus on the development of a digital twin for the visualisation of spatial data from the USP. This was done based on two focusgroups and multiple qualitative interviews with different stakeholders focusing on requirements. Based on the output of the previous stage, we designed and demonstrated the DT prototype that visualizes static and dynamic environmental data collected from the USP. The DT prototype has then been evaluated with multiple stakeholders in the evaluation stage, leading to several improvements to be designed in the next iteration of the prototype. We disseminated the DT prototype and our findings both towards stakeholders using different mechanisms such as events and presentations as well as publishing a first paper (this work).

4 Data acquisition and sensor integration

The prototype leverages real-time environmental data collected via an IoT sensor network deployed across two main streets within the USP. These sensors measure a range of environmental parameters.

The collected data is transmitted to the SamenMeten API, managed by the RIVM (Dutch national institute for Public Health and the Environment). This API provides access to dynamic environmental data, collected in an hourly interval. An example snapshot of the dynamic data is presented in Table 1. Furthermore, we also collected static data for analysis and visualization in the digital twin. In total, we integrated seven GIS map layers, being sound, rainfall, tropical days, heatmap tropical days, shadowmap buildings and shadowmap bicycle and walking lanes. These were collected using public data sets with the help of the municipality and province of Utrecht.

Table 1: Example of Collected Data

ID	Name	NO2 ($\mu\text{g}/\text{m}^3$)	PM2.5 ($\mu\text{g}/\text{m}^3$)	Calibrated PM2.5 ($\mu\text{g}/\text{m}^3$)	Temperature ($^{\circ}\text{C}$)	Humidity (%)	Pressure (hPa)	Longitude	Latitude
1	SSK_USP02	2.45	2.5	2.28	18.77	56.47	1018	5.169	52.085
3	SSK_USP01	8.63	19.02	13.33	20.15	85.9	1019	5.164	52.084
4	SSK_USP03	17.8	20.3	25.74	-1.3	67.68	1017.63	5.172	52.085
5	SSK_USP04	15.57	17.93	13.93	6.33	73.05	1027.67	5.172	52.085
6	SSK_USP05	11.35	13.24	10.93	7.33	75.05	1028.67	5.172	52.085
7	SSK_USP06	9.35	11.24	9.93	8.33	76.05	1029.67	5.172	52.085

5 Data visualisation

To facilitate effective data interpretation, collected environmental data is visualized within a Digital Twin framework using ESRI ArcGIS software. The GIS-Based approach enables spatial analysis of the USP. This allows stakeholders to identify high risk zones and design targeted interventions. An example of two combined datasets is the general shadowmap and the shadowmap with bicycle and walking lanes in the area, presented in Figure 1. The grey areas represent shadow coverage from trees or buildings, while the cyan lanes represent walking and bicycle lanes. The digital twin quite effectively presents the spots that are not covered by shade. This information, for example, is relevant to combine with the heatindex-datasets, but this is omitted due to space constraints.

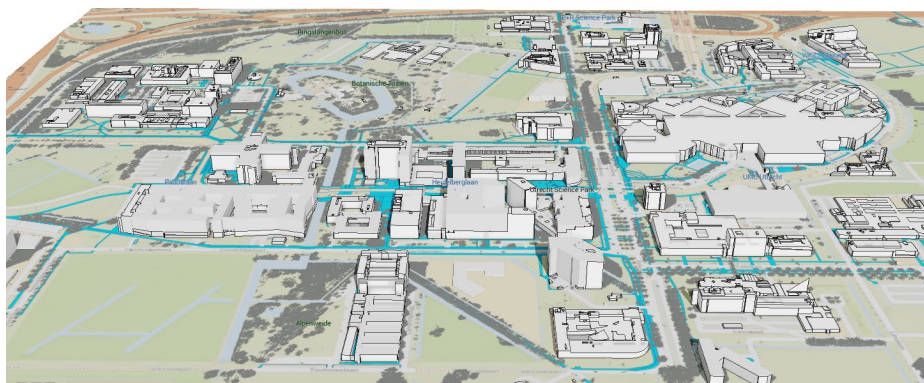


Figure 1: Example data visualization of shadow coverage and bicycle and walking lanes in ArcGIS (several areas are visible that show no shade at all for large areas)

To improve the development, we will first test it with all stakeholders. These include area development advisors, researchers, GIS experts, and software developers. These sessions will, for example, focus on decision-making utility and the relevance of the data.

6 Future work and research agenda

Future work will focus on several key areas to improve the artifact. A scalable database infrastructure still needs to be implemented to store and analyse historical environmental data from the IoT-network. Allowing different stakeholders to track environmental impact over time. See Figure 2 for the proposed database architecture for the next iteration of the artifact.

The data architecture is designed to facilitate integration with various external tools to improve the quality of the digital twin facilitated in the ESRI ArcGIS software. The system also needs to incorporate automated data validation processes to ensure safety, accuracy and consistency in historical records. Allowing different stakeholders to track environmental impact over time. Enhancing the Microsoft Azure Database structure to efficiently manage and store large volumes of historical data beyond the initial 200 entries currently supported by the SamenMeten-API is also to be added to the prototype in a future iteration.

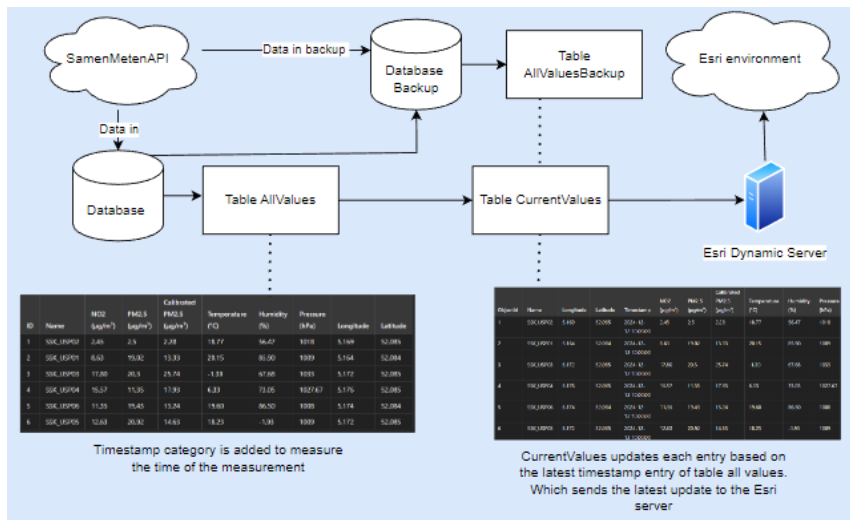


Figure 2: Dataflow process of Historical data

The prototype can be further enhanced by the development of additional GIS layers to visualize complex datasets, such as underground infrastructure and living experience surveys, enhancing stakeholder understanding of subsurface networks and user experiences. Continuous expansion of real-time data visualization capabilities, integrating various external and internal datasets, thus providing comprehensive insights into the multiple factors influencing Utrecht Science Park. Research and implementation of Geospatial data within the Digital Twin, translating complex sensor and GIS data into visualizations.

Extensive deployment of sensors across the entire Utrecht Science Park presents logistical and practical challenges that cannot be addressed within a short timeframe. Moreover, as areas within the park undergo redevelopment, existing sensors may need relocation to ensure continuous and accurate data collection. Therefore, future research should prioritize developing robust methodologies for extrapolating sensor data beyond their immediate measurement range. To support these extrapolations, nearby sensors could serve as validation points to ensure accuracy and reliability. However, this approach has limitations, particularly concerning parameters influenced by highly localized microclimatic conditions, such as soil moisture. Also, the current density of sensors is very low and we will focus on expanding the sensor network to improve validity of the data in a next iteration.

Furthermore, a future iteration should include an accessible public web portal that transparently displays non-sensitive information including real-time sensor data visualizations, historical trends, and interactive maps. Also, the prototype should include a secure, researcher-exclusive portal with advanced analytical tools, data export capabilities, and customizable visualization options for in-depth scientific exploration and analysis.

Lastly, it is important to gather insights from all stakeholders regarding the challenges experienced. Further empirical research is necessary because the prototype is not merely a technical implementation; it is a socio-technical solution and must therefore be investigated as such. To further develop the prototype, the values of stakeholders should be further identified and tested using future iterations. This is fundamental for the successful sustainable usage of digital twins as a technology for spatial planning in the context of the USP.

References

- Azadi, S., Kasraian, D., Nourian, P., & Wesemael, P. V. (2025). What Have Urban Digital Twins Contributed to Urban Planning and Decision Making? From a Systematic Literature Review Toward a Socio-Technical Research and Development Agenda. *Smart Cities*, 8(1), 32. <https://doi.org/10.3390/smartcities8010032>
- Botín-Sanabria, D. M., Mihaita, A.-S., Peimbert-García, R. E., Ramírez-Moreno, M. A., Ramírez-Mendoza, R. A., & Lozoya-Santos, J. D. J. (2022). Digital Twin Technology Challenges and Applications: A Comprehensive Review. *Remote Sensing*, 14(6), 1335. <https://doi.org/10.3390/rs14061335>
- El-Agamy, R.F., Sayed, H.A., Al Akhatatneh, A.M. et al. Comprehensive analysis of digital twins in smart cities: a 4200-paper bibliometric study. *Artif Intell Rev* 57, 154 (2024). <https://doi.org/10.1007/s10462-024-10781-8>
- Fernandez, J. B., Dhingra, M., Mandal, S., O'Connor, N. E., Kerr, A., McGuinness, K., Intizar, A., Cudden, J., & Mahon, K. (2023). Smart DCU Digital Twin: Towards Smarter Universities. *2023 IEEE Smart World Congress (SWC)*, 1–6. <https://doi.org/10.1109/SWC57546.2023.10448967>
- Hevner, March, Park, & Ram. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75. <https://doi.org/10.2307/25148625>
- Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., Konstantinou, E., Heaton, J., & Schooling, J. (2020). Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus. *Journal of Management in Engineering*, 36(3), 05020004. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000763](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000763)
- Mazzetto, S. (2024). A Review of Urban Digital Twins Integration, Challenges, and Future Directions in Smart City Development. *Sustainability*, 16(19), 8337. <https://doi.org/10.3390/su16198337>
- Mongo, M., & Daidj, N. (2025). Digital twin: A driver of sustainable smart cities? Evidence from a bibliometric analysis. *Environmental Economics and Policy Studies*. <https://doi.org/10.1007/s10018-025-00437-4>

- Nochta, T., Wan, L., Schooling, J. M., & Parlikad, A. K. (2021). A Socio-Technical Perspective on Urban Analytics: The Case of City-Scale Digital Twins. *Journal of Urban Technology*, 28(1–2), 263–287. <https://doi.org/10.1080/10630732.2020.1798177>
- Oakes, B. J., Zhang, H., Hatledal, L. I., Feng, H., Frasheri, M., Sandberg, M., Gil, S., & Gomes, C. (2024). Case Studies in Digital Twins. In J. Fitzgerald, C. Gomes, & P. G. Larsen (Eds.), *The Engineering of Digital Twins* (pp. 257–310). Springer International Publishing. https://doi.org/10.1007/978-3-031-66719-0_12
- 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>
- Ranatunga, S., Ødegård, R. S., Onstein, E., & Jetlund, K. (2024). Digital Twins for Geospatial Decision-making. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4-2024, 271–278. <https://doi.org/10.5194/isprs-annals-X-4-2024-271-2024>
- Vardoulakis, S., Dear, K., & Wilkinson, P. (2016). Challenges and Opportunities for Urban Environmental Health and Sustainability: The HEALTHY-POLIS initiative. *Environmental Health*, 15(S1), S30. <https://doi.org/10.1186/s12940-016-0096-1>

RESEARCH IN PROGRESS

BRIDGING THE DIGITAL DIVIDE FOR OLDER ADULTS: AN INVESTIGATION OF BARRIERS AND FACILITATORS

MILENA HEAD, KHALED HASSANEIN

McMaster University, DeGroote School of Business, Ontario, Canada
headm@mcmaster.ca, hassank@mcmaster.ca

As Information and Communication Technologies (ICT) become increasingly pervasive, older adults often face significant barriers to digital inclusion. This research-in-progress paper investigates the multi-faceted barriers and facilitators affecting ICT use and embracement among older adults. Barriers include cognitive decline, physical limitations, lack of digital skills, and psychosocial factors such as low confidence and distrust. Facilitators encompass inclusive interface design and motivational support tailored to older adults' needs. The study employs a mixed-methods approach, combining behavioral surveys, interviews, and neurophysiological measures to capture comprehensive insights into older adults' ICT experiences. By involving older adults in the co-design of ICT interfaces and identifying effective motivational strategies, this research aims to develop empirically driven tools and guidelines to enhance digital inclusion and well-being for older adults, ultimately bridging the digital divide.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.49](https://doi.org/10.18690/um.fov.4.2025.49)

ISBN

978-961-286-998-4

Keywords:

older adults,
digital divide,
barriers,
facilitators,
interface design



University of Maribor Press

1 Introduction

As Information and Communication Technologies (ICT) become increasingly pervasive, those that are left behind in technology access and use are highly disadvantaged (Wei et al. 2011). The COVID-19 pandemic highlighted the vital importance of increasing digital inclusion as migration to digital life rapidly accelerated. Older adults have frequently been excluded and marginalized with respect to technology (Weil et al. 2021) and this reality became even more vivid due to the pandemic. This digital exclusion gap is deeply widened when we intersect lower socioeconomic factors such as race, gender and new immigrant status with the older adults' segment (Yoon et al. 2020).

2 Digital Divides for Older Adults

When one thinks of a digital divide, not having access to technologies is what typically comes to mind (Compaine 2001). However, more recently, there has been an appreciation that digital divides exist at different levels, which are successive and interrelated (Loh & Chib 2021; Wei et al. 2011). The first level, Digital Access Divide (*access*), refers to inequities in access of hardware, software and connectivity. The second level, Digital Capabilities Divide (*use*), refers to inequities in ability (i.e., skills and/or access to training) and attitude (e.g., self-efficacy, trust, motivation) to use ICT. The third level, Digital Outcome Divide (*embracement*), refers to inequities of benefits derived from ICT use. Equality in ICT access and capabilities may not necessarily result in equivalent outcomes, as certain groups of people derive greater benefits from ICT than others (Van Deursen & Helsper 2015). In this research the focus is on how ICT is embraced and integrated in one's life in order to achieve positive tangible outcomes that can result in offline benefits to social, emotional, physical and intellectual wellness (Wilson et al. 2023; Van Deursen & Helsper 2015). With increased ICT and Internet penetration within developed countries, the issue of the digital access divide has waned (Loh & Chib 2021). As such, this research focuses on the digital divides of ICT use and embracement among older adults.

2.1 Barriers for Older Adult ICT Use and Embracement

Older adults may encounter various barriers when it comes to using and embracing ICT. Cognitive barriers emerge due to natural cognitive decline associated with neurocognitive, structural, or functional changes in the brain as people age (Harada et al. 2013). For example, age-related decreases in fluid cognitive abilities such as reasoning, pattern recognition and problem solving as well as decline in memory and attention can impact effective learning and use of new technologies (de Bruin et al. 2012). Physical barriers such as diminished motor control, vision, hearing, hand-eye coordination, flexibility of joints, touch sensation, among others, can impact performance (errors and efficiency) of ICT use and, ultimately, older adults' willingness to continue using technologies for their potential benefits (Wildenbos et al. 2018). Knowledge-based barriers arise from a lack of digital skills and training, further impacting older adults' cognitive ability to navigate websites, use applications, and resolve common issues, often stemming from limited exposure to technology in earlier years (Wilson et al. 2023; Wilson et al. 2021). Older adults need more time to learn new ICT-related skills (Wildenbos et al. 2018), emphasizing the importance of training and support (Marston et al. 2019). Finally, psychosocial barriers also play a substantial role in ICT use and embracement among older adults. These barriers encompass issues like low confidence (Horst et al. 2021), distrust and anxiety of ICT (Gomez-Hernandez et al. 2022), lack of interest (Nymberg et al. 2019), insufficient support (Nymberg et al. 2019), and concerns about privacy and online scams (Gomez-Hernandez et al. 2022). Particularly for older adults, self-stereotyping may become a self-fulfilling prophecy that impacts ICT use and embracement (Wagner et al. 2010; Wilson et al. 2023).

Extant research has tended to examine the impact of these barriers on ICT use in isolation rather than holistically understanding how these multi-faceted barriers may interact to not only reduce overall use of ICT and exacerbate digital exclusion but limit the opportunity to use these technologies for positive tangible outcomes for older adults (Wilson et al. 2021; Lu et al. 2023). Such barriers become more pronounced when considering the intersections of age with factors such as race, ethnicity, socioeconomic status, sex, disability, etc. (Loh & Chib 2021).

2.2 Facilitators for Older Adult ICT Use and Embracement

2.2.1 Interface Design

The interface design of ICT can either create a barrier to use, which can be insurmountable especially for marginalized users, or facilitate effective use and embracement of technologies. In an effort to accommodate the broad diversity of users who may face multiple individual barriers (such as older adults), ICT designers have employed universal design (also described as accessible design) approaches (Nussbaumer 2011). However, this “one size fits all” approach focuses on designing with the average user in mind where needs and preferences of some individuals are prioritized over others and it ignores the fact that these needs may change over time (Patrick & Hollenbeck 2021). A more inclusive approach to ICT interface design should take into account personal identifiers such as age, race, economic status, gender, etc., where design is done for groups that are typically underrepresented or ignored (also known as equity-based design) (Patrick & Hollenbeck 2021). For the underrepresented group of older adults, equity-based interface designs have tended to focus on overcoming visual and cognitive load barriers through larger font sizes, high-contrast colours and simplified navigation, instructions, and other design elements (Wilson et al. 2023).

A drawback of the traditional ICT design approaches is that they do not capture or address the impact of digital designs beyond the immediate use of the design. For example, they do not consider impacts on users’ offline well-being (Sin et al. 2021). Additionally, they tend to follow a “top-down” approach where the design process is based on technologists’ or geriatricians’ preconceptions of the needs of older adults (White et al. 2022). To create ICT interfaces that are truly useful, reduce barriers and have positive online and offline impact, older adults must be meaningfully engaged in the design process. Not involving older adults in the design of ICT has been recognized as a significant barrier to their use and embracement of these technologies (Wilson et al. 2023). As such, older adults still largely remain a ‘relevant but absent’ group in the design/development of ICT (Xie et al. 2012).

2.2.2 Motivation

ICT interfaces that are tailored to the unique needs of older adults are essential to remove obstacles for use. However, technology design alone is not sufficient to bridge digital divides. Alongside effective interface design, motivational facilitators need to inspire and support older adults in their journey towards ICT embracement for positive online and offline outcomes (Tyler et al. 2020). Self-efficacy theory, with foundations in social cognitive theory (Bandura 1997), is a well-established motivational theory that has been posited to positively influence older adults' learning and adoption of technology (Tyler et al. 2020). According to this theory, one's self-belief in one's competence (for example with ICT) is shaped by: (1) mastery experiences (experienced previous success or failure); (2) vicarious experiences (observation of success); (3) verbal persuasion (feedback from influential people) and (4) social ties and support networks (social milieu). The appropriateness and effectiveness of motivational facilitators may differ by individuals (varied backgrounds and barriers) and by stages of ICT use and embracement (Loh & Chib 2021).

3 Proposed Methodology

An initial model for this proposed research is shown in Figure 1. A combination of methods will be employed – from behavioural surveys and interviews to neurophysiological measures, ensuring a complete grasp of older adults' ICT experiences while capturing conscious and subconscious interactions.

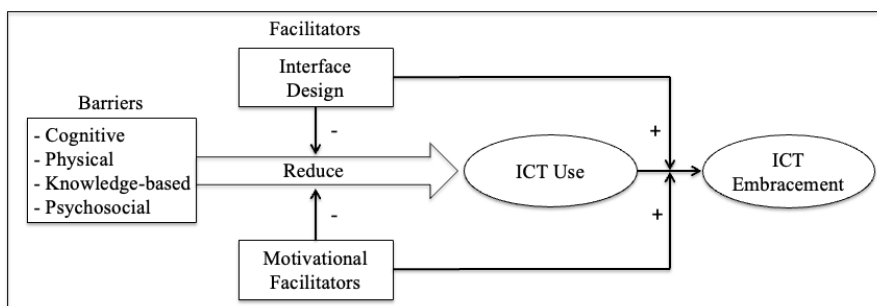


Figure 1: Barriers and Facilitators of ICT Use and Embracement

Source: Own

An exploratory study will be conducted to gain a holistic understanding of how different barriers may interact in their effects on ICT use and embracement by older adults. This investigation will also explore facilitators that motivate older adults to effectively engage with ICT. Interviews (n=15), focus groups (4 x 5 older adults each, n=20) and surveys (n=200) will be conducted across a diverse range of older adults, capturing the nuances of their ICT barriers and facilitators.

Based on the findings from the above exploratory study and in consultation with the older adult community, three interface designs will be developed to attenuate the negative impacts of multi-faceted barriers (examples may include attentional enhancement and social cue designs). Similarly, three most salient motivational facilitators will be identified, drawing from the above findings and self-efficacy theory (Bandura 1997). A controlled experiment will follow a three-by-three balanced analysis of variance design (3 interface designs x 3 motivational facilitators) with 270 older adults (30 replications per cell). To ensure generalizability, the sample will be drawn across a diverse range of older adults. Using the selected motivational facilitator, participants will be asked to interact with the selected interface design during several (approximately 4) one-hour sessions over the span of three to six months. At the end of each session, participants will be asked to complete a survey to express their views on the manipulated interface design.

While traditional self-reported human computer interaction methods may yield significant insights, such methods can suffer from shortcomings due to inherent biases (e.g.; retrospective, subjectivity, etc.). As such, each session will be conducted at the Digital Transformation Research Centre of a major Canadian university where neurophysiological equipment will allow for additional unobtrusive measurement of behavioural (e.g., eye movements), physiological (e.g., Electrocardiogram (EKG), Electromyography (EMG), Galvanic Skin Response (GSR)), and cognitive (e.g., Electroencephalography (EEG)) activities of diverse older adults engaged with ICT. Triangulating these methods with traditional self-reported methods will allow for a much deeper understanding of true older adult experiences by analyzing their underlying cognitive, biological, affective and attitudinal correlates.

3 Expected Contributions

The proposed research will provide: (1) An in-depth understanding of multi-faceted barriers and motivational facilitators for effective ICT use among older adults. Behavioural methods as well as neurophysiological measures will be utilized to provide rich insights into conscious and subconscious processes. (2) An examination of the efficacy of various interface designs and motivational facilitators via experimental comparisons across older adults with diverse subsets of multi-faceted barriers. (3) The development of empirically driven interface tools and guidelines to support diverse older adults across various ICT contexts. Overall, this research promises to help older adults remain connected, informed, and engaged in an increasingly digital world, ultimately promoting digital inclusion and well-being.

References

- Bandura, A., (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- 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.
- Compaine, B.M. (2001), *The Digital Divide: Facing a Crisis or Creating a Myth?* MIT Press, Cambridge, MA, pp. xvi-357.
- Gomez-Hernandez, M., Adrian, S. W., Ferre, X., & Villalba-Mora, E. (2022). Implicit, explicit, and structural barriers and facilitators for information and communication technology access in older adults. *Frontiers in Psychology*, 13, 874025.
- Harada, C. N., Love, M. C. N., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in geriatric medicine*, 29(4), 737-752.
- Horst, B. R., Sixsmith, A., Simeonov, D., and Mihailidis, A. (2021). Demographic and psychographic factors of social isolation during the COVID-19 pandemic: the importance of technology confidence. *Front. Public Health*, 9:749515.
- Loh, Y. A. C., & Chib, A. (2021). Reconsidering the digital divide: an analytical framework from access to appropriation. *Information Technology & People*, 35(2), 647-676.
- Lu, X., Jiang, J., Head, M., Kahai, S. S., & Yang, J. (2023). Synergistic Impacts of online and offline social participation on older adults' subjective well-being: evidence from the Canadian longitudinal study on ageing. *European Journal of Information Systems*, 1-18.
- Marston, H. R., Genoe, R., Freeman, S., Kulczycki, C., & Musselwhite, C. (2019). Older adults' perceptions of ICT: Main findings from the technology in later life (TILL) study. *In Healthcare* (Vol. 7, No. 3, p. 86).
- Nussbaumer, L. L. (2011). *Inclusive design: A universal need*. New York, NY: Fairchild Books.
- Nymberg, V. M., Bolmsjö, B. B., Wolff, M., Calling, S., Gerward, S., & Sandberg, M. (2019). 'Having to learn this so late in our lives...' Swedish elderly patients' beliefs, experiences, attitudes and expectations of e-health in primary health care. *Scandinavian journal of primary health care*, 37(1), 41-52.
- Patrick, V. M., & Hollenbeck, C. R. (2021). Designing for all: Consumer response to inclusive design. *Journal of consumer psychology*, 31(2), 360-381.

- 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.
- Tyler, M., De George-Walker, L., & Simic, V. (2020). Motivation matters: Older adults and information communication technologies. *Studies in the Education of Adults*, 52(2), 175-194.
- Van Deursen AJ and Helsper EJ (2015) The third-level digital divide: who benefits most from being online? In Robinson L, Cotten SR and Schulz J (eds), *Communication and Information Technologies Annual*. Bingley, UK: Emerald Group Publishing, 29–52.
- Wagner N., Hassanein K., Head M. (2010). Computer use by older adults: A multi- disciplinary review. *Computers in Human Behavior*, 26, 870-882.
- 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.
- Weil, J., Kamber, T., Glazebrook, A., Giorgi, M., & Ziegler, K. (2021). Digital inclusion of older adults during COVID-19: Lessons from a case study of older adults technology services (OATS). *Journal of Gerontological Social Work*, 64(6), 643-655.
- White, P. J. (2022). Designing products for older people's social and emotional needs: A case study. *Anthropology & Aging*, 43(2), 24-39.
- Wildenbos, G. A., Peute, L., & Jaspers, M. (2018). Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). *International journal of medical informatics*, 114, 66-75.
- Wilson, G., Gates, J. R., Vijaykumar, S., & Morgan, D. J. (2023). Understanding older adults' use of social technology and the factors influencing use. *Ageing & Society*, 43(1), 222-245.
- Wilson, J., Heinsch, M., Betts, D., Booth, D., & Kay-Lambkin, F. (2021). Barriers and facilitators to the use of e-health by older adults: a scoping review. *BMC public health*, 21, 1-12.
- Xie, B., Druin, A., Fails, J., Massey, S., Golub, E., Franckel, S., & Schneider, K. (2012). Connecting generations: developing co-design methods for older adults and children. *Behaviour & Information Technology*, 31(4), 413-423.
- Yoon H, Jang Y, Vaughan PW and Garcia M (2020) Older adults' Internet use for health information: digital divide by race/ethnicity and socioeconomic status. *Journal of Applied Gerontology* 39, 105–110.

DOCTORAL CONSORTIUM

ACCEPTANCE AND DEVELOPMENT OF BEYOND BANKING ECOSYSTEMS

SEBASTIAN VOIGT,¹ ALEXANDER HOLLAND²

¹UCAM Universidad Católica San Antonio de Murcia, Murcia, Spain
sgvoigt@alu.ucam.edu

²FOM University of Applied Sciences, Essen, Germany
alexander.holland@fom.de

This paper aims to examine the acceptance and development of Beyond Banking Ecosystems. It analyzes current market trends, the phenomena and classifications these ecosystems exhibit, and how their acceptance among bank customers can be evaluated. Current market studies are first reviewed to validate the research questions of relevance and acceptance of Beyond Banking services, followed by expert interviews from banking sectors to gain deeper insights into existing activities and assess potential business models. Additionally, a survey is conducted to gauge customer interest and acceptance. The results of this paper form the basis for the future design of a scalable ecosystem beyond banking. Its development will provide bank decision-makers with knowledge and a practical tool for strategically aligning beyond-banking initiatives.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.50](https://doi.org/10.18690/um.fov.4.2025.50)

ISBN

978-961-286-998-4

Keywords:

beyond banking,
open banking,
ecosystem,
acceptance,
development,
bank customer



University of Maribor Press

1 Introduction

Beyond banking, developing additional revenue streams for banks is highly topical. Since the launch of API interfaces under the Payment Services Directive 2 (PSD2) agreement in 2019 to introduce more extensive financial data disclosure, bank customers can easily choose third-party products via their online banking and various platforms (Deloitte, n.d.). This Open Banking strategy signifies a significant shift from a closed financial system to an open, customer-focused ecosystem influenced by regulation. The vision is for banks to become platform providers that cover more than just pure financial services.

1.1 Beyond Banking

Beyond Banking refers to banking transcending conventional financial practices, focusing on revenue generation beyond the standard banking model (Junghanns & Niebudek, 2019). Collaborating with third-party providers can facilitate the creation of novel business models that extend beyond traditional banking offerings. This concept includes various innovative financial products, services, and technologies designed to elevate customer experiences, enhance efficiency, and broaden access to financial services (Frankenberger et al., 2024).

1.2 Ecosystems and categorizations

Beyond Banking can unlock its full potential by leading and shaping the open banking era. Open banking (Cumps et al., 2020) reveals or makes accessible all key contact points in the banking industry: products, customers, and ecosystems (Figure 1).

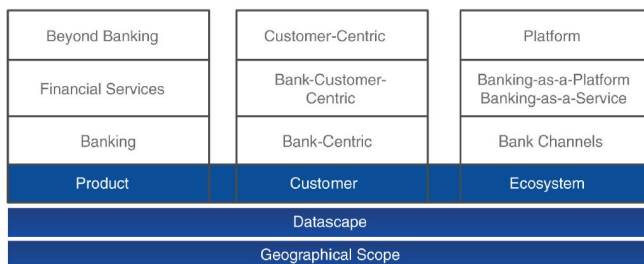


Figure 1: Framework for an Open Banking Strategy

Source: Cumps et al., 2020

Gómez-Márquez (2023) defines an ecosystem as a collection of organisms that live and interact within a specific environment. In the banking sector (Cumps et al., 2020), we can discuss a customer-centric strategy within the platform ecosystem where the customer engages with the bank and its partners. Additionally, this marks the expansion of the banking ecosystem and easier access to financial services, specifically the ability to reach beyond banking services through a platform that connects with third-party providers at this second development stage 2.0. Just as Li et al. (2016) predicted that the market would shift from a one-sided to a two-sided structure, the banking market is now experiencing a similar situation: a bilateral relationship between banks and third-party providers. This relationship should regularly exchange data streams and create value for customers. The recent emergence of Beyond Banking Ecosystem (BBE) 3.0, featuring an intra-institutional approach (Knörrer et al., 2021), has become particularly evident among cooperative banks. In this model, banks serve as market orchestrators, linking customers and facilitating the balance of supply and demand as intermediaries. This may increase the presence of various Third Party Providers (TPP) services. Based on historical observations and studies, BBEs can be categorized into three distinct stages:

Stage 1.0: Initial Beyond Banking relationships exist in isolation, lacking an automated ecosystem approach (Liang & Savage, 1990).

Stage 2.0: A B2C inter-institutional approach allows bank customers to receive offers from third-party providers (TPPs) (Li et al., 2016).

Stage 3.0: An intra-institutional approach (C2C) emerges, enabling bank customers to act as suppliers and consumers within the bank, with the bank operating as a marketplace (Knörrer et al., 2021).

2 Problem definition

Beyond banking offers excellent potential for banks to develop new sources of income. Recent studies (Chung et al., 2020) indicate that progress is underway in the initial two expansion stages. Nevertheless, limited research focused on the banking sector, and almost no exploration at the ecosystem level 3.0 incorporates an intra-institutional strategy. This strategy emphasizes a customer-to-customer (C2C) business model and the bank acting as a market orchestrator, leveraging machine

learning and data analytics for customer matching. Banks are aware of the usefulness of Beyond Banking, the potential added value for customers, and want to expand their business model via digital ecosystems (zeb-consulting, n.d.), but are initially sticking to purely financial products. This is because there is still a lack of clear business cases and the necessary skills to realize the potential of Beyond Banking.

This paper also investigates Beyond Banking's revenue potential, which developments and activities are taking place in expansion stage 2.0, what experience has already been gained with stage 3.0, and to what extent initial insights can be gained in direct comparison with both stages. Another important perspective is the acceptance on the demand side, i.e., the extent to which acceptance for Beyond Banking Services occurs on the market side.

3 Methodology

Once the fundamental definitions and relevance of Beyond Banking are clarified, alongside an analysis of the current market conditions and research validating its significance today, the concepts of Beyond Banking 2.0—focused on inter-institutional customer matching—and Beyond Banking 3.0—centered on intra-institutional customer matching—are examined. Standardized expert interviews (Hopf, 2004) utilizing a qualitative content analysis methodology are employed to gain deeper insights. The transcription and interview protocols will adhere to the guidelines established by Lamnek and Krell (2016). This study predominantly follows a deductive categorization approach, where pre-formulated questions and existing categories from the expert interviews align with the research questions and hypotheses derived from the literature review. This methodology allows for systematically assigning categories to specific text segments (Mayring, 2022).

Subsequently, further sub-questions are formulated to test specific hypotheses concerning the main questions. For the primary and secondary questions, main and subcategories were developed along the interview guide (Table 1).

Here, RQ1 addresses the revenue potential and developments of Beyond Banking (BB) 2.0 and the experiences of banking experts with BB 3.0. RQ2 deals with various business models for BB 2.0 and 3.0 and Beyond Banking Sales Signals as data-driven recommenders. At the same time, RQ3 focuses on frameworks for creating Beyond

Banking Ecosystems, key influencing factors and criteria, and challenges and risks for the ecosystem constructs.

Table 1: Coding System

Research Questions	Categories	Sub Code 1	Number of citations (frequency)	Frequency in % of 670 citations
RQ1			483	72,09%
	Revenue potential Beyond Banking		262	39,10%
		Customer Adoption Barriers	22	2,95%
		Ecosystem Integration Limitations	28	3,76%
		Provider Impatience	9	1,21%
		Indirect Revenue Potential	70	9,40%
		Data-driven Sales Opportunities	49	6,58%
		Diversified Beyond-Banking Services	72	9,66%
		Assessment	12	1,61%
	Developments Beyond Banking 2.0		130	19,40%
		Observations	13	1,74%
		Added values for customers and banks	70	9,40%
		Co-operations and offers	47	6,31%
	Experiences Beyond Banking 3.0		39	5,82%
		Yes/No	11	1,48%
		C2C-matching	12	1,61%
		B2C-matching	16	2,15%
	Acceptance and organisational effects		52	7,76%
RQ2			92	13,73%
	Business Models		42	6,27%
		Hypothesis evaluation	11	1,48%
		Use/Business Cases 2.0	18	2,42%
		Use/Business Cases 3.0	13	1,74%
	Possibilities of Dimensionalisation and Scaling		21	3,13%

Research Questions	Categories	Sub Code 1	Number of citations (frequency)	Frequency in % of 670 citations
		Decision-makers	8	
	Beyond Banking Sales Signals		29	4,33%
		Yes/No	13	1,74%
		If yes, utilising potential and determining	6	0,81%
		Establishment and use	10	1,34%
RQ3			34	5,07%
	Guideline/Concept/Framework for creation of BBE		15	2,24%
		Yes/No	12	1,61%
		concepts and frameworks	3	0,40%
	Key influencing factors and Criteria		19	2,84%
	Challenges and Risks		61	9,10%
		Challenges for BB 2.0 and 3.0	13	1,74%
		Legal risks	17	2,28%
		IT risks	18	2,42%
		Other risks	13	1,74%

Source: own illustration

After conducting the interviews, the responses from interviewees will be coded and categorized according to a structured coding agenda. Extensive research led to selecting MAXQDA software (Fries, 2016) for qualitative data analysis (QDA) and interview evaluation. The software's capability to integrate AI for automatic transcription and translation of interview texts was particularly persuasive, allowing the author to efficiently proceed with manual editing and coding (Cao et al., 2023).

Before data collection, it is crucial to identify suitable experts based on their skills and working environments relevant to the studied topic. Kuckartz (2018) emphasizes that the quality of information significantly hinges on the participant selection for expert interviews or evaluations. Furthermore, Creswell (2009) argues that qualitative approaches are warranted for concepts or phenomena that lack thorough research. Therefore, Mayring's deductive categorization is complemented by an iterative process that enables ongoing formative assessments of category

reliability and the created coding agenda, allowing for necessary modifications. Lastly, quantitative analysis elements such as frequency or correlation assessments may be included in the evaluation.

The primary goal of engaging with experts is highlighting banks' current efforts in adopting open banking models, particularly in Beyond Banking Ecosystems. The study will explore how to dimensionalize and scale these Ecosystems, prompting banks to consider which use cases they will pursue to unlock Beyond Banking opportunities. Additionally, it will examine the sales strategy and the extent to which Beyond Banking Sales Signals are utilized to enhance profitability. Interviews were conducted with 10 experts from private, regional, and international banks regarding inter-institutional customer matching (Beyond Banking 2.0) and the focused intra-institutional customer matching method (Beyond Banking 3.0). These bank experts and professionals from non-banks, start-ups, and IT or platform service providers in the open and Beyond Banking space were consulted via this methodology. Moreover, the insights gathered from these experts will shed light on key influencing factors and evaluation criteria for devising a scalable BBE, thus providing preliminary validation of the hypotheses. Case studies (Gerring, 2007) will underscore initial efforts toward market acceptance of a scaled Beyond Banking Ecosystem and the corresponding customer data matching.

This survey aims to deliver quantitative and representative evidence (Watson, 2015) indicating the market demand for a scaled Beyond Banking Ecosystem with various stages of expansion alongside the emergence of significant customer benefits. Customers are encouraged to contemplate whether they would be interested in receiving non-banking offerings from their banks and if they are willing to permit banks to utilize and share more data to enable improved, targeted, and personalized marketing strategies.

At least 300 bank customers will be surveyed to ensure a representative customer sample. The following hypotheses will be validated by accepting a scaled BBE of 2.0 to 3.0 and confirming customer-to-customer assignment (C2C) through data analysis, an uncharted market area.

4 Preliminary/Expected results

The initial current market environment and studies confirm the relevance and revenue potential of Beyond Banking. Numerous studies indicate that the banking business model is shifting from purely financial services to a comprehensive everyday provider (Sancar & Sgardelli, 2021). In contrast, regional banks primarily concentrate on establishing a digital ecosystem at the regional level (Knörrer et al., 2021). Europe is falling behind in international comparisons. There is significant activity in beyond-banking case studies worldwide. For example, JP Morgan Chase & Co. is set to launch a full-service travel solution (de Backer et al., 2023). This is also confirmed by case studies on German regional and international banks, especially from Asia (Fastnacht, 2023; Knörrer et al., 2021).

In further research, expert interviews from banking circles should shed light on the provider side's acceptance and development of BBE. Limited knowledge and activities regarding BBE 3.0 and new data-driven sales recommenders (Beyond Banking Sales Signals) are expected, as well as standardized governance concepts in frameworks, blueprints, etc. Furthermore, the initially deductive category approach along the lines of the questions in the pre-prepared interview guide followed an inductive approach, particularly in the category "Revenue potential Beyond Banking," because experts provided much-differentiated feedback on this, which had to be analyzed. In summary, it can be concluded that most experts see (high) potential in BB under the sub-code "Assessment." This means that four experts rate the potential as high to very high. Two experts rate the potential as existing. The other two experts have a differentiated view on this, namely that the monetary revenue is questionable or depends on the use case. However, there is potential for indirect revenue. One expert rates the potential as relatively low, and another expert could not provide an estimate.

In addition to evaluating expert knowledge, customer demand is also essential to find out, with the help of a survey, whether there is any interest in and acceptance of Beyond Banking Services. It is also likely to be researched here that customers are receptive to this new range of services, provided that the added value services are convincing and there is a willingness to transfer personal data through consent.

5 Future development

The compiled results will encompass additional attributes and influencing factors, serving as the foundation for a practical design concept for BBE 3.0, which will be carried out in future development. The concept for creating a scalable BBE 3.0 will be ensured using the Design Science Research (DSR) methodology (vom Brocke & Hevner, 2023).

The approach begins with a systematic review of the literature review. This is followed in several iteration cycles by a sharpened target image, which is helped by expert interviews and validated by a focus group as practical suitability. In addition, an updated framework as a decision support system for creating a thriving Beyond Banking Ecosystem can be built on existing decision support systems.

References

- Creswell, J. W. (2009). Editorial: Mapping the field of mixed methods research. *Journal of Mixed Methods Research*, 3(2), 95–108. https://doi.org/10.1177/1558689808330883/ASSET/1558689808330883.FP.PNG_V03.
- Deloitte. (n.d.). Wie Open Banking die Finanzbranche transformiert | Deloitte Deutschland. Retrieved March 25, 2024, from https://www2.deloitte.com/de/de/pages/financial-services/articles/o_pen-banking.html.
- Fasnacht, D. (2023). Offene und digitale Ökosysteme. Mehrwert durch Branchen- und Technologiekonvergenz. <https://doi.org/10.1007/978-3-658-42494-7>.
- Friese, S. (2016). Qualitative data analysis software: The state of the art. *KWALON*, 21(1). <https://doi.org/10.5117/2016.021.001.005>.
- Gao, J., Choo, K. T. W., Cao, J., Lee, R. K. W., & Perrault, S. (2023). CoAICoder: Examining the Effectiveness of AI-assisted Human-to-Human Collaboration in Qualitative Analysis. <http://arxiv.org/abs/2304.05560>.
- Gerring, J. (2007). *Case study research: Principles and practices*. Cambridge University Press, New York, USA.
- Gómez-Márquez, J. (2023). A new definition and three categories for classifying ecosystems. *Academia Biology*. <https://doi.org/10.20935/ACADBIOL6072>.
- Hevner, A., & Brocke, J. vom. (2023). A Proficiency Model for Design Science Research Education. *Journal of Information Systems Education*, 34(3). <https://aiselaisnet.org/jise/vol34/iss3/2>.
- Hopf, C. (2004). Qualitative interviews: An overview. In U. Flick, E. von Kardorff and I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 203–208). Rowohlt Taschenbuch Verlag.
- Junghanns, H., & Niebudek, M. (2019). Platform Banking & Digital Ecosystems. pwc.de. Retrieved April 5, 2024, from <https://www.pwc.de/de/finanzdienstleistungen/study-platform-banking-and-digital-ecosystems.pdf>.
- Katzenschlaeger, A., Singh, A., Forst, F., & de Backer, P. (2023). Banks as facilitators of everyday life. Retrieved May 2, 2024, from https://www.adlittle.com/sites/default/files/viewpoints/ADL_Banks_as_facilitators_2023.pdf.

- Knörrer, D., Mosen, M., Moormann, J., & Schmidt, D. (2021). *Digitale Ökosysteme: Strategien, KI, Plattformen*. Frankfurt School Verlag/efiport. Frankfurt a.M., Germany.
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse Methoden, Praxis, Computerunterstützung*. Beltz Juventa. Weinheim, Basel. Germany.
- Lamnek, S., & Krell, C. (2016). *Qualitative Sozialforschung: mit Online-Material*. Beltz. Weinheim, Basel. Germany.
- Liang, J. N., & Savage, D. T. (1990). The Nonbank Activities of Bank Holding Companies. *Federal Reserve Bulletin*, 76 (p. 280).
- Mayring, P. (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13th ed.). Julius Beltz. Weinheim. Germany.
- Reiter, A., Stonig, J., & Frankenberger, K. (2024). Managing multi-tiered innovation ecosystems. *Research Policy*, 53(1), 104905. <https://doi.org/10.1016/J.RESPOL.2023.104905>.
- Sancar, B., & Sgardelli, H. (2021). Vertriebschancen in einer Bank durch Nutzung ihrer Daten und ihres Ökosystems. Retrieved March 4, 2025, from https://www.eurogroupconsulting.de/wp-content/uploads/2021/11/DataNavigator_08.pdf.
- Standaert, W., Muylle, S., & Cumps, B. (2020). Opening the gates: A framework for an open banking strategy. *Journal of Digital Banking*, Volume 4/Number 4/Spring 2020, 364-376(13).
- Watson, R. (2015). Quantitative research. *Nursing Standard*, 29(31), 44–48. <https://doi.org/10.7748/NS.29.31.44.E8681>.
- zeb. (n.d.). Dem „Beyond Banking“ fehlen die Business Cases | zeb. Retrieved March 26, 2024, from <https://zeb-consulting.com/de-DE/dem-beyond-banking-fehlen-die-business-cases>.
- Zhu, X., Song, B., Ni, Y., Ren, Y., & Li, R. (2016). Platforms—From the One-Sided Market to the One-Sided + Two-Sided Market. *Business Trends in the Digital Era*, 43–61. https://doi.org/10.1007/978-981-10-1079-8_3.

DOCTORAL CONSORTIUM

BETTER TOGETHER: CONDITIONS FOR SUSTAINABLE INNOVATION AIMED AT IMPROVING QUALITY OF CARE IN HEALTH DELIVERY ORGANIZATIONS

PIETER J. KIEVIT

Open University of the Netherlands, Beta faculty, Heerlen, the Netherlands
vockievit@gmail.com

The resilience of health care organizations over time in a changing and ever challenging socio-economic environment depends on effective innovation. Innovation is based on translating research outcomes into interventions aiming at improving quality of care which is defined as a combination of six dimensions: effectiveness, efficiency, safety, equity, accessibility and patient-centeredness of care. The majority of research papers on innovation and quality improvement deal with the conditions for dissemination and implementation (D&I) but does not address the conditions for long term valorisation of research outcomes or adopting them beyond the implementation period. Factors determining successful D&I are ordered in frameworks consisting of five domains: characteristics of the intervention, internal context, external context, adopters skills and design and management of the implementation process. In general, these frameworks provide a descriptive tool or taxonomy, but miss explanatory power which will eventually be found in a theory of good decisional practice.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.51](https://doi.org/10.18690/um.fov.4.2025.51)

ISBN

978-961-286-998-4

Keywords:

Innovation,
quality of care,
implementation,
health delivery
organizations,
beyond banking ecosystems



University of Maribor Press

1 Introduction

The resilience of health care organizations in a changing and ever challenging socio-economic environment depends on effective innovation. Innovation is based on translating research outcomes into interventions aiming at improving quality of care which is defined as a combination of six dimensions: effectiveness, efficiency, safety, equity, accessibility and patient-centeredness of care. The majority of research papers on innovation and quality improvement deal with the conditions for dissemination and implementation (D&I) but does not address the conditions for long term valorization of research outcomes or adopting them beyond the implementation period. In this proposal, the terms sustainable innovation or sustainable quality improvement will be used for the time being. Factors determining successful D&I are ordered in frameworks consisting of five domains: characteristics of the intervention, internal context, external context, adopters skills and design and management of the implementation process. In general, these frameworks omit the conditions for sustainable innovation. They provide a (limited) descriptive tool or taxonomy, but miss explanatory power which will be found in a theory of good decisional practice.

This research proposal aims at clarifying the concept of sustainability in health care innovation and at constructing and providing a tool for determining an organizations innovating capacity, based on the prevailing operative decision premises. Research will be conducted, based on the process model of Design Science Research Methodology because of the combination of methodological rigor and practical utility.

2 Problem definition

This proposal concentrates on health delivery organizations (aka hospitals). Healthcare organizations are under growing external and internal pressure to become more efficient in containing healthcare costs while delivering the same or better quality of care to an aging population with more complex comorbidities. This demands constant innovation aimed at quality improvement. Health care quality in this context is defined as a combination of six dimensions and their interactions: effectiveness, efficiency, safety, equity, accessibility and patient-centeredness of care (IOM 2001). Health care innovation does not include medical innovation which

focuses on patients and somatic conditions by better diagnostics, evidence based treatments and interactive revalidation supported by shared decision making between patients and professionals. Health care innovation is defined as the creation and implementation of concepts, ideas, technologies and services aimed at processes, structures and outcomes and leading to quality improvement (Donabedian 1966/2005). Both strategies depend on one-another: medical practice takes place within and by means of the health delivery organization and the organization finds her ultimate *raison d'être* in care and cure for patients, but this research proposal is restricted to health care innovation among other reasons because research into medical innovation would require forms of patient related clinical research for which we are neither authorized nor qualified. Dissemination and implementation of innovation is studied by the discipline of implementation science. Implementation science develops (applied) theories and tools to plan, manage and evaluate valorization of research outcome or evidence based knowledge in (among others) health care practice (Wensing et al. 2022; Home page | Implementation Science ([biomedcentral.com](https://www.biomedcentral.com/implementation))).

The conditions for failure or success of dissemination and implementation (D&I) projects are listed in implementation frameworks, arranged in domains. These are: characteristics of the intervention, internal context, external context, adopters skills and quality of the implementation process. Almost all frameworks in implementation science and practice use the five domains, albeit with varying operationalization. Although the total number of such frameworks at present is about 170, only a fraction of these are more or less frequently in use (Damschroder et al. 2009; 2022) and in most cases the application is sub-optimal (Moullin 2020). Not included in most publications on innovation in health care or elsewhere are two factors which are essential when implementing evidence based interventions: sustainability of the intervention (lasting beyond the implementation phase) and the quality of the decisions made in that context which have to be both effective and acceptable.

Sustainability

The first challenge is sustainability. Sustainable innovation depends on more than the correct application of a framework of conditions for managing D&I projects (Fleischer et al. 2015). To begin with, “sustainable innovation” is in dire need of

conceptual clarification. With an implementation science focus, sustainability has been provisionally defined as: “after a defined period of time, the program (...) and/or implementation strategies continue to be delivered and/or individual behavior change is maintained, continuing to produce benefits for individuals/systems” (Moore et al. 2017). In the present literature on innovation and implementation theory, the elements or factors conditional for or leading to sustainable quality improvement in health care or in other sectors of society are either not addressed or treated as very similar to the frameworks for D&I projects (Shelton et al. 2018; Khalil, Kynoch 2021). In a first, exploratory phase, this research proposal will contribute to clarification of the concept of sustainability in innovation, thereby contributing to its applicability and explanatory power. It will list the conditions for sustainability and present them in a framework.

Decision making

The second challenge is decision making. In order to reconcile differences in and between local logics or situational rationalities in D&I projects in complex organizations, decisions have to be both effective in terms of the objective of the implementation and sufficiently acceptable to the internal and external stakeholders. In spite of a multitude of (descriptive) theories about and case-studies of decision making in medical practice and elsewhere, at present there is no broadly accepted standard for “good decisional practice” comparing to the one for “good clinical practice” (WHO 2005). In order to be reliable and effective, decisions on investments reckoning costs and benefits and the expected interacting effects on dimensions of quality will have to meet a set of guidelines, covering the way a decision making process is designed, conducted, monitored, recorded, analyzed and reported. This research proposal has as its main purpose to contribute to a concept of “good decisional practice” as one of the vital factors for successful implementation and a determining factor for the innovation capacity of any organization.

3 Methodology

In this proposal, research interest is limited to the conditions for sustainable innovation, or the durability of innovation and its effects beyond the actual implementation phase and to the standards and procedures for making effective and

acceptable decisions in the process of dissemination and implementation. By developing design and methods, the process model of design science research methodology will be followed (Peffer, Tuunanen, Rothenberger, Chatterjee 2007; Brocke, Hevner, Maedche 2020). This model offers transparency, combining the demands of methodological rigor with practical utility in first designing an artifact which improves something for stakeholders and subsequently empirically investigating the performance of that artifact in its context: “validation in context” (Wieringa 2014 p V. The DSRM process model contains six phases.

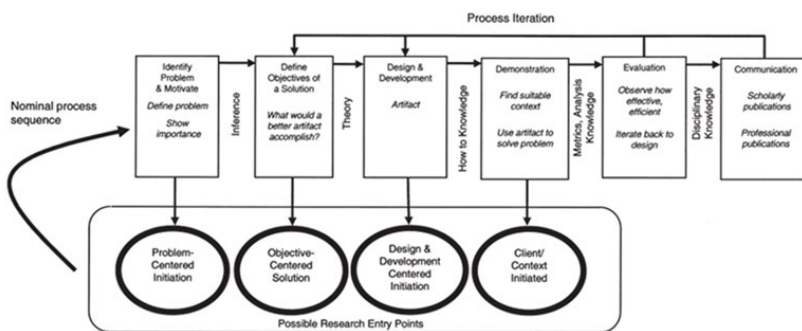


Figure 1. DSRM Process Model

Figure 1: DSRM Process Model

Source: (Peffer et al. 2007)

The six phases of the DSRM model are realized in five research projects, each answering a specific research question, including a sixth phase, presenting the outcomes to the community of health care professionals and valorization specialists. In applying this model as blueprint for the present research proposal, the coherence between the phases in the sequence becomes explicit, the phases can be evaluated in their merits and the phases can be repeated as an additional quality check. In this way, the model contributes to the objectivity, reliability and validity of the research project and its components.

3.1 Identifying the problem: what are the conditions for successful D&I projects and what are the conditions for sustainability in innovation?

The process of translating the outcome of (applied) research into evidence based interventions leading to quality improvement has been the object of research by (for instance) the discipline of implementation science for a long time (Peters 2013).

This research project will start with a literature search, leading to a survey of actual facilitators and barriers to innovation projects, comparing them with the elements of standard frame-works, based on empirical studies and reviews. Inclusion will not be restricted to papers referring to the health care sector as windfalls and setbacks or complications with dissemination and implementation are universal and not typical for care providers.

The second literature review searches for definitions of sustainability in innovation and the validated knowledge concerning the conditions for this sustainability. Given the conceptual ambiguity of sustainable implementation as a process and desired outcome and the lack of standardized frameworks in this field, the review will concentrate on available knowledge on the topic since 2015. In order to make the review up-to-date with expert opinion, a number of additional interviews with a limited number of innovation managers and specialists will be included and reported in the publication.

Both questions will be answered using grounded theory as a method for reviewing literature in qualitative research (Wolfswinkel 2011).

Outcome: The reviews will lead to an understanding of the actual use of implementation frameworks, the wide variety of operationalization of the categories and of the large number of various elements determining sustainability of quality improvement and the organizational resilience through innovation.

3.2 Objectives for a solution: what are the relevant aspects of a theory of decision making?

Even the best implementation framework as blueprint and management tool requires adequate decision making in preparation and management of a project or program, especially in a complex organizational context and when solid investments are required. Decisions have to be both effective and accepted in dealing with the risks and opportunities of the investments in dimensions of health care quality and their interactional effects. Not the quality of a tool like a framework but the handling of that tool determines its effectivity in the implementation process and this handling depends on adequate decision making, leading to decisions which are both acceptable to the organization and effective in the process of D&I.

In this process, different interests based on situational rationalities and local logics have to be weighed, compared and acknowledged or rejected, hence decided upon in order to implement and adopt interventions aimed at quality improvement. In order to describe and analyze the decision making process itself and its contribution to the innovation readiness of an organization, this research proposal relies on Luhmann's theory of the organization as a special kind of social system (Luhmann 1985; 2000, also Schoeneborn 2011; Blaschke 2012; Dobusch 2015).

In this theory, a social system is nothing but the totality of all communication taking place, reflecting the process of autopoiesis or development of identity in self-awareness expressed internally—defining the system's elements and their possible relationships—and externally in temporal and spatial relationships: to the system's remembered past and expected future and to the environment it perceives as “alter” or the other. In this experience of the temporal and spatial other as alter, a system experiences itself simultaneously in its identity as auto. In Luhmann's theory, the organization is denied an ontological status (or at least this status is not granted). Instead, organizations are conceived of as a specific kind of social system characterized by the fact that all communication is instrumental and aimed at decision making (Luhmann 2018 p. 49; Seidl 2005; Aal 2022). Decisions are reached by exchanging and confronting arguments that derive their impact from decision premises (Luhmann 2018 p. 181 ff) which form the organization's set of genes and provide the only way an organization can define itself and reduce the limitless complexity of the possible to a form of imaginable and thereby manageable reality.

Alternative theories of decision making – presupposing the organization as ontological context of communicative interaction, resulting – or not resulting – in a decision, lead to the problem that the framework of conditions for innovation presented in the classification of objects in a domain of interest in the shape of a taxonomy (Nickerson 2017; Michie 2013) does in no way determine the communicative performance or the quality of decision making and its outcome, nor the other way around. Taxonomies provide a descriptive model, what is needed, is an explanatory one. Based on Luhmann's theory, we will propose that the referral to decision premises to add impact to argument reveals the organizations propensity to or capacity for innovation or conservation by the share nature of the premises as a set of organizational genes.

3.3 Design and development: the artifact or an instrument to determine a system's innovation readiness.

How can we measure an organizations willingness and capacity to innovate? The challenge in this phase of research is the selection of a communicative situation (or discourse) and constructing an artifact in combination with an appropriate methodology to analyze innovation capacity or innovation readiness (Allen 2017 p. xxxvi/56 ff). In order to construct the artifact, a number of decision premises will be selected and presented. They will be derived from the three areas where complexity reduction and risk management are vital and which are crucial in any form of innovation or changing current practice: power of decision making, costs and benefits and identity. The premises as the genes of an organization become manifest though DNA/RNA actualization in decisions on project planning, managing the process and reporting on the outcome. This process can be visualized and analyzed by applying means and methods of pragmatics or discourse analysis based on applied linguistics. The outcome will be presented in a coding scheme (Schreier 2012 p. 58-79) which forms the central tool for the analysis and interpretation of discourse or the exchange of arguments leading to a decision in a specific situation. With this instrument or artifact the quality of decisions made in specific D&I projects and programs can be appraised, contributing to an assessment of the probable outcome and determining the strength and restrictions of the organization's innovation capacity.

3.4 Demonstration: how to put the artifact to the test? Proof of concept.

In order to test the validity of the artifact, it will be applied in a sufficient number of multi-case studies (Yin 2018) to score the operative set of decision premises of a team with a well-established and confirmed reputation and track record for sustainable innovation. The hypothesis to be tested is, that there will be a positive correlation between the “decision premises profile” of the teams and their proven innovation capacity. The way to establish the correlation will be by structural equation modelling, linking latent variables in the domain of decision premises to well established variables like a team’s proven innovation track record (Cook 2011; Phillips 2002; Kaplan 2000) by means of scheme analysis of one or more instances of yet to decide decision making discourse. The case studies will be executed with assistance from a group of quality specialists in the community of Dutch top clinical teaching hospitals. The number, nature and location of the teams has yet to be decided, based on a power analysis. Aim of the set of case studies is to calibrate the tools by measuring and comparing the realized surface set of decision premises with the well-established track record of the teams over a large enough set of data. It will be followed by the adjustment of the tools if necessary.

Outcome: a (presumably) valid tool to diagnose social systems – or innovation teams – in their innovation capacity as expressed in the predominant quality of decision making.

3.5 Evaluation : how to put the artifact to the test? Communication in action

Top clinical teaching hospitals are a worldwide category of health delivery organizations providing complex medical care, acting as centers of medical teaching and education and contributing to patient related clinical research. In the Netherlands the 27 top clinical teaching hospitals form a specific group, known as the Samenwerkende Topklinische Ziekenhuizen or STZ (Samenwerkende Topklinische Ziekenhuizen - STZ). This association will be asked to provide a number of sites to test the instrument on validity, reliability and effectiveness (Leung 2015). Sites could be either innovation projects or implementation teams. The tests will take place in the shape of action research (O’Leary 2017, p 415 ff), the researcher observing and actively participating in an Observe-Reflect-Plan-Act cycle in a limited

number of actual innovation projects aimed at sustainable implementation of quality improving interventions. Aim is to observe communication-in-action, being the decision making process where arguments are presented, exchanged and reacted upon, leading to a decision which is both effective in the project and acceptable to the team as a social system. Through this participating action research, it will be possible to scrutinize the actual decision making processes in innovation and the strategies for conferring pragmatic power to arguments, determining the quality of decision making in a particular case or a number of cases (McCarthy, Matthiessen, Slade 2010). In order to guarantee the appropriate level of research integrity and quality, this phase of the project will be monitored by a limited team of experts in action research, hospital management and decision making. In this way, the artifact and its efficacy can be tested, either validating or falsifying its claim to be able to assess an organizations innovation readiness and the quality of its decision making process.

3.6 Communication

Each of the five phases of this research project will lead to a publication in a peer-reviewed international journal in the domain of implementation science and/or health care management. On top of that, a number of practice oriented publications and conferences will be delivered in order to introduce the concepts of sustainability in (health care) innovation and effective decision making or good decisional practice by deploying decision premises to a relevant audience from management, education and consultancy.

4 Preliminary/Expected results

Implementing the research proposal will lead to a number of outcomes adding to facilitating the implementation of evidence based interventions aimed at quality improvement in health care.

4.1 Scientific importance and added value

As yet, implementation theory has not presented a dedicated framework of conditions for planning, monitoring and evaluating sustainable innovation of interventions aimed at quality improvement in health care. The present research

proposal will provide an outline of conditions for sustainable innovation, based on literature research and the expert opinion of innovation managers. The outline will be presented in the form of a framework. It is not just the framework of facilitator and barrier factors which determines the outcome of health care innovation, it is also – and predominantly – the process of decision making and the effectiveness and acceptability of the decisions resulting from the process. Frameworks are a necessary condition but they are not a sufficient cause, they provide a taxonomy but have no explanatory power. The present research proposal will provide a theory of good decision making based on Luhmann's theory of social systems. This research will develop and test a set of decision premises meant to describe and analyze the decision process as such, dedicated to innovation and implementation of quality improvement. The set of decision premises will not just analyze the process of decision making but will also be used to assess the innovation capacity or innovation readiness of the organization as a social system thereby providing a tool for feasibility studies and team development.

4.2 Expected scientific output

This research proposal will extend the concept of frameworks for D&I projects into the area of sustainable innovation by providing a set of critical conditions for sustainability in health care. It will elaborate further on Luhmann's theory of the organization as a specific form of social system by operationalizing the crucial element of decision premises as organizational genes determining its innovation capacity or its innovation readiness. The output will consist of a number of research papers (at least one per phase) in international peer-reviewed journals in the field of health care, innovation and implementation sciences and of a number of professional publications in media dedicated to research valorization and the translation process from knowledge to practice. Aim is to bring the importance of sustainable innovation and the conditions for sustainability as well as the conditions for "Good Decisional Practice" as a parallel to "Good Clinical Practice" to the attention of the health care and implementation research community.

4.3 Practical significance

The development of a framework for sustainable implementation of health care innovation will contribute to successful planning, managing, monitoring and evaluating such projects as part of a continuous innovation PDCA-cycle. The validated set of decision premises as criteria for effective decision making will help to analyze, monitor and if necessary improve the decision making process, fulfilling one of the basic conditions for successful planned innovation, either incremental or disruptive. In focusing on the explanatory power of the decision making process to clarify barriers and facilitators for sustainable innovation, this research will contribute to the development of standardized guidelines for "good decisional practice".

5 Future development.

Both the implementation framework for sustainable innovation and the review framework for making decisions which are acceptable as well as effective will have positive effects on health care practice when converted into applications for everyday use in planning and implementing quality strategies, feasibility studies, project planning and management and outcome evaluation on behalf of a continuous PDCA improvement cycle.

Applying the frameworks into tools, guidelines, standardized action plans and evaluation formats will take place in close proximity to health delivery situations in mixed teams consisting of health care practitioners, managers and implementation specialists.

References

- Aal, Erik B.W. The significance of Luhmann's theory on organisations for project governance. *Project Leadership and Society* 3 (2022) 100070. Corvinus University of Budapest, Institute of Strategy and Management. (<https://doi.org/10.1016/j.plas.2022.100070>).
- Allen, M. (Ed.) (2017). *The SAGE Encyclopedia of Communication Research Methods Volume 1*. Thousand Oaks, California, SAGE Publications, Inc.
- Blaschke, Steffen, Schoeneborn, Dennis, Seidl, David (2012). Organisations as networks of communication episodes: Turning the Network perspective inside out. In *Organization Studies*, 33(7), 879-906.

- Brocke J. vom, Hevner A., Maedche A. (2020). Introduction to Design Science Research. In: vom Brocke J., Hevner A., Maedche A. (eds) *Design Science Research. Cases*, Cham. https://doi.org/10.1007/978-3-030-46781-4_1
- Cook, Guy (2011). Discourse analysis. In James Simpson (Ed). *The Routledge Handbook of Applied Linguistics*, p. 431-445. London/New York, Routledge.
- Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.* 2009;4(1):1–5.
- Damschroder, Laura J., Reardon, Caitlin M., Widerquist, Opra, Marilla A., Lowery, Julie. The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation Science* (2022) 17:75 (<https://doi.org/10.1186/s13012-022-01245-0>).
- Dobusch, L. Schoeneborn, D. (2015). Fluidity, Identity and Organizationality, the communicative constitution of Anonymous. In *Journal of Management Studies* 52:8 December 2015. (<https://doi.org/10.1111/joms.12139>).
- Donabedian, Avedis. Evaluating the Quality of Medical Care. In *Milbank Q.* 2005 Dec; 83(4): 691–729. doi: 10.1111/j.1468-0009.2005.00397.x
- Fleischer A.R., Semenic S.E., Ritchie J.A., Richer M.C. & Denis J.L. The sustainability of healthcare innovations: a concept analysis. *Journal of Advanced Nursing* 2015 71(7). (<https://doi.org/10.1111/jan.12633>).
- Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington (DC): National Academies Press (US); 2001. PMID: 25057539.
- Kaplan, David (2000). *Structural Equation Modeling: Foundations and Extensions*. Thousand Oaks, California, SAGE Publications, Inc.
- Khalil, Hanan and Kynoch, Kathryn. Implementation of sustainable complex interventions in health care services: the triple C model. *BMC Health Services Research* (2021) 21:143. (<https://doi.org/10.1186/s12913-021-06115-x>).
- Leung L. Validity, reliability, and generalizability in qualitative research. *J Family Med Prim Care* 2015;4:324-7 (<https://doi.org/10.4103/2249-4863.161306>).
- Luhmann, N. (1984). *Soziale Systeme. Grundriß einer allgemeinen Theorie*, Frankfurt: Suhrkamp.
- Luhmann, N. (2000). *Organisation und Entscheidung*, Westdeutscher Verlag, Opladen/Wiesbaden.
- Luhmann, Niklas, Becker, Dirk (2018). *Organization and decision*. Cambridge, Cambridge University Press.
- McCarthy, Michael, Matthiessen, Christian, Slade, Diana (2010). *Discourse Analysis in: Norbert Schmitt (ed.) An Introduction to Applied Linguistics*, London, Hodder & Stoughton Ltd.
- Michie, Susan et al. (2013). The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. In *Annals of Behavioral Medicine*, Volume 46, Issue 1, August 2013, Pages 81–95, <https://doi.org/10.1007/s12160-013-9486-6>.
- Moore JE, Mascarenhas A, Bain J, Straus SE. Developing a comprehensive definition of sustainability. *Implement Sci.* 2017 Sep 2;12(1):110. doi: 10.1186/s13012-017-0637-1. PMID: 28865479; PMCID: PMC5581411.
- Moullin, Joanna C. et al. (2020). Ten recommendations for using implementation frameworks in research and practice. (<https://doi.org/10.1186/s43058-020-00023-7>).
- Nickerson, Robert C. Varshney, Upkar and Muntermann, Jan (2017). *A method for taxonomy development and its application in information systems*. Augsburg, Opus Bibliothek 93439.pdf.
- Nundy, Shantanu, Cooper, Lisa A. , Mate, Kedar S. The Quintuple Aim for Health Care Improvement. A New Imperative to Advance Health Equity in *JAMA*. 2022;327(6):521-522. Doi:10.1001/jama.2021.25181.
- O'Leary, Zina (2017). *The Essential Guide to Doing Your Research Project*. Thousand Oaks, Sage.

- Peffer, Ken, Tuunanen, Tuure, Rothenberger, Marcus A., Chatterjee, Samir. A Design Science Research Methodology for Information Systems Research in *Journal of Management Information Systems* / Winter 2007–8, Vol. 24, No. 3, pp. 45–77. <https://DOI.org/10.2753/MIS0742-1222240302>.
- Peters, David H., Adam, Taghreed, Adam, Alonge, Olakunle, Agyepong, Irene, Aku, Tran, Nhan. Implementation research: what it is and how to do it? in *BMJ* 2013;347:f6753.
- Phillips, Nelson, Hardy, Synthia (2002). Discourse analysis. Investigating processes of social construction. Qualitative research methods series 50. Thousand Oaks, Sage publications.
- Schoeneborn, D. (2011). Organization as communication, a Luhmannian Perspective. In *Management Communication Quarterly*, 25(4), 663-689. (<https://doi:10.1177/0893318911405622>).
- Schreier, Margrit (2012). *Qualitative Content Analysis in Practice*. Los Angeles, London, Sage Publications.
- Seidl, David (2005). Luhmann's theory of autopoietic social systems. Ludwig-Maximilians-Universität München.
- Shelton, Rachel C., Rhoades Cooper, Brittany and Wiltsey Stirman, Shannon. The Sustainability of Evidence-Based Interventions and Practices in Public Health and Health Care. *Annu. Rev. Public Health* 2018. 39:55–76. (<https://doi.org/10.1146/annurev-publhealth040617-014731>).
- Wensing, M. et al. Evidence for objects of implementation in healthcare: considerations for Implementation Science and Implementation Science Communications. In *Implementation Science* (2022) 17:83 (<https://doi.org/10.1186/s13012-022-01249-w>).
- WHO (2005). Handbook for Good Clinical Research Practice (GCP) Guidance for Implementation (untitled (who.int))
- Wieringa, Roel J. (2014). *Design Science Methodology for Information Systems and Software Engineering*. Heidelberg, Springer.
- Wolfswinkel, Joost F., Furtmueller, Elfi, Wilderom, Celeste P.M.: Using grounded theory as a method for rigorously reviewing literature in: *European Journal of Information Systems* 2011, 1-11.
- Yin, Robert K. (2018). *Case Study Research and Applications: Design and Methods*. Thousand Oaks, California, SAGE Publications, Inc.

DOCTORAL CONSORTIUM

SAFEGUARDING PUBLIC VALUES BY DESIGN: A SOCIO-TECHNICAL APPROACH TO DIGITAL TWINS IN REGIONAL SPATIAL PLANNING UNDER THE ENVIRONMENTAL PLANNING ACT

JOHN VAN MEERTEN

HU University of Applied Sciences Utrecht, Digital Ethics, Utrecht, the Netherlands
john.vanmeerten@hu.nl

In the context of climate change and housing challenges, Digital Twins (DTs) are a promising technology. However, there is little experience regarding good use of this technology. This research proposal explores how DTs can safeguard public values in spatial planning under the Environmental Planning Act (Omgevingswet). Combining a socio-technical perspective with Value-Sensitive Design (VSD), the study focuses on integrating values into DT technology. Expected studies include identifying key values, designing a normative framework, and measuring its impact. This contributes to improved decision-making and spatial planning by regional governments.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.52](https://doi.org/10.18690/um.fov.4.2025.52)

ISBN

978-961-286-998-4

Keywords:

digital twins,
public values,
spatial planning,
provincial government,
environment act



University of Maribor Press

1 Introduction

The Netherlands is facing multiple societal challenges, such as climate change, requiring adaptation to risks like flooding and heat stress. The housing market is strained by the demand for affordable and sustainable homes. Moreover, the transition to renewable energy and reduction of CO₂-emissions marks a major energy shift (IPLO, 2024b).

To help tackle these challenges, the Environmental Planning Act (Omgevingswet), effective from January 1, 2024, was introduced (Rijksoverheid, 2024). This law simplifies and combines environmental regulations, making it easier to address complex issues (IPLO, 2024a). It promotes an integrated approach to national environmental policy, crucial for addressing issues like housing. Provinces play a key role in translating national policies into regional actions, empowered by the Act to develop tailored environmental visions and plans.

Supporting this, digital twins (DTs) are increasingly used at the provincial level for spatial planning, providing dynamic, (real-time) digital replicas of physical systems. These DTs, enhanced by the Internet of Things and AI, offer predictive capabilities. For instance, the Municipality of Almere and the Province of Flevoland are experimenting with DTs in various applications (GEONOVUM, 2024), while the Province of Utrecht and the Province of South Holland use DTs for urban development (Provincie Utrecht, 2024) and the "Samen fietsen" project (Provincie Zuid-Holland, 2024).

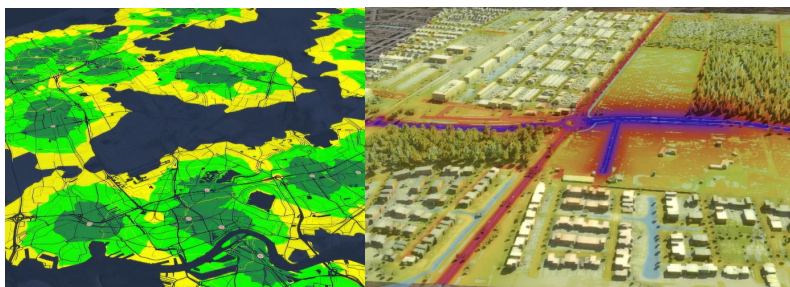


Figure 1: DT Example: Left: bike accessibility simulation, right: noise simulation

DTs also facilitate navigating the complexities of the Environmental Planning Act, modernizing the planning process through integrated decision-making and stakeholder collaboration (i.e., DTs can simulate policy impacts, helping policymakers make informed, balanced decisions). However, their deployment in complex environments poses challenges, particularly in aligning with public values (e.g., public safety). Misalignments can result in unfair outcomes, erode trust in decision-making, and provoke resistance to policy implementation. These risks highlight the need for a framework to align DTs with public values, preventing consequences like environmental damage, inequality, or inefficiency. (Friedman et al., 2013; Spiekermann, 2012).

Ignoring public values in spatial planning can have serious repercussions. A historical example is Robert Moses' parkway designs in New York, which restricted low-income and minority groups from accessing public beaches by designing low overpasses that buses couldn't pass through (Valentine, 2020). In the Netherlands, recent urban densification efforts aimed at accommodating population growth and maximizing land use have met resistance from residents (Claassens et al., 2020), who fear reduced public spaces, altered community dynamics, and deteriorating living conditions. This opposition underscores the challenge of aligning spatial development with community needs, highlighting the importance of incorporating public values into planning.

This research advocates operationalizing a socio-technical perspective (i.e., considers the interaction between people, technology, and the social environment) through Value Sensitive Design (VSD) to address these risks. VSD incorporates human values into technology development across conceptual, empirical, and technical investigations, ensuring ethical and societal considerations are integrated into technology design and usage (Friedman et al., 2013). This approach aims to create a normative framework that upholds public values in DT applications, promoting responsible technology adoption. Collaborations are arranged with partners, including the Province of Utrecht, Vereniging van Nederlandse Gemeenten (VNG), Data- en Kennishub Gezond Stedelijk Leven (DKH-GSL), and Netherlands 3D.

MRQ: What conditions are necessary for Digital Twins to safeguard public values by design in regional spatial planning under the Environmental Planning Act?

Sub-questions:

1. What public values emerge and how are they prioritized in regional spatial planning under the Environmental Planning Act? (**RQ1**)
2. What is the state-of-the-art in applying Digital Twins from a socio-technical perspective in regional spatial planning under the Environmental Planning Act? (**RQ2**)
3. How can a normative framework be designed to safeguard public values in the use of Digital Twins in regional spatial planning under the Environmental Planning Act? (**RQ3**)
4. How can this normative framework be implemented in regional spatial planning under the Environmental Planning Act? (**RQ4**)
5. How can the impact of the normative framework be measured and what effects does it have in regional spatial planning under the Environmental Planning Act? (**RQ5**)

3 Methodology and Expected results

The research has five studies: RQ1 explores public values, RQ2 reviews state-of-the-art DT, RQ3 designs a normative framework, RQ4 develops a code of conduct and prototype to operationalize the framework, and RQ5 tests its impact and effect, see the following figure:

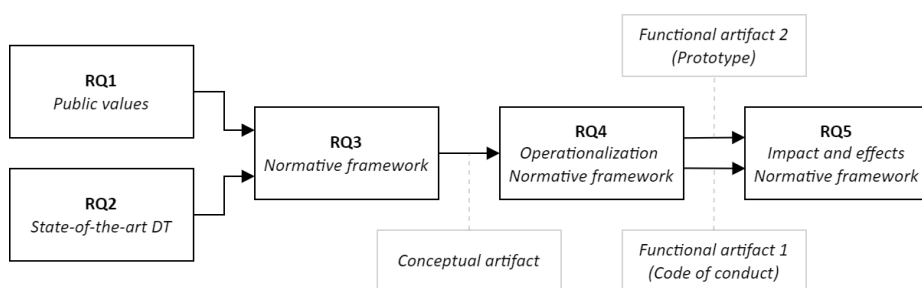


Figure 3: Research overview

Study 1 (RQ1): What public values emerge and how are they prioritized in regional spatial planning under the Environmental Planning Act?

The first study identifies and prioritizes public values (Bozeman, 2019; Jørgensen & Bozeman, 2007).

The identification of public values will use two methods. The first is a scoping review, following the framework outlined by Paré (2015), which will survey literature from the Netherlands and countries with similar (cultural) contexts at both conceptual and empirical levels. This phase aligns with VSD's *conceptual investigation*, uncovering stakeholders' values

Subsequent focus groups (Morgan, 1996) will include diverse stakeholders—civilians, policymakers, council members, and data scientists—who will first be asked to identify what they consider to be public values and elaborate on their meaning. To facilitate this discussion, we will use value scenarios, a proven VSD-technique (Friedman & Hendry, 2019). Participants will then prioritize public values using Q-methodology (Brown, 1996), a tool for collecting subjective data to capture diverse perspectives. This empirical phase will continue until data saturation is achieved, with findings validated by multiple senior researchers, aligning with the *empirical investigation* of VSD. This structured approach ensures a comprehensive understanding of public values into spatial planning initiatives.

Study 2 (RQ2): What is the state-of-the-art in applying Digital Twins from a socio-technical perspective in regional spatial planning under the Environmental Planning Act?

NASA coined DT technology in the 1960s, and it has advanced since (El-Agamy et al., 2024). Understanding the current state-of-the-art in DT technology and its application within this study's context is crucial. To achieve this, two methods will be used. First, a qualitative systematic review following Paré's (2015) methodology will be conducted. This approach will consult empirical sources and provide a narrative synthesis, examining DTs through a socio-technical lens to understand both technical and social dimensions.

Second, a survey based on Krosnick's methodology (Krosnick, 1999) will be administered to stakeholders to explore how DT technology is currently used. This survey will leverage the extensive network of the supervision team (DKH-GSL, VNG, Netherlands 3D, Province of Utrecht, and others) to ensure a broad and

representative sample. The survey will gather data on current practices, challenges, and perceptions regarding DTs in regional spatial planning.

These methods offer a holistic view of DT technology and its socio-technical implications. The results will inform the design of a normative framework in subsequent research.

Study 3 (RQ3): How can a normative framework be designed to safeguard public values in the use of Digital Twins in regional spatial planning under the Environmental Planning Act?

The results of studies one and two will be integrated to develop a normative framework (we define a normative framework as the standards and criteria to which a technology must adhere), working from both a scientific knowledge base and a practical perspective, as outlined in Hevner's (2004) Design Science Research (DSR) model. Leveraging Peffers' (2007) interpretation of DSR, along with the results from studies one and two, the development process of the normative framework will start from the design and development-centered approach step.

Three methods will be employed: a scoping review, in-depth interviews, and focus groups. The scoping review will identify and analyse methods for building a normative framework. This review, part of the *conceptual investigation* phase of VSD, will draw from the academic and practical knowledge base to pinpoint successful principles and methodologies.

In-depth interviews with technical-ethical experts will explore how public values relate to the components of digital twins and how these values can be effectively incorporated into the normative framework. This step falls under the *empirical investigation* phase, focusing on understanding the value-technology relationship through empirical data.

Finally, focus groups with stakeholders will refine the initial draft of the normative framework. These groups will assess the alignment of design choices with the target audience and ensure the language used is clear and consistently interpreted. This phase aligns with the *technical investigation* phase of VSD.

This study will construct and refine the normative framework (conceptual artifact).

Study 4 (RQ4): How can this normative framework be implemented in regional spatial planning under the Environmental Planning Act?

Study four explores the operationalization of the normative framework developed in Study three, specifically for regional spatial planning under the Environmental Planning Act. The study focuses on advancing the normative framework into a code of conduct (functional artifact 1) and creating a prototype (functional artifact 2) for semi-automatically validating DTs. These developments occur within the Develop/Build phase as described by Hevner et al. (2004) and the Design & Development phase by Peffers (2007), in alignment with VSD principles.

The study consists of two main phases. The first uses focus groups Morgan (1996) to explore the application of the normative framework in regional spatial planning. Participants, including citizens, policymakers, and council members, will discuss how the framework can guide the manual validation of DTs, focusing on technology, process, and people. This socio-technical approach ensures a holistic examination of the framework's practical application, culminating in the creation of a code of conduct.

The second phase is the development of a prototype (functional artifact 2) for the semi-automatic validation of DTs, aimed at assessing their compliance with public values. Building on previous studies, the prototype will be developed in accordance with the phases outlined by Hevner et al. (2004) and Peffers (2007), this aligns with the *technical investigation* of VSD.

The goal is to produce both a code of conduct and a prototype for DT validation, both grounded in the normative framework established in previous studies.

Study 5 (RQ5): How can the impact of the normative framework be measured and what effects does it have in regional spatial planning under the Environmental Planning Act?

In this study, case studies and/or projects will be selected to implement the developed code of conduct (functional artifact 1) and prototype software (functional

artifact 2). Various regional projects in the Netherlands, such as the Geodan project in North Brabant (*Provincie Noord-Brabant*, 2023), the optimal climate-proof design by RIVM and Tygron (RIVM, 2021), and Dutch Metropolitan Innovations (DMI) in Flevoland (Mureau, 2023), provide potential case study sites. The specific cases will be selected from one of our partners mentioned earlier.

A case study offers a unique opportunity to examine the prototype within its natural context, capturing the dynamics and complexities that experimental settings often lack. It allows for multiple perspectives, offering a deep understanding of real-world phenomena (Runeson & Höst, 2009).

This research adopts an exploratory approach to understand the effects and implications of both artifacts. Qualitative methods (interviews with direct and indirect stakeholders, including policymakers and community representatives) and quantitative methods (metrics comparing previously developed digital twins with those developed using the prototype) will be used. Information will be collected before and after the prototype's application to enable accurate comparisons. These methods align with the *technical investigation* of VSD, examining how technology safeguards public values from a socio-technical perspective.

This study will enhance our understanding of how the code of conduct and the prototype functions within regional spatial planning and the Environmental Planning Act, providing insights into their impact and effectiveness.

Functional artifact 1 (code of conduct) can be used in both the design and operation phases of a DT, while functional artifact 2 (prototype) is used once a DT is in operation. These artifacts create the conditions to safeguard public values in DTs by design, see the following figure:

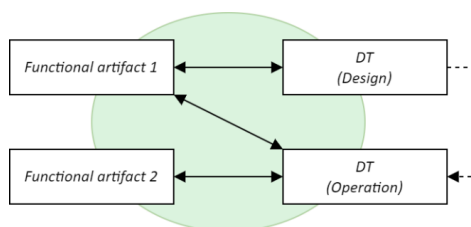


Figure 3 - Functional artifact 1 (code of conduct) and functional artifact 2 (prototype)

4 Scientific and social impact

The primary scientific contribution lies in the development (RQ1, RQ2, RQ3), operationalization (RQ4), and validation (RQ5) of a normative framework with a socio-technical perspective through VSD. This interdisciplinary approach addresses a crucial gap in existing literature, where socio-technical approaches often fail to adequately safeguard public values (Abbas & Michael, 2023; Baxter & Sommerville, 2011). Recent studies have emphasized the urgency of integrating public values systematically into socio-technical designs to address societal challenges and ethical concerns (Royakkers et al., 2018; Sattlegger, 2024). By systematically incorporating values into the design and implementation of DTs, this research advances the fields of socio-technical systems, VSD and DTs. The integration of VSD with a socio-technical perspective offers a novel theoretical contribution. VSD's emphasis on values throughout the design process aligns well with the complex and dynamic nature of regional policy making, thereby enhancing the theoretical understanding of how values can be systematically integrated into digital technologies. This theoretical advancement has broader implications for the field of information systems, particularly in terms of ethical considerations and the societal impact of digital innovations.

Moreover, the research provides a comprehensive analysis of the state-of-the-art of DTs, revealing their current applications and limitations within regional-level spatial planning. This detailed overview contributes to the academic discussion by highlighting the technological and methodological advancements required for the effective deployment of DTs in complex policy environments (Yao et al., 2023).

The iterative design methodology employed in this research, encompassing conceptual, empirical, and technical investigations (VSD (Friedman et al., 2013)) as well as a DSR approach, demonstrates a rigorous scientific approach. This methodology not only contributes to the robustness of the normative framework but also facilitates its empirical validation within the context of the Environmental Planning Act. The scientific relevance of this research lies in the development of the first scientifically validated and operationalized framework for DTs, which accounts for diverse stakeholders, societal challenges, and values, specifically tailored for use by a regional governmental agency.

The societal impact of this research is significant, given the complexity and challenges faced by regional governments in managing spatial planning. The implementation of DTs, guided by a normative framework that safeguards public values, can enhance the decision-making process in regional policy making. By simulating the potential consequences of policy decisions, DTs enable policymakers to make more informed, transparent, and equitable choices, improving the quality of life for citizens.

One of the key societal benefits of this research is its potential to foster social cohesion and equity. By ensuring that diverse public values are considered in spatial planning, the research promotes inclusive and participatory governance. This is particularly important in addressing the needs of various demographic groups, thereby reducing inequalities in access to resources and opportunities. The emphasis on public trust and transparency in the decision-making process further enhances civic engagement and support for policy initiatives.

Environmental sustainability is another aspect of the societal impact. The integration of environmental considerations into spatial planning ensures that DTs help promote sustainable development practices, mitigate environmental degradation, and protect natural resources for future generations. By enabling policymakers to assess the environmental impact of their decisions, DTs can contribute to more sustainable and responsible governance.

The research also addresses practical challenges such as infrastructure overload, resource shortages, and affordable housing. By providing functional artifact 1 and 2 for validation and optimizing policy solutions, DTs help regional governments devise innovative strategies to, e.g., manage urban growth, improve transportation systems, and promote affordable housing options. This capability is particularly valuable in rapidly expanding cities, where effective spatial planning is crucial for maintaining the quality of life.

In the context of the Environmental Planning Act, the research supports the legislation's goal of integrated and collaborative decision-making among stakeholders. By visualizing differing interests and perspectives, DTs facilitate a more comprehensive understanding of stakeholder needs, thereby enhancing the legitimacy and acceptance of policy decisions.

This research not only advances scientific knowledge in the field of DTs, VSD and socio-technical systems but also delivers tangible social benefits by enhancing the effectiveness, transparency, and inclusivity of regional policy making. By addressing technical and social dimensions, the research ensures that the deployment of DTs in spatial planning aligns with public values, contributing to more sustainable and equitable urban development.

5 Future development

This chapter is not yet applicable, as I have only recently started my PhD research. At this stage, I am not yet able to provide an answer to this chapter.

References

- Bozeman, B. (2019). Public values: citizens' perspective. *Public Management Review*, 21(6), 817–838. <https://doi.org/10.1080/14719037.2018.1529878>
- Brown, S. R. (1996). Q Methodology and Qualitative Research. [Http://Dx.Doi.Org/10.1177/104973239600600408](http://Dx.Doi.Org/10.1177/104973239600600408), 6(4), 561–567. <https://doi.org/10.1177/104973239600600408>
- Claassens, J., Koomen, E., & Rouwendal, J. (2020). Urban density and spatial planning: The unforeseen impacts of Dutch devolution. *PLOS ONE*, 15(10), e0240738. <https://doi.org/10.1371/JOURNAL.PONE.0240738>
- El-Agamy, R. F., Sayed, H. A., AL Akhatatneh, A. M., Aljohani, M., & Elhosseini, M. (2024). Comprehensive analysis of digital twins in smart cities: a 4200-paper bibliometric study. *Artificial Intelligence Review* 2024 57:6, 57(6), 1–54. <https://doi.org/10.1007/S10462-024-10781-8>
- Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design*. The MIT Press. <https://doi.org/10.7551/mitpress/7585.001.0001>
- Friedman, B., Kahn, P. H., Borning, A., & Hultgren, A. (2013). Value Sensitive Design and Information Systems. *Philosophy of Engineering and Technology*, 16, 55–95. https://doi.org/10.1007/978-94-007-7844-3_4/FIGURES/5
- GEONOVUM. (2024). *Digital Twins | Geonovum*. <https://www.geonovum.nl/themas/digital-twins>
- Hevner, March, Park, & Ram. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75. <https://doi.org/10.2307/25148625>
- IPL0. (2024a). *Maak kennis met de Omgevingswet | Informatiepunt Leefomgeving*. <https://iplo.nl/regelgeving/omgevingswet/introductie/maak-kennis-omgevingswet/>
- IPL0. (2024b). *Maatschappelijke opgaven | Informatiepunt Leefomgeving*. <https://iplo.nl/thema/maatschappelijke-opgaven/>
- Jørgensen, T. B., & Bozeman, B. (2007). Public Values. *Administration & Society*, 39(3), 354–381. <https://doi.org/10.1177/0095399707300703>
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50, 537–567. <https://doi.org/10.1146/ANNUREV.PSYCH.50.1.537>
- Morgan, D. L. (1996). Focus groups. *Annual Review of Sociology*, 22(Volume 22, 1996), 129–152. <https://doi.org/10.1146/ANNUREV.SOC.22.1.129/CITE/REFWORKS>

- Mureau, F. (2023). *Ontwikkeling Digital Twin voor Flevoland en Almere van start – DMI-ecosysteem*. <https://dmi-ecosysteem.nl/blog/ontwikkeling-digital-twin-voor-flevoland-en-almere-van-start/>
- 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>
- Peffer, 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>
- Provincie Noord-Brabant. (2023). <https://www.geodan.nl/projecten/waar-wij-aan-werken/digital-twin-provincie-noord-brabant/>
- Provincie Utrecht. (2024). *Programma Gezond en Veilig | provincie Utrecht*. <https://www.provincie-utrecht.nl/onderwerpen/gezonde-en-veilige-leefomgeving/programma-gezond-en-veilig#ggo-maptable>
- Provincie Zuid-Holland. (2024). *Samen Verder Fietsen - Provincie Zuid-Holland*. <https://www.zuid-holland.nl/onderwerpen/verkeer-vervoer/samen-verder-fietsen/>
- Rijksoverheid. (2024). *Omgevingswet | Rijksoverheid.nl*. <https://www.rijksoverheid.nl/onderwerpen/omgevingswet>
- RIVM. (2021). *RIVM zoekt samenwerking met Digital Twins om impact vergroening te visualiseren | Atlas Natuurlijk Kapitaal*. <https://www.atlasnatuurlijkkapitaal.nl/nieuws/rivm-zoekt-samenwerking-met-digital-twins-om-impact-vergroening-te-visualiseren>
- Runeson, P., & Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14(2), 131–164. <https://doi.org/10.1007/S10664-008-9102-8/TABLES/27>
- Spiekermann, S. (2012). *The challenges of privacy by design*. <https://doi.org/10.1145/2209249.2209263>
- Valentine, A. (2020). The Wrong Complexion For Protection, How Race Shaped America's Roadways And Cities. NPR.

DOCTORAL CONSORTIUM

ENHANCING AI ADOPTION IN EUROPEAN SMEs: A REFERENCE MODEL FOR SECURE AND ETHICAL INTEGRATION

MARIUS GOEBEL

Open University, Department of Information Science, Heerlen, the Netherlands
marius.goebel@ou.nl

The rapid advancement of artificial intelligence (AI) offers significant benefits but also poses challenges, particularly for European Small and Medium Enterprises (SMEs). The aim of this dissertation project is to develop a reference model combining simplified Enterprise Architecture (EA) and Value-Sensitive Design (VSD) to facilitate secure and ethical AI adoption in European SMEs. AI systems can enhance efficiency and reduce costs but also introduce risks such as legal issues and reputational harm. SMEs, crucial to the European economy, face barriers like limited resources and regulatory concerns. Design Science Research (DSR) shall be employed to develop a reference model tailored to SMEs' needs, addressing technical, ethical, and social considerations. By addressing these challenges, the research aims to foster innovation and enhance SMEs' competitiveness, promoting broader AI adoption across the EU.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.53](https://doi.org/10.18690/um.fov.4.2025.53)

ISBN

978-961-286-998-4

Keywords:

AI adoption,
SMEs,
enterprise architecture,
value-sensitive design,
ethical AI,
design science research



University of Maribor Press

1 Problem Definition and Motivation

Rapid advancements in artificial intelligence (AI) have led to its widespread adoption across various sectors, with the potential to revolutionize industries and transform the way we live and work. However, the integration of AI technologies also brings forth potential risks and challenges that need to be effectively managed. (Cousido-Gonzalez & Palacios-Alonso, 2022, Wirtz et al., 2020, Wilner, 2018) The field of AI not only concerns the subject area of IT, but also a broad field of social and ethical subject complexes. (Jones, 2020) In this context, Butterworth argues that as AI applications become increasingly complex, unknown ‘ethical surprises’ are to be expected. This requires a more comprehensive consideration of possible ethical and social implications. (Butterworth, 2018)

AI systems execute complex financial transactions, flag potential terrorists using facial recognition software, and perform document reviews. (Scherer, 2015) Thus it is possible that AI decision-making systems can cause physical damage, legal risks, regulatory risks, reputational risks or operational and financial losses for organizations that are applying them. (Chiu & Lim,) As an example, an AI system that supports medical professionals may provide inaccurate diagnoses or cause a violation of privacy. (Mannes, 2020)

According to Gartner’s trend report, AI is one of the top technology trends that is expected to have a significant impact on businesses and society in the coming years. (Gartner, 2023) However, despite the potential benefits of AI in efficiency gains or cost reductions, many organizations, particularly Small and Medium Enterprises (SMEs), struggle in adopting AI and achieving value. (Akoh, 2024, De Bellefonds et al., 2024, Agrawal & Vidyapeeth, 2024). In this work, the focus shall be set to European SMEs. SMEs are of particular interest as they form the backbone of the European economy, accounting for 67% of employment within the EU. (Robu, n. d.) Further actual AI adoption rates among European organizations indicate that they are still facing various challenges: Denmark 15.2 %, Finland 15.2 %, Luxembourg 14.4 %, Belgium 13.8 %, Netherlands 13.4 %, EU average 8 % (Eurostat, 2021). One of the main problems with AI adoption for EU SMEs is the lack of resources (money, personnel, time ...), expertise, and knowledge to implement and integrate AI technologies. SMEs often have limited budgets and lack the technical expertise to implement AI solutions, which can be complex and require

significant investment. Moreover, SMEs are hesitating to adopt AI due to concerns about regulatory and compliance, data privacy, and ethical issues. (Watney & Auer, 2021, Ulrich & Frank, 2021, Javier De Vicente Mohino et al., 2021)

As an IT security consultant working with SMEs, I have identified several recurring patterns and challenges that impact the ability of SMEs to adopt advanced technologies such as AI securely and effectively. These observations are: Inadequate documentation of IT systems, inadequate documentation of processes, resource constraints, and reactive rather than proactive approaches, in other words: lack of adherence to principles of Enterprise Architecture.

Several nations and organizations have already made initial efforts to support the adoption of AI by addressing potential threats and unethical use through the development of standardization and regulatory frameworks within various specifications. (Butcher & Beridze, 2019, Schmitt, 2022, Lorenz, 2020, Djeflal et al., 2022) This also applies to the European Union with its AI Act. The EU emphasizes ethical and transparent AI, (European Commission, 2018) potentially helping to build trust and confidence in this technology, which is critical for its widespread adoption. However, the ongoing efforts in regulations and guiding principles that companies must adhere to when adopting AI technology seems to increase the barriers to the adoption of AI technology for SMEs in the EU. (Malmborg & Trondal, 2021)

The EU is a fragmented market, with diverse languages, cultures, and economic traditions. (Spence, 2000) While this diversity can also be a challenge for AI adoption, European states also have many similarities in common, which could be an advantage in technology adoption. The regulatory frameworks most of the countries of Europe are increasingly governed by the EU. The EU has a common degree of intellectual and cultural heritage. (Spence & Perrini, 2009) This heritage might be what makes ethics so important to European policy makers. However the implementation of social and ethical responsibility by European SMEs is still a heterogeneous landscape. (Murillo & Lozano, 2006)

AI is often discussed from a regulatory or ethical standpoint, but few works target Enterprise Architecture. (Schneider et al., 2022) Scharre, Horowitz et al. recognize the potential for the leading AI provider to become the ruler of the world. (Scharre

et al., 2018) In order to accelerate European innovation and promote the rapid adoption of AI technology in Europe, an interest arises in finding out, if the AI adoption of European SMEs can be increased by providing SMEs with the tools, expertise, and knowledge to implement and integrate AI technologies.

Based on my experience working with SMEs, I postulate that SMEs need to adopt simplified Enterprise Architecture (EA) principles, combined with Value-Sensitive Design (VSD), to document IT systems and processes effectively, enabling the secure and ethical integration of AI into their operations.

2 Approaching the Problem

In order to approach the problem, the following provides an initial insight into methodologies and frameworks that could represent possible solution components. This work is to be deepened in the course of the dissertation by means of a scientific literature review in order to obtain as comprehensive an overview as possible. The aim here is to form a theoretical basis that can provide a comprehensive framework for understanding and overcoming the many challenges of introducing AI in SMEs.

In their studies ‘Emotional responses to human values in technology’ Görnemann and Spiekermann point out, that ‘the goal of technology development should be to precisely identify the direct and indirect impacts in different areas of users’ lives in order to build technology that is – in the ethical sense of the word – good for humans’ (Görnemann & Spiekermann, 2022). Hofstede emphasizes the importance of prioritizing the social aspects of AI, advocating for the development of socially aware algorithms prior to the creation of intelligent ones. (Hofstede, 2019) This perspective highlights the significance of considering the social context and human factors in the design and implementation of AI technologies. This leads us to the Value Sensitive Design approach, which incorporates ethical considerations into the development of systems and ensures that the technology is in line with the moral and social values of its users. (Friedman et al., 2013) Value Sensitive Design could be crucial for AI integration in this context, as it offers a dynamic and adaptive approach that ensures the continuous alignment of AI technologies with the human workforce and the organization’s needs.

Further the integration of AI technologies requires careful consideration of both technical and strategic aspects. An examination of Enterprise Architecture (EA) frameworks is essential to bridge the gap between IT execution and business strategy. (Lankhorst, 2009) EA helps organizations align their business goals, processes, data, applications and technology to operate efficiently and achieve their strategic objectives. (Kotusev, 2016) This alignment is essential for maximizing operational efficiency, fostering innovation, and maintaining competitive advantage. (Luftman & Kempaiah, 2007) EA acts as a strategic enabler for AI adoption by providing the necessary structure, governance, and security. (Saat et al., 2010) Nevertheless SMEs are facing issues in the implementation of EA frameworks due to their limited resources, simpler structures, and dynamic needs. (Bernaert et al., 2014)

To integrate AI into European SMEs operations, they need a tool set, specifically designed to meet their unique requirements, incorporating best practices derived from policy-maker guidelines, individual risk tolerance, and cultural elements of both the human workforce and the organization as a whole. This set of tools must meticulously assess each SME individually to align with regulatory standards, organizational culture, governance, risk management, and the seamless integration of human and AI capabilities. AI applications should adopt to SME's simple Business-IT alignment needs, and fit into their day to day activities.

The objective of this doctoral research is to create an artifact that represents a new reference model for AI integration within the enterprise architecture, tailored to the specific needs of European SMEs seeking to integrate AI into their operations. By considering their specific requirements, objectives, and the context of their AI adoption aspirations, SMEs can leverage the reference model to navigate the complex landscape of AI implementation. the reference model meticulously assesses each input to align with regulatory standards, organizational culture, governance, risk management, and the seamless integration of human and AI capabilities. The reference model guides SMEs through the technical, ethical and social considerations of deploying AI solutions and also ensures that the proposed strategies are practical, regulation-compliant, and tailored to enhance the SME's operational efficiency, innovation, and competitive edge. Success factors for a reference model, especially in the context of SMEs could be: Alignment with SME

needs; Ease of adoption; Scalability; Flexibility; Interdisciplinary approach; Cost-Effectiveness; Guidance on Change Management.

3 Research Objective

By researching the adoption of AI we can gain a deeper understanding about A) problems related to the governance and design of socio-technical systems (Bauer & Herder, 2009) and B) the requirements to information system development in AI integration projects (Benbya & McKelvey, 2006). Also, AI systems, particularly those that learn and improve through experience, can be seen as complex adaptive systems themselves.

The adoption of AI technologies of SMEs in Europe is an important and timely research topic. SMEs are the backbone of the European economy, accounting for over 99,8 % of all businesses in the EU and providing 64,4 % of total employment. (European Commission, 2023) However, SMEs often face significant challenges in adopting new technologies, including a lack of resources (money, personnel, time ...), expertise, and knowledge. (Ulrich & Frank, 2021, Javier De Vicente Mohino et al., 2021)

By focusing research on AI integration for SMEs in Europe, we can help to address these challenges and unlock the potential benefits of AI for this critical sector. AI has the potential to significantly improve the efficiency and competitiveness of SMEs, enabling them to automate routine tasks, make better decisions, and develop new products and services. The basis of the research inquiry lies in determining the feasibility of developing an AI adoption reference model tailored to European SMEs, with the primary goal of reducing associated challenges and risks taking into account complexity and ethics. The overall vision is to enable AI adoption and consequently foster innovation throughout the European Union.

The main research question to be answered is: To what extent can small and medium-sized enterprises effectively adopt AI-driven innovation while ensuring legal compliance and ethical responsibility?

Based on the main research questions the following sub research question need to be answered:

1. What are the key challenges European SMEs face in adopting AI-driven innovation while maintaining legal compliance and ethical responsibility?
2. What specific requirements must an artifact for AI adoption meet to address SMEs' constraints while ensuring compliance and ethical practices?
3. What components and structures should the artifact include to guide SMEs in securely and responsibly adopting AI technologies and how can the artifact be integrated into an SME's Enterprise Architecture?
4. How can the artifact be applied in real world SME scenarios to validate its effectiveness and practicality?
5. How effectively does the artifact address the identified challenges, and what metrics can be used to assess its success?

4 Methodology

The application of Design Science Research (DSR) methodology is planned for this dissertation. DSR is a systematic and iterative approach to problem-solving that focuses on creating and evaluating innovative solutions to complex challenges. (Chatterjee et al., 2023) DSR is ideal for research that aims to develop new frameworks, models, methods, or tools (artifacts) to address specific challenges. DSR combines rigorous research with practical application. DSR bridges the gap between theory and practice by grounding solutions in rigorous research while ensuring they are actionable in practice. (Weigand & Johannesson, 2023) Thus DSR empowers researchers and practitioners to address real-world problems. (Hevner et al., 2004) Following the DSR principles according to Hevner, the subsequent five studies are planned for achieving the goal of the dissertation:

1. Problem understanding – What are the key challenges European SMEs face in adopting AI-driven innovation while maintaining legal compliance and ethical responsibility?

Scientific Paper 1

Title: Barriers to AI Adoption in European SMEs: A Study on Ethical and Legal Challenges

Focus: This paper identifies and categorizes the main obstacles to AI adoption in SMEs, including limited resources, compliance issues, and ethical considerations.

2. Define Artifact – What specific requirements must an artifact for AI adoption meet to address SMEs' constraints while ensuring compliance and ethical practices?

Scientific Paper 2

Title: Defining Reference Model Requirements for Ethical and Compliant AI Adoption in SMEs

Focus: This paper outlines the objectives and design principles for a solution, considering the constraints and priorities of SMEs.

3. Design and Development – What components and structure should the artifact include to guide SMEs in securely and responsibly adopting AI technologies?

Scientific Paper 3

Title: A Tailored Reference Model for AI Integration in SMEs: Addressing Enterprise Architecture and Ethics

Focus: This paper presents the design of an Enterprise Architecture integrated reference model that aligns technological capabilities with ethical and legal requirements for SMEs.

4. Demonstration – How can the proposed artifact be applied in real world SME scenarios to validate its effectiveness and practicality?

Scientific Paper 4

Title: Application of a Reference Model for AI adoption: Case Studies in SMEs

Focus: This paper documents the implementation of the reference model, illustrating its application and practical impact.

5. Evaluation of the Artifact – How effectively does the artifact address the identified challenges, and what metrics can be used to assess its success?

Scientific Paper 5

Title: Evaluating the Impact of an AI Adoption Reference Model for SMEs: A Metrics-Based Approach

Focus: This paper evaluates the reference model using qualitative and quantitative metrics, assessing its ability to improve AI adoption while ensuring compliance and ethics.

Acknowledgements

The author gratefully acknowledges feedback and insightful comments provided by Prof. Dr. Ir. Johan Versendaal and Dr. Ir. Roger W. H. Bons.

References

- Agrawal, P. and Vidyapeeth, B. (2024). Adoption Of Artificial Intelligent Technologies In SMEs Sector. Article in Journal of Mountain Research.
- Akoh, E. (2024). Adoption of artificial intelligence for manufacturing SMEs 'growth and survival in South Africa: A systematic literature review. *International Journal of Research in Business and Social Science* (2147- 4478), 13(6):23–37.
- Bauer, J. M. and Herder, P. M. (2009). Designing Socio-Technical Systems. *Philosophy of Technology and Engineering Sciences*, pages 601–630.
- Benbya, H. and McKelvey, B. (2006). Toward a complexity theory of information systems development. *Information Technology and People*, 19(1):12–34.
- Bernaert, M., Poels, G., Snoeck, M., and De Backer, M. (2014). Enterprise Architecture for Small and Medium-Sized Enterprises: A Starting Point for Bringing EA to SMEs, Based on Adoption Models. pages 67–96.
- Butcher, J. and Beridze, I. (2019). What is the State of Artificial Intelligence Governance Globally? *RUSI Journal*, 164(5-6):88–96.
- Butterworth, M. (2018). The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law & Security Review*, 34(2):257–268.

- Chatterjee, S., Sundarraj, R. P., and Dutta, K. (2023). Design Science Research in Information Systems and Technology. *IEEE Explore*, 70(3).
- Chiu, I. H.-Y. and Lim, E. W. Managing Corporations' Risk in Adopting Artificial Intelligence: A Corporate Responsibility Paradigm.
- Cousido-Gonzalez, M. and Palacios-Alonso, D. (2022). Artificial Intelligence serving National Security: the Spanish case. In 2022 11th Mediterranean Conference on Embedded Computing, MECO 2022. Institute of Electrical and Electronics Engineers Inc.
- De Bellefonds, N., Charanya, T., Franke, M. R., Apotheker, J., Forth, P., Grebe, M., Luther, A., De Laubier, R., Lukic, V., Martin, M., Nopp, C., and Sassine, J. (2024). Where's the Value in AI?
- Djeflal, C., Siewert, M. B., and Würster, S. (2022). Role of the state and responsibility in governing artificial intelligence: a comparative analysis of AI strategies. *Journal of European Public Policy*, 29(11):1799–1821.
- European Commission (2018). Artificial Intelligence for Europe.
- European Commission (2023). Annual Report on European SMEs 2022/2023. Technical report, Joint Research Centre, Brussels.
- Eurostat (2021). Use of artificial intelligence in enterprises.
- Friedman, B., Kahn, P. H., Borning, A., and Hultgren, A. (2013). Value Sensitive Design and Information Systems. In *Philosophy of Engineering and Technology*, volume 16, pages 55–95. Springer Nature.
- Gartner (2023). Die wichtigsten strategischen Technologie-Trends für 2024. Technical report, Gartner.
- Görnemann, E. and Spiekermann, S. (2022). Emotional responses to human values in technology: The case of conversational agents. *Human-Computer Interaction*.
- Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). DESIGN SCIENCE IN INFORMATION SYSTEMS RESEARCH 1. Technical Report 1.
- Hofstede, G. J. (2019). GRASP agents: social first, intelligent later. *AI and Society*, 34(3):535–543.
- Javier De Vicente Mohino, J. J., Mallouli, W., Francisco Ruiz, J. F., and Van Haastrecht, M. (2021). GEIGER: Solution for small businesses to protect themselves against cyber-threats. In *ACM International Conference Proceeding Series*. Association for Computing Machinery.
- Jones, M. (2020). Artificial Intelligence – The Need for New Safety Standards and Methodologies. *Journal of System Safety*, 55(3):11–22.
- Kotusev, S. (2016). The History of Enterprise Architecture: An Evidence-Based Review. *Journal of Enterprise Architecture*, 12(1):29.
- Lankhorst, M. (2009). Enterprise Architecture at Work.
- Lorenz, P. (2020). AI Governance through Political Fora and Standards Developing Organizations Mapping the actors relevant to AI governance. Technical report, Stiftung neue Verantwortung, Berlin.
- Luftman, J. N. and Kempaiah, R. M. (2007). An Update on Business-IT Alignment. *MIS Q. Executive*.
- Mannes, A. (2020). Governance, Risk, and Artificial Intelligence. *AI Magazine*, 41(1):61–69.
- Malmborg, F. and Trondal, J. (2021). Discursive framing and organizational venues: mechanisms of artificial intelligence policy adoption. *International Review of Administrative Sciences*.
- Murillo, D. and Lozano, J. M. (2006). SMEs and CSR: An approach to CSR in their own words. *Journal of Business Ethics*, 67(3):227–240.
- Robu, M. (n. D.), S. THE DYNAMIC AND IMPORTANCE OF SMES IN ECONOMY, [http://www.ifc.org/ifcext/globalfm.nsf/AttachmentsByTitle/MSME-Country-Indicators-Analysis-Note/\\$FILE/MSME-](http://www.ifc.org/ifcext/globalfm.nsf/AttachmentsByTitle/MSME-Country-Indicators-Analysis-Note/$FILE/MSME-)
- Saat, J., Franke, U., Lagerström, R., and Ekstedt, M. (2010). Enterprise Architecture meta models for IT/business alignment situations. *Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC*, pages 14–23.

- Scharre, P., Horowitz, M. C., Allen, G. C., Frederick, K., Cho, A., Saravalle, E., and Kania, E. (2018). What Every Policymaker Needs to Know. Technical report, Center for a New American Security, Washington.
- Scherer, M. U. (2015). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. SSRN Electronic Journal.
- Schmitt, L. (2022). Mapping global AI governance: a nascent regime in a fragmented landscape. *AI and Ethics*, 2(2):303–314.
- Schneider, J., Abraham, R., Meske, C., and Vom Brocke, J. (2022). Artificial Intelligence Governance For Businesses. *Information Systems Management*.
- Spence, L. (2000). Teaching business ethics: are there differences within Europe, and is there a European difference? *Business Ethics: A European Review*, 9(1):58–64.
- Spence, L. J. and Perrini, F. (2009). Practice and politics: Ethics and social responsibility in SMEs in the European Union. *African Journal of Business Ethics*, 4(2):20–31.
- Ulrich, P. and Frank, V. (2021). Relevance and Adoption of AI technologies in German SMEs – Results from Survey-Based Research. *Procedia Computer Science*, 192:2152–2159.
- Watney, C. and Auer, D. (2021). Encouraging AI adoption by EU SMEs. Technical report, Progressive Policy Institute.
- Weigand, H. and Johannesson, P. (2023). How to Identify your Design Science Research Artifact. *IEEE Xplore*
- Wilner, A. S. (2018). Cybersecurity and its discontents: Artificial intelligence, the in-ternet of things, and digital misinformation. *International Journal*, 73(2):308–316.
- Wirtz, B. W., Weyerer, J. C., and Sturm, B. J. (2020). The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration. *International Journal of Public Administration*, 43(9):818–829.

DOCTORAL CONSORTIUM

THE JOURNEY TO SUSTAINABLE DIGITAL BUSINESS MODELS: EMBEDDING TWIN TRANSFORMATION FOR LONG- TERM SURVIVAL

TARA VESTER

Utrecht University of Applied Sciences, Utrecht, Netherland
tara.vester@hu.nl

This PhD explores how organizations overcome resource constraints and develop dynamic capabilities within the Twin Transformation (TT) to move toward sustainable digital business models. TT—the convergence of digital and sustainability transformations—requires organizations to reshape their business models to stay resilient in an evolving landscape. Yet, as both TT and sustainable digital business models are still emerging concepts, the implications and necessary organizational adaptations remain unclear. The complexity of the TT, with its dilemmas and trade-offs, adds further challenges. This research investigates how internal and external dynamics influence TT adoption, and how organizations orchestrate resources and build capabilities to overcome barriers. Using a qualitative, process-oriented approach findings are expected to show that the TT must be approached as a fundamental organizational change, with similar constraints and capability challenges emerging across organizations. This study contributes to the discourse on TT, organizational change, and sustainable digital business model discourse, offering practical guidance to strengthen organizational resilience.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.54](https://doi.org/10.18690/um.fov.4.2025.54)

ISBN

978-961-286-998-4

Keywords:

sustainable digital business
model,
twin transformation,
purpose,
resource orchestration,
dynamic capabilities,
MLEs



University of Maribor Press

1 Introduction

Two major transformative forces are influencing how organizations adapt their business models to stay future proof: the sustainability transformation and digital transformation. The sustainable transformation necessitates a shift towards a business model based on multiple-value-co-creation, integrating ecological, social, and governance dimensions into business models that have traditionally prioritized economic value (Barnes et al., 2024; Breiter et al., 2024), thereby refining their *raison d'être*. Those that fail to adapt may struggle to comply with regulations such as the Corporate Sustainability Reporting Directive (CSRD) exposing themselves to regulatory penalties (Novicka & Volkova, 2024), lessened employee attractiveness (N. M. P. Bocken & Geradts, 2020) and potential loss of market trust (George & Schillebeeckx, 2022). Simultaneously, the digital transformation compels organizations to integrate transformative technologies—such as platforms, Artificial Intelligence, blockchain, and Internet of Things—into their business model, fundamentally redefining how they operate and innovate (Breiter et al., 2024). Organizations that fail to embrace digital transformation risk operational inefficiencies (Martínez-Peláez et al., 2024) and diminished market resilience (Graf-Drasch et al., 2023), making it increasingly difficult to sustain long-term success as their competitors *will* embrace it. These two transformations are not occurring in isolation, but are converging, leading to what is known as the Twin Transformation (TT) (Graf-Drasch et al., 2023; Kürpick et al., 2024). According to Christmann et al. (2024, p. 7) the TT represents “a value-adding interplay... leveraging digital technologies for enabling sustainability and leveraging sustainability for guiding digital progress.”, as shown in Figure 1. However, while the TT offers the potential to harness the synergies of both transformations for sustainable digital value creation, it is not a straightforward process in which their combined benefits automatically add up. Instead, the TT introduces new dilemmas and trade-offs that organizations must navigate (Raihan, 2024). Existing literature highlights two challenges that organizations might encounter in the TT process: sustainability traps and digitalization traps. In sustainability traps, innovations support sustainability goals but lack economic feasibility, making them financially unsustainable (Xu et al., 2024). Notably, this type of trap does not (yet) explicitly incorporate the role of digitalization. Conversely, in digitalization traps technological advancements increase negative environmental impact (N. Bocken, 2023; Rosati et al., 2024), such as the high energy consumption in AI-driven processes. These challenges

underscore the complexity of TT, requiring organizations to move beyond isolated sustainability or digital strategies towards a more integrated and holistic approach to TT. The need for a holistic transformation aligns with Christmann et al.'s (2024, p.5) definition of TT, which I will follow in this PhD: “a fundamental organisational change process that enables organizations to address digital and societal challenges synergistically by harnessing the power of digital transformation to enable sustainable transformation and leveraging sustainable transformation to redesign digital transformation”. Because the TT is recognized as a broader geopolitical, social, economic, and regulatory shift (European Commission, 2022), it requires more than superficial adjustments. Organizations need to fundamentally rethink their structures, strategies, and capabilities to fully align with TT imperatives and position themselves for long-term survival (Plotnytska et al., 2024; Riso & Morrone, 2023). So, by actively integrating TT into their business model, organizations could strengthen their competitiveness (van Erp & Rytter, 2023), and could potentially avoid risks that could ultimately weaken their resilience, with potential consequences such as financial instability or even bankruptcy. In other words, organizations have both a strategic imperative and a broader responsibility to effectively navigate TT—not only to secure their own survival but also to contribute to a more sustainable world.

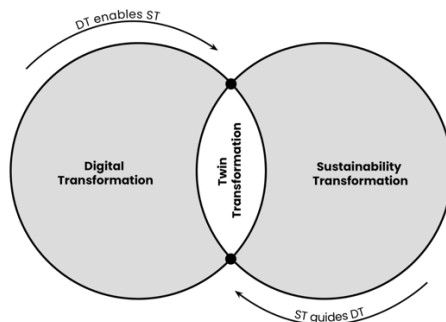


Figure 1: The Twin Transformation dynamic

Source: (Christmann et al., 2024, p. 7)

Before we can start to understand how organizations navigate the challenges and opportunities of TT, it is essential to establish a clear framework that defines what a business model incorporating the TT entails. Without such a foundation, discussions on the TT risk remaining abstract, making it difficult to assess progress, identify

strategic pathways, and develop actionable insights for organizations undergoing this transformation. Within the TT, the objective is to foster the development of Sustainable Digital Business Models (SD-BMs)—business models that integrate sustainability and digitalization to create, deliver, and capture value in a way that is both environmentally responsible and technologically advanced. This raises the question: how do ‘sustainable digital’ aspects enhance or redefine existing business model constructs? First, what exactly is an SD-BM? Currently, no clear definition exists, although Böttcher et al. (2024) attempt to conceptualize the term within the TT literature by focusing on archetypes, and, along with Xu et al. (2024), call for more research on this topic. Existing frameworks for sustainable business models and digital business models offer relevant foundations to build upon. For example, sustainable business models are defined by Geissdoerfer et al. (2018, pp. 403-404) as models that “incorporate pro-active multi-stakeholder management, the creation of monetary and non-monetary value for a broad range of stakeholders, and hold a long-term perspective”. In turn, Bican and Brem (2020, p. 10) offer a framework on how digital business models might sustainably relate to innovation, describing how they “enhance resource optimization, characterized by intangibility, businesses’ uniqueness, and core values, centering around experience, platform and content”. Encouragingly, organizations are already beginning to rethink sustainable digital value by leveraging digital technologies, such as exchange platforms, to e.g. facilitate resource sharing (Aagaard & Vanhaverbeeke, 2024). Additionally, Green IT initiatives demonstrate how digital transformation can actively support sustainability goals by reducing environmental impacts throughout the ICT value chain (Fors et al., 2024). Thus, although the concept of SD-BMs is still in need of clearer frameworks, the existing theoretical foundations combined with practical examples of business models that integrate both transformations, offer valuable starting points. These insights make it possible to establish a discursive foundation for further exploring and refining what an SD-BM is and how it can guide organizations in practice.

To navigate towards a SD-BM it is important to realize that organizations do no longer modify their business models in a vacuum, since “addressing complex sustainability challenges with uncertain payoffs may require financial trade-offs and lengthy experimentation with a broad range of external stakeholders (N. M. P. Bocken & Geradts, 2020)”. In more abstract terms, business models are shaped by, and in turn, shape the specific segment of its external environment that encompasses

the opportunities and expectations relevant to its operations, referred to as the existence-relevant space (Rüegg-Stürm & Grand, 2019), visualized in Figure 2. As the TT is a broader geopolitical, social, economic, and regulatory shift (European Commission, 2022), the TT influences this space one way or another. Because organizations not only adapt to but also actively shape this space, the TT therefore influences both the external environment and the organizations' internal dimensions, such as governance, strategy, structure, and culture. Within the existence-relevant space organizations must continuously balance the continuum of 'renewal' and 'optimization' of their business model (Rüegg-Stürm & Grand, 2019) to sustain organizational value creation over time. Excessive renewal can lead to instability, while over-reliance on optimization risks stagnation. TT adds complexity by requiring organizations to manage both digital and sustainability transformations while navigating this balance. The continuous interaction between complex internal and external dynamics, combined with the increasing need for business models to provide 'complex value' to different stakeholders (Barnes et al., 2024, p. 3), creates additional challenges in effectively transforming toward a business model that ensures long-term survival. This brings us to a key question: where do organizations begin their TT journey within the complex interaction with their existence-relevant space, and how does this initial orientation shape the specific resources and capabilities required to transition towards a SD-BM?

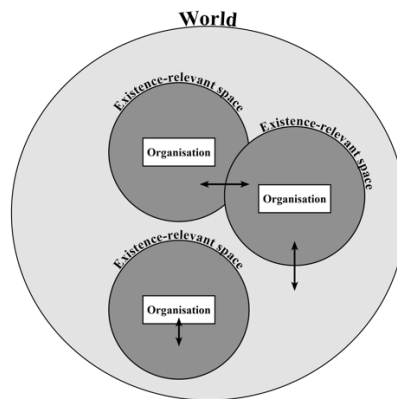


Figure 2: Own Interpretation of Organizations in Relation to the Existence-relevant Space
Based on Rüegg-Stürm & Grand, 2019

How organizations start their journey towards SD-BMs is important, as it influences whether the TT is embedded in the organization or remains a superficial exercise. This is important as it determines how organizations will further allocate resources and develop dynamic capabilities to strategically adapt their business models to align with both digitalization and sustainability imperatives. While some scholars have attempted to conceptualize TT pathways (Aagaard & Vanhaverbeeke, 2024; Böttcher et al., 2024), empirical insights remain scarce (Christmann et al., 2024; Jonkers & Vester, 2024). Aagaard and Vanhaverbeeke (2024) propose a quadrant model that maps organizations based on their degree of sustainability integration (high or low) and digital intensity (high or low), outlining potential pathways they may take. Breiter et al. (2024) adopt a maturity model approach, identifying three pathways leading to the “True Twin Transformer”. Although these models contribute to the discourse, they do not address a fundamental question: what determines whether TT is adopted as a fundamental entrepreneurial opportunity or merely a marketing task (cf. Rüegg-Stürm & Grand, 2019)? Purpose literature (George et al., 2023) can help analyze the early-stage dynamics of TT adoption to clarify how initial strategic decision-making practices shape resource allocation. Additionally, organizational culture plays a pivotal role in the digital transformation towards sustainable development (Philbin et al., 2022), supported by Martínez-Peláez et al. (2024) who write that “success [in the TT] depends on creating adaptive business models and fostering a culture that embraces change, innovation and dynamic capabilities”. However, because the TT literature lacks knowledge on organizational change management (Pacoli, 2022), it leaves a gap in understanding how culture can support TT integration. Therefore, understanding how internal and external influences shape TT adoption is crucial, as these factors directly impact how organizations manage resource constraints and navigate the transition toward SD-BMs.

Once the TT adoption in the context of TT has been examined, a critical gap remains in understanding the specific resources and capabilities required to transform towards a SD-BM (Christmann et al., 2024; Feroz et al., 2023). While some research is emerging in the DT literature (Chen & Tian, 2022; Peretz-Andersson et al., 2024) and circular economy literature (Kristoffersen et al., 2021), the intersection of these two domains has not been sufficiently examined. Resources, defined as combinations of assets (Barney et al., 2001), can be *physical* (e.g., plant and equipment), *human* (e.g., attributes of managers and workers), and *organizational* (e.g.,

structures and technologies) (Barney et al., 2001). To achieve long-term survival, these resources must be continuously adjusted to align with the existence-relevant space. However, the question remains: how can firms effectively manage and develop these resources within the TT context? To address this, Resource Orchestration (RO) theory (Sirmon et al., 2011) provides a framework for understanding how firms structure, bundle and leverage resources. When effectively combined, these resources form capabilities (Barney et al., 2001), as illustrated in Figure 3 (Ahuja & Chan, 2017, p. 81). Dynamic Capabilities (DC) theory extends the RO theory by explaining how organizations develop and adapt capabilities over time (Teece, 2007), critical to transform business models (N. M. P. Bocken & Geradts, 2020). The DC framework, comprising sensing, seizing, and transforming capabilities, enables organizations to respond effectively to external shifts and sustain competitive advantage (Teece et al., 1997). RO and DC therefore together provide complementary lenses for analyzing how firms adjust their resources while moving towards SD-BMs. In summary, while RO explains how firms mobilize and configure their resources, DC clarifies how these resource configurations evolve into adaptive capabilities to sustain business model resilience through time.



Figure 3: The Process View of IT Innovation

Source: (Ahuja & Chan, 2017, p. 81)

2 Problem Definition

As organizations face increasing pressure from TT, they need more empirical examples on how to navigate the intersection of digitalization and sustainability. While these two transformations have the potential to reinforce one another, they also introduce dilemmas and trade-offs, making it challenging for organizations to determine how to modify their business models to sustain long term survival. This has significant implications for how organizations manage resource constraints arising from TT and how they develop (dynamic) capabilities to effectively respond to external shifts. This study aims to identify common patterns among organizations that successfully integrate TT, with the objective of understanding how internal and

external dynamics enable organizations to navigate TT's complexities and constraints.

The main research question is: How can organizations embed the Twin Transformation in their organization for sustainable digital business model development? To answer this question the following sub-research questions are formulated.

SQ1: What constitutes a sustainable digital business model?

SQ2: How do organizations initiate their journey towards a sustainable digital business model?

SQ3: What resource constraints do organizations encounter when moving towards a sustainable digital business model and how do they cope with them?

SQ4: How do organizations develop the necessary dynamic capabilities in order to move towards a sustainable digital business model?

SQ5: What patterns can be identified in organizations embedding a sustainable digital business model?

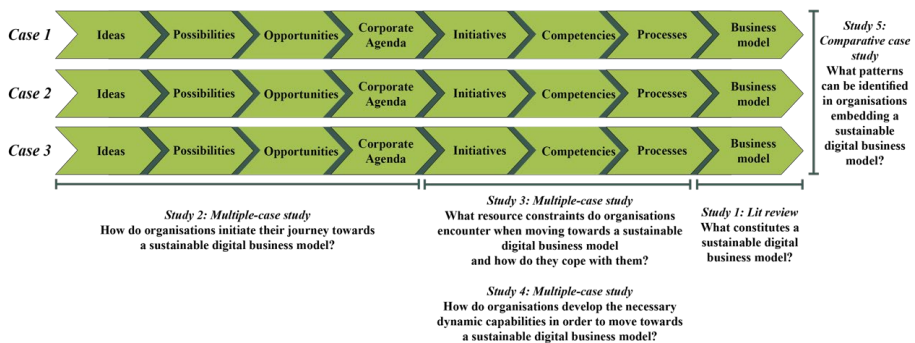


Figure 4: Visualization of Sub-question Alignment for SD-BM Development, Based on Rüegg-Stürm and Grand's (2019, p. 176) Strategy Development Framework

3 Methodology

The complexity and evolving nature of TT, coupled with its predominantly conceptual literature (Christmann et al., 2024; Jonkers & Vester, 2024), necessitate an explorative approach. This PhD adopts social constructionism, which views reality as socially constructed through shared meanings, interactions, and language (Berger & Luckmann, 1991). This perspective is grounded in a relational process ontology (Langley, 1999; Van de Ven & Poole, 2005), emphasizing that organizations continuously adapt their business models in response to changing internal and external dynamics. I therefore take a temporal approach within three organizations over the period of two years to capture both the intertwined context and more contextual specificity that shape SD-BMs (Hoorani et al., 2023). By integrating multiple forms of knowledge-exchange throughout the research period, I ensure temporal data access while simultaneously creating value for the case organizations throughout their journey. This continuous data collection will ensure sufficient and sufficiently rich data to capture in-depth, embedded, and processual insights, while allowing for theoretical grounding in the phenomenon (Van de Ven, 2016). Through ongoing reflection and practitioner engagement, this study aims to provide a sincere, credible, and meaningful account (cf. Tracy, 2010) of how organizations shape SD-BMs. I conduct the following studies: A scoping review to build common language (SQ1), three multiple-case studies, analyzing adoption, resource constraints and dynamic capabilities separately (SQ2-4) and a study to identify overarching patterns across the three case studies (SQ5).

To support this approach, I combine social constructionism with a strong process orientation (Chia, 2002), as it allows for understanding transformation as an unfolding process, examining how organizations make sense of and embed TT in practice through time. Before a new initiative like TT becomes embedded within an organization's structural foundation, several organizational processes must unfold. Understanding how this happens is the core focus of my PhD. To provide theoretical structure for designing this process, I draw on the Strategy Development Framework (Rüegg-Stürm & Grand, 2019). This framework is particularly valuable as it helps to clarify the underlying creation and decision-making processes that ultimately lead to value creation—which, in this case, translates to the development of an SD-BM. In my PhD, I aim to provide clarity on the TT embedding process by systematically studying different phases of this process. While Figure 4 illustrates

how my studies cover all phases towards business model development, the studies themselves will be conducted in a sequential order–1, 2, 3, 4, and 5. Even if the visual representation suggests a different order, my research follows a progression where each study builds upon the insights of the previous one. Ultimately, this approach will allow me to construct a more comprehensive and integrated overview of the processes underlying SD-BMs development within organizations.

This study focuses on frontrunner organizations who act as catalysts for the TT. Frontrunner organizations actively shape their existence-relevant space through collaborative and coordinative activities (Kilpatrick & Conroy, 2024), influencing the actions and strategies of other actors operating within, and extending from, this space. By examining how frontrunners create enabling conditions and overcome barriers, this study provides insights that can inspire and guide other organizations in moving towards SD-BMs. Mid-to-large-size enterprises (MLEs) provide a structured setting for examining TT adoption over time, with defined processes and decision-making structures suited for temporal and process research. Unlike smaller firms, MLEs face more complex structural constraints, requiring deliberate strategic adaptations to embed TT effectively. I follow the classification of MLEs as outlined by the Sociaal-Economische Raad (n.d.), in Table 1.

Table 1: Definitions of Mid-to-large size Companies (Sociaal-Economische Raad, n.d.)

	Mid-size	Large
Assets	No more than €25 million	More than €25 million
Net sales	No more than €50 million	More than €50 million
# employees	Less than 250	250 or more

3.1 Study 1: What constitutes a sustainable digital business model?

This study uses a scoping review, supplemented by focus groups, to establish a foundational understanding of SD-BMs within the TT. As SD-BMs lack a clear definition, this research provides an initial framework to guide the PhD. Focus groups add depth by capturing practitioners' interpretations of TT, ensuring alignment with real-world applications. Data collection includes academic literature and industry reports on business models, sustainability, and digitalization. Given the TT's limited academic discourse, industry reports offer practical insights. Focus groups with sustainability and digital transformation leads further explore how these concepts are applied in practice. For data analysis, the scoping review systematically

maps existing knowledge, while focus group insights are coded and analyzed to identify key themes and contextual applications, refining the SD-BM framework. The planning of this study can be found in Figure 5 in the appendix.

3.2 Study 2: How do organizations initiate their journey towards a sustainable digital business model?

This study employs a multiple case study based on dynamic capability lens (Teece, 2007; Teece et al., 1997) to examine early-stage TT adoption. By integrating Purpose literature (George et al., 2023) I explore how the TT is initially approached (e.g., top-down or bottom-up, fundamental entrepreneurial opportunity or marketing task (Rüegg-Stürm & Grand, 2019) to understand how this orientation later shapes resource orchestration (Study 3) and dynamic capability development (Study 4) in the following sub studies. Data collection focuses on frontrunner mid-to-large enterprises (MLEs), which have likely been engaged in TT for some time, necessitating a retrospective approach. Using the “timeline and eye-opener workshop” method (Van Mierlo et al., 2010) key stakeholders will reflect on challenges, successes, and experiences in their early TT journey. The key stakeholders involved are expected to vary across case organizations, depending on their structure, industry, and stage of TT adoption. Following this phase, interim reflection workshops will track how organizations introduce, frame, and embed TT in real time, providing longitudinal insights into their transformation process. The planning of this study can be found in Figure 6 in the appendix.

3.3 Study 3: What resource constraints do organizations encounter when moving towards a sustainable digital business model and how do they cope with them?

This multiple-case study applies the resource orchestration lens (Sirmon et al., 2007, 2011) to examine how organizations prioritize, restructure, and allocate resources while balancing the demands and trade-offs of digitalization and sustainability. Despite emerging research on digital transformation (Chen & Tian, 2022; Peretz-Andersson et al., 2024) and circular economy (Kristoffersen et al., 2021), resource orchestration at their intersection remains underexplored (Jonkers & Vester, 2024). Data collection includes semi-structured interviews and observations with key decision-makers (e.g., executives, sustainability officers, digital transformation leads). This study will

analyze the underlying processes through which organizations restructure, prioritize and deploy resources in response of the demands of the TT. Using pattern mapping, it will identify common strategies for navigating and overcoming resource constraints. The planning of this study can be found in Figure 7 in the appendix.

3.4 Study 4: How do organizations develop the necessary dynamic capabilities in order to move towards a sustainable digital business model?

This multiple-case study uses a dynamic capabilities lens (Teece, 2007; Teece et al., 1997) to explore how organizations develop the capabilities needed for the TT. While literature on dynamic capabilities in TT is emerging (Christmann et al., 2024; Feroz et al., 2023), empirical evidence on how organizations develop the necessary dynamic capabilities remains scarce (Christmann et al., 2024). Additionally, existing dynamic capability studies predominantly focus on ecological aspects, while the integration of social, economic, and governance dimensions into these strategies remains underexplored (Feroz et al., 2023). Building on SQ2 and SQ3, this study examines how an organization's Purpose and resource orchestration influence re- and upskilling strategies, which are essential as the TT necessitates re- and upskilling across the entire organization on many different levels (Jonkers & Vester, 2024). Doing so, it is essential to align with evolving competency frameworks, such as the Circular Economy Education Map (Vitti et al., 2025). Data collection includes workshops with industry, government, and education stakeholders to leverage triple-helix interactions. This aligns with ongoing initiatives like the Nationale Coalitie Duurzame Digitalisering (NCDD) and the ErasmusPlus project Digital4Sustainability. Internal document analysis of training programs and job descriptions will offer insights into learning strategies, while observational studies, where feasible, will assess engagement with skill-building initiatives like on-the-job training and mentoring. The Digital4Sustainability project's training materials will be piloted to evaluate their effectiveness. Data analysis will involve process tracking to capture the evolution of DCs and qualitative coding to identify capability-building patterns and decision-making processes in workforce development. The planning of this study can be found in Figure 8 in the appendix.

3.5 Study 5: What patterns can be identified in organizations embedding a sustainable digital business model?

This comparative case study, complemented by a qualitative survey, analyzes the three organizations from SQ2-4 to identify factors explaining similarities in TT adoption, resource allocation, and dynamic capability development. It bridges TT and organizational change literature by examining the dynamics that shape these patterns (Pacoli, 2022). Data collection includes workshops, observational studies, and a qualitative survey to validate findings across a broader set of organizations. After 2–3 years of research, emerging patterns will be analyzed through iterative feedback loops and a cross-case synthesis to distill transferable process principles across sectors and structures. The planning of this study can be found in Figure 9 in the appendix.

4 Expected Results

This research is expected to reveal common patterns among organizations that are reconfiguring their business models in response to the challenges and opportunities of the TT. In this process, organizations are anticipated to encounter resource constraints and dynamic capability challenges to facilitate this transformation. Organizations are likely to face similar types of dilemmas, encountering common barriers such as knowledge gaps, limited adoption of digital technologies, misalignment internal incentives, and insufficient cross-functional coordination. Despite variations in business models, these barriers are expected to reflect shared underlying structural challenges. By examining front-runner organizations, the study expects to identify key enablers to overcome these constraints, such as an innovation-stimulating culture, adaptive governance structures, collaborative cross-functional practices, and strategic flexibility—all of which support organizations in effectively embedding TT toward resilient SD-BMs.

This brings me to the most significant limitation of this study, which stems from the decision to examine SD-BM development as a process embedded in complex internal and external dynamics. While this perspective is essential to capture the interconnectedness of the various elements and to generate insights that are truly meaningful for practice, this approach also means that the findings are based on an in-depth study of three organizations. As a result, the outcomes are highly context-

specific and not directly transferable to other settings, possibly limiting the generalizability of the results. Moreover, the reliance on qualitative data introduces the risk of interpretation bias, despite efforts to ensure methodological rigor through triangulation and transparent data analysis procedures. To mitigate these limitations, Study 5 incorporates a qualitative survey to validate and refine the emerging patterns across a broader set of organizations, thereby enhancing the robustness and applicability of the conclusions. Ultimately, the findings should be viewed as contextually grounded insights that can inform—but not prescribe—strategic approaches to Twin Transformation in other organizational settings.

5 Future Development

This PhD offers a starting point for understanding how organizations develop SD-BMs while overcoming resource constraints and dynamic capability limitations. As the TT continues to evolve, future research can expand across sectors and organizational types, exploring industry-specific approaches to SD-BM development. A key area for further study is the role of digital technologies, such as AI-powered expectation models, which could generate predictive insights to guide SD-BM design. As organizations become more interdependent on their external environment, future research should also examine the role of supply chains and network collaborations in fostering collaborative SD-BMs, driving systemic change in the existence-relevant space. Finally, combining qualitative and quantitative methods could enhance generalizability and validation, offering a more robust understanding of how TT is embedded across different contexts.

References

- Aagaard, A., & Vanhaverbeeke, W. (2024). *Business Model Innovation* (A. Aagaard, Ed.). Springer International Publishing. <https://doi.org/10.1007/978-3-031-57511-2>
- Ahuja, S., & Chan, Y. E. (2017). Resource Orchestration for IT-enabled Innovation. In *Kindai Management Review* (Vol. 5). https://www.researchgate.net/publication/328811215_Resource_Orchestration_for_IT-enabled_Innovation
- Barnes, J., Hansen, P., Kamin, T., Golob, U., Darby, S., van der Grijp, N. M., & Petrovics, D. (2024). Creating valuable outcomes: An exploration of value creation pathways in the business models of energy communities. *Energy Research & Social Science*, 108, 103398. <https://doi.org/10.1016/j.erss.2023.103398>
- Barney, J., Wright, M., & Ketchen, D. J. (2001). The resource-based view of the firm: Ten years after 1991. *Journal of Management*, 27(6), 625–641. <https://doi.org/10.1177/014920630102700601>

- Berger, P. L., & Luckmann, T. (1991). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Penguin Books. <https://amstudugm.wordpress.com/wp-content/uploads/2011/04/social-construction-of-reality.pdf>
- Bican, P. M., & Brem, A. (2020). Digital Business Model, Digital Transformation, Digital Entrepreneurship: Is There A Sustainable “Digital”? *Sustainability*, 12(13), 5239. <https://doi.org/10.3390/su12135239>
- Bocken, N. (2023). Business Models for Sustainability. In *Oxford Research Encyclopedia of Environmental Science*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199389414.013.842>
- Bocken, N. M. P., & Geradts, T. H. J. (2020). Barriers and drivers to sustainable business model innovation: Organization design and dynamic capabilities. *Long Range Planning*, 53(4), 101950. <https://doi.org/10.1016/j.lrp.2019.101950>
- Böttcher, T. P., Empelmann, S., Weking, J., Hein, A., & Krcmar, H. (2024). Digital sustainable business models: Using digital technology to integrate ecological sustainability into the core of business models. *Information Systems Journal*, 34(3), 736–761. <https://doi.org/10.1111/isj.12436>
- Breiter, K., Crome, C., Oberländer, A. M., & Schnaak, F. (2024). Dynamic Capabilities for the Twin Transformation Climb: A Capability Maturity Model. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-024-10520-y>
- Chen, H., & Tian, Z. (2022). Environmental uncertainty, resource orchestration and digital transformation: A fuzzy-set QCA approach. *Journal of Business Research*, 139, 184–193. <https://doi.org/10.1016/j.jbusres.2021.09.048>
- 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>
- European Commission. (2022). *Twinning the green and digital transitions in the new geopolitical context*. <https://publications.jrc.ec.europa.eu/repository/handle/JRC129319>.
- Feroz, A. K., Zo, H., Eom, J., & Chiravuri, A. (2023). Identifying organizations’ dynamic capabilities for sustainable digital transformation: A mixed methods study. *Technology in Society*, 73, 102257. <https://doi.org/10.1016/j.techsoc.2023.102257>
- Fors, P., Kreps, D., & O’Brien, A. (2024). *Green IT: The Evolution of Environmental Concerns Within ICT Policy, Research and Practice* (pp. 25–48). https://doi.org/10.1007/978-3-031-61749-2_2
- Geissdoerfer, M., Vladimirova, D., & Evans, S. (2018). Sustainable business model innovation: A review. *Journal of Cleaner Production*, 198, 401–416. <https://doi.org/10.1016/j.jclepro.2018.06.240>
- George, G., Haas, M. R., McGahan, A. M., Schillebeeckx, S. J. D., & Tracey, P. (2023). Purpose in the For-Profit Firm: A Review and Framework for Management Research. *Journal of Management*, 49(6), 1841–1869. <https://doi.org/10.1177/01492063211006450>
- George, G., & Schillebeeckx, S. J. D. (2022). Digital transformation, sustainability, and purpose in the multinational enterprise. *Journal of World Business*, 57(3), 101326. <https://doi.org/10.1016/j.jwb.2022.101326>
- Graf-Drasch, V., Kauffeld, L., Kempf, L., Maria Oberländer, A., Teuchert, A., & Maria, A. (2023). *Driving twin transformation - the interplay of digital transformation and sustainability transformation*. https://aisel.aisnet.org/ecis2023_rp/255
- Hoorani, B. H., Plakoyiannaki, E., & Gibbert, M. (2023). Understanding time in qualitative international business research: Towards four styles of temporal theorizing. *Journal of World Business*, 58(1), 101369. <https://doi.org/10.1016/j.jwb.2022.101369>
- Jonkers, I., & Vester, C. N. W. (2024). *Academic Literature Review Digital4Sustainability*.
- Kilpatrick, C., & Conroy, K. M. (2024). Driving Digital Sustainability in Global Value Chains: Multinational Enterprises as Chief Orchestrators. *Academy of Management Perspectives*. <https://doi.org/10.5465/amp.2023.0160>

- Kristoffersen, E., Mikalef, P., Blomsma, F., & Li, J. (2021). The effects of business analytics capability on circular economy implementation, resource orchestration capability, and firm performance. *International Journal of Production Economics*, 239, 108205. <https://doi.org/10.1016/j.ijpe.2021.108205>
- Kürpick, C., Schreiner, N., Krauß-Kodytek, L., Kühn, A., Plass, S., & Scholz, T. (2024). Capabilities for the Strategic Alignment of Sustainability and Digitalization in Manufacturing: Insights from Theory and Practice. *2024 IEEE 65th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, 1–6. <https://doi.org/10.1109/ITMS64072.2024.10741924>
- Langley, A. (1999). Strategies for Theorizing from Process Data. *Academy of Management Review*, 24(4), 691–710. <https://doi.org/10.5465/amr.1999.2553248>
- Martínez-Peláez, R., Escobar, M. A., Félix, V. G., Ostos, R., Parra-Michel, J., García, V., Ochoa-Brust, A., Velarde-Alvarado, P., Félix, R. A., Olivares-Bautista, S., Flores, V., & Mena, L. J. (2024). Sustainable Digital Transformation for SMEs: A Comprehensive Framework for Informed Decision-Making. *Sustainability*, 16(11), 4447. <https://doi.org/10.3390/su16114447>
- Novicka, J., & Volkova, T. (2024). Regulation of Sustainability Reporting Requirements—Digitalisation Path. *Sustainability*, 17(1), 138. <https://doi.org/10.3390/su17010138>
- Pacolli, M. (2022). Importance of Change Management in Digital Transformation Sustainability. *IFAC-PapersOnLine*, 55(39), 276–280. <https://doi.org/10.1016/j.ifacol.2022.12.034>
- Peretz-Andersson, E., Tabares, S., Mikalef, P., & Parida, V. (2024). Artificial intelligence implementation in manufacturing SMEs: A resource orchestration approach. *International Journal of Information Management*, 77, 102781. <https://doi.org/10.1016/j.ijinfomgt.2024.102781>
- Philbin, S., Viswanathan, R., & Telukdarie, A. (2022). Understanding how digital transformation can enable SMEs to achieve sustainable development: A systematic literature review. *Small Business International Review*, 6(1), e473. <https://doi.org/10.26784/sbir.v6i1.473>
- Plotnytska, S., Chédotel, F., & Glińska-Neweś, A. (2024). *Towards sustainable digital transformation for SMEs: an agenda for future studies*. <https://www.oecd.org/cfe/smes/2090740.pdf>
- Raihan, A. (2024). A review of the potential opportunities and challenges of the digital economy for sustainability. *Innovation and Green Development*, 3(4), 100174. <https://doi.org/10.1016/j.igd.2024.100174>
- Riso, T., & Morrone, C. (2023). To Align Technological Advancement and Ethical Conduct: An Analysis of the Relationship between Digital Technologies and Sustainable Decision-Making Processes. *Sustainability (Switzerland)*, 15(3). <https://doi.org/10.3390/su15031911>
- Rosati, P., Lynn, T., Kreps, D., & Conboy, K. (2024). *Digital Sustainability: Key Definitions and Concepts* (pp. 1–24). https://doi.org/10.1007/978-3-031-61749-2_1
- Rüegg-Stürm, J., & Grand, S. (2019). *Managing in a Complex World*. utb GmbH. <https://doi.org/10.36198/9783838552996>
- 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>
- Sirmon, D. G., Hitt, M. A., Ireland, R. D., & Gilbert, B. A. (2011). Resource orchestration to create competitive advantage: Breadth, depth, and life cycle effects. In *Journal of Management* (Vol. 37, Issue 5, pp. 1390–1412). <https://doi.org/10.1177/0149206310385695>
- Sociaal-Economische Raad. (n.d.). *Wat is een grote, middelgrote, kleine en micro-onderneming?* Retrieved February 27, 2025, from <https://www.ser.nl/nl/thema/duurzaamheid/faq/groot-middel-klein>
- 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<509::AID-SMJ882>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z)

- Tracy, S. J. (2010). Qualitative quality: Eight a"big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*, 16(10), 837–851. <https://doi.org/10.1177/1077800410383121>
- Van de Ven, A. H. (2016). Grounding the research phenomenon. *Journal of Change Management*, 16(4), 265–270. <https://doi.org/10.1080/14697017.2016.1230336>
- Van de Ven, A. H., & Poole, M. S. (2005). Alternative Approaches for Studying Organizational Change. *Organization Studies*, 26(9), 1377–1404. <https://doi.org/10.1177/0170840605056907>
- van Erp, T., & Rytter, N. G. M. (2023). Design and operations framework for the Twin Transition of manufacturing systems. *Advances in Production Engineering & Management*, 18(1), 92–103. <https://doi.org/10.14743/apem2023.1.459>
- Van Mierlo, B., Regeer, B., Van Amstel, M., Arkesteijn, M., Beekman, V., Bunders, J., De Cock Buning, T., Elzen, B., Hoes, A.-C., & Leeuwis, C. (2010). *Reflexieve monitoring in actie Handvatten voor de monitoring van systeeminnovatieprojecten* (Vol. 1).
- Vitti, M., Trevisan, A. H., Ocampo, H. R., Cuentas, V. K., Sarbazvatan, S., Terzi, S., & Sassanelli, C. (2025). A competency map for circular economy education. *Procedia Computer Science*, 253, 336–345. <https://doi.org/10.1016/j.procs.2025.01.096>
- Xu, G., Zhang, J., & Wang, S. (2024). How Digitalization and Sustainability Promote Digital Green Innovation for Industry 5.0 through Capability Reconfiguration: Strategically Oriented Insights. *Systems*, 12(9), 341. <https://doi.org/10.3390/systems12090341>

Appendix

		2025					2026	
		jul	aug	sep	oct	nov	dec	jan
RQ1	Scoping review							
	Orientation & Literature Collection							
	Reading & Thematic Analysis							
	Conducting Focus Groups							
	Writing the Scoping Review							
	Feedback & Revisions							
	Submission							

Figure 5: Planning Study 1

		2025		2026						
		nov	dec	jan	feb	mar	apr	may	jun	jul
RQ2	Multiple case study: TT adoption									
	Research Strategy & Planning									
	Retrospective Data Collection									
	Data Analysis & Writing									
	Finalization									
	Submission									

Figure 6: Planning Study 2

		2026				2027						
		sep	oct	nov	dec	jan	feb	mar	apr	may	jun	sep
RQ3	Multiple case study: resource orchestration											
	Research Strategy & Planning											
	Data Collection – Interviews & Document Analysis											
	Data Collection – Process Mapping & Workshops											
	Data Analysis & Writing											
	Finalization											
	Submission											

Figure 7: Planning Study 3

		2027				2028						
		sep	oct	nov	dec	jan	feb	mar	apr	may	jun	aug
RQ4	Multiple case study: dynamic capabilities											
	Research Strategy & Planning											
	Data Collection – Interviews & Document Analysis											
	Data Collection – Observational Studies & Workshops											
	Data Analysis & Writing											
	Finalization											
	Submission											

Figure 8: Planning Study 4

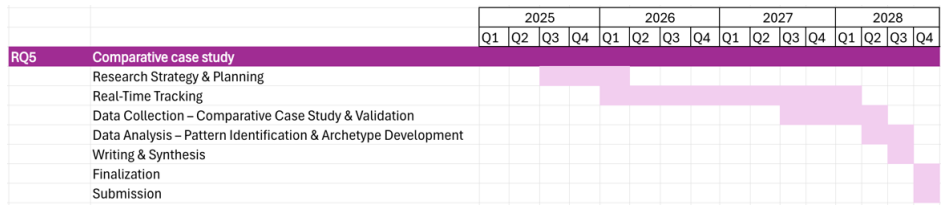


Figure 9: Planning Study 5

DOCTORAL CONSORTIUM

ORGANIZING FOR CYBER RESILIENCE: BALANCING RESILIENCE PARADIGMS FOR SUSTAINABLE IT BUSINESS VALUE

ZIGGY VAN GIEL

University of Antwerp, Faculty of Business and Economics, Antwerp, Belgium
ziggy.vangiel@uantwerpen.be

Organizational cyber resilience has been mentioned in literature to take a business-first perspective on ensure both short-term and long-term achievement of business objectives through technology. Nevertheless, it appears that it is still mostly approached from a technical perspective. When a business-first perspective is used, literature makes suggestions that are neither based on theory, nor on empirical findings. By using a design science research approach and relying on a combination of literature, surveys, interviews and case studies, this PhD project aims to develop strategies to achieve organizational cyber resilience by proposing specific organizational capabilities.

DOI

[https://doi.org/
10.18690/um.fov.4.2025.55](https://doi.org/10.18690/um.fov.4.2025.55)

ISBN

978-961-286-998-4

Keywords:

cyber resilience,
IT governance,
digital strategy,
engineering resilience,
ecological resilience



University of Maribor Press

1 Introduction

The increasing reliance of organizations on IT for the achievement of their business objectives in combination with the increasing frequency and sophistication of external threats pose significant risks to organizations' survival. For that reason, there has been a call to transition from a traditional IT risk management approach to one centered around cyber resilience. Organizational cyber resilience is in this context defined as “the ability to deliver the intended outcome despite adverse cyber events”. As compared to other concepts such as cybersecurity, it is said to take the business as a starting point. Despite organizational cyber resilience being positioned as a business-driven approach, existing literature reveal a persistent technical perspective. This gap underscores the need for research that integrates IT business value into cyber resilience.

This doctoral project aims to provide insights into how organizations can effectively organize and structure themselves to balance engineering and ecological resilience perspectives in the context of IT, ensuring both continuity through fast recovery and sustainable survival in an evolving external environment. Finally, this enables the organization to deliver short-term continuity and ensure long-term sustained IT business value.

2 Problem definition

Despite the growing body of knowledge on cyber resilience, existing literature predominantly focus on technical and operational aspects, often neglecting the organizational, business, and strategic dimensions. Nevertheless, cyber resilience is said to take the business as a starting point instead of the technical considerations. Because cyber resilience evolved in parallel from different domains and is often applied to specific problems, there is still conceptual ambiguity regarding cyber resilience.

Next, academic literature presents abilities that are said to lead to organizational cyber resilience. However, those propositions are neither based on empirical observations, nor on existing theoretical foundations. They are rather suggested abilities that would increase the chance of an organization to be cyber resilient. However, applications of those frameworks to case studies has yielded differing

results regarding the relevance of them. Also, there is no differentiation based on specific contingency factors. However, from contingency theory, and its application to IT governance theory, we know that there is no silver bullet that would work for every organization.

Finally, cyber resilience has been discussed as the next evolution of both cybersecurity and IT risk management. Combined with the conceptual relationship between cyber resilience and the broader (organizational) resilience domain, there is no consensus on the outcome of organizational cyber resilience. Depending on the specific niche, it can be business continuity, security, or something else.

Combining all the above, this research focuses on the following elements. Firstly, there is a need to provide conceptual clarity on what organizational cyber resilience entails, how it can be defined, and what fundamental paradigms form its basis. Secondly, based on the conceptual clarity and definition of organizational cyber resilience, the outcomes of it should be identified. Thirdly, theoretically grounded organizational capabilities for cyber resilience should be proposed that are validated through empirical observations. Combined these elements should answer the following research question: *How can organizations design and implement strategies to enhance their organizational cyber resilience?*

To conclude, this project is centered around (1) the antecedents or enablers of organizational cyber resilience, (2) the conceptualization, definition and theoretical paradigms underlying cyber resilience, and (3) the outcomes of cyber resilience for organizations. This is visualized in Figure 1.

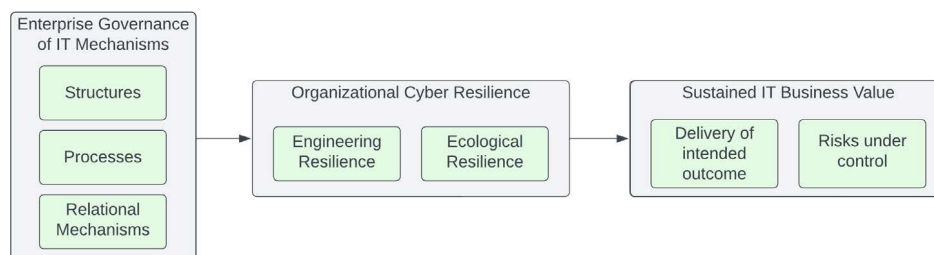


Figure 1: Conceptual model

3 Methodological approach of the different work packages

This doctoral project adopts a mixed-methods and design science approach conducted in different phases. While the overarching methodology can be specified as design science research with the aim to design strategies to improve organizational cyber resilience, different separate research initiatives can be identified. First, based on existing literature a bibliometric analysis and systematic literature review (SLR) are conducted. While the former aims to uncover trends in cyber resilience literature, the latter aims to provide conceptual clarity on what organizational cyber resilience entails, what theoretical perspectives are being used, and what is needed to achieve it. Consequently, organizational agility has been identified as an important aspect of organizational cyber resilience. Therefore, based on an international survey, the impact of different aspects of enterprise governance of IT on agility is analyzed. Next, because the capabilities proposed in literature are rather suggestions and not based on theory, a next phase employs a design science approach to propose capabilities that balance rigor and relevance. Finally,

3.1 Literature review: Bibliometric analysis & SLR

The first phase consisted out of a quantitative and qualitative literature review. First the quantitative literature review was a bibliometric analysis focusing on cyber resilience to uncover the trends in the domain. For this the process described by Zupic and Cater (2015) on bibliometric methods was followed to analyze the current trends in cyber resilience research while considering the potential evolution towards IT business value. A search string was evaluated on February 6th, 2024, and re-evaluated on February 7th, 2025.

Next, also a scoping literature review on organizational cyber resilience was performed. The review followed the five-stage methodological framework proposed by Arksey and O'Malley (2005), with enhancements from Levac et al. (2010) to ensure rigor and transparency. A scoping review is particularly suitable for clarifying key concepts, identifying thematic characteristics, and identifying knowledge gaps (Munn et al., 2018; Peters et al., 2021). The literature selection process is visualized in Figure 2. After the selection, the data was thematically analyzed to uncover patterns

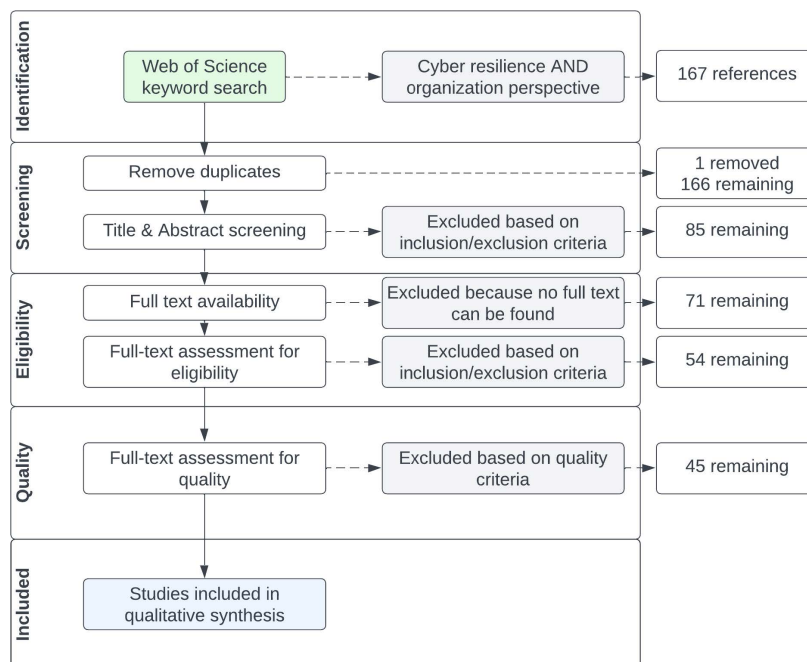


Figure 2: Scoping review literature search and selection strategy

3.2 Quantitative data collection: International survey based on COBIT 2019

The previous qualitative literature identified the organizational capabilities that are instrumental for organizational cyber resilience. Based on these capabilities, it became apparent that the operationalization of organizational agility based on Chakravarty et al. (2013) and Tallon and Pinsonneault (2011) highly aligns with organizational cyber resilience.

In collaboration with ISACA, an international survey was conducted in 2023 to measure the self-reported achievement of COBIT 2019 objectives, alignment goals, enterprise goals. Next, the survey also included questions on organizational agility based on Chakravarty et al. (2013) and Tallon and Pinsonneault (2011). As such organizational agility was operationalized through questions on three dimensions: entrepreneurial agility, adaptive agility, and business process agility.

We received 848 unique responses to the survey. However, 229 responses were incomplete as the survey was terminated before finalizing it, totaling to 619 potential valid responses. All questions on the achievement of the COBIT 2019 objectives, alignment goals and enterprise goals used a 5-point Likert scale with the option to answer “Don’t know”. Respondents that answered “Don’t know” to every question or that gave the same answer to every question (e.g., score 2 for every question) were dropped. The reasoning was that it is very unlikely that a company has the same maturity for every aspect. The more plausible explanation would be that the respondent went quickly through the entire survey. As a result, an additional 51 observations were dropped, bringing the final total of valid responses to 568.

The data analysis included different elements. First, using some descriptive statistics an overview is proved of the sample on different dimensions such as sector, company size, threat landscape, compliancy requirements, or strategic role of IT (Nolan & McFarlan, 2005). Next, using Shapiro-Wilk test the normality of the data was assessed and assumed to be not-normally distributed. As the data originates from a Likert-scale questionnaire, the data is ordinal. Therefore, non-parametric techniques are used such as Mann-Whitney U test or partial least squares path modeling. The former is used to compare whether there are any differences between different sub-sets, while the latter is used to answer the research question on how IT governance impacts organizational agility.

3.3 System-theoretical development of organizational capabilities

Existing literature presents different abilities that organizations should possess to improve chances of cyber resilient behavior. However, they are neither empirically validated, nor theoretically grounded. Also, the theoretical paradigms used in contemporary organization cyber resilience literature appears to be highly aligned with the viable systems model of Stafford Beer (e.g., (Beer, 1979, 1981, 1985; Espejo & Reyes, 2011)). Namely, the ecological and engineering resilience perspective align with the different systems that aim to either focus on internal stability and synergies, or on adaptation to environmental changes.

As a result, a separate work package will aim to propose high-level organizational capabilities, theoretically grounded in the viable systems model. Finally, using case studies and interviews these high-level capabilities will be validated empirically.

3.4 Qualitative data collection: Case studies & Interviews

There is a close collaboration with an IT consulting company during this PhD project. This project is namely funded by the consulting firm by employing the PhD researcher. This means that the researcher has access to completed and ongoing consulting projects relating to different aspects of this PhD. Next, the researcher has also the ability to be actively involved in those projects. The ambition is to include those consulting projects in the future phases of this PhD project.

Obviously, as the PhD researcher is employed by the consulting firm, there is a clear risk of independence when the researcher participates in the consulting projects with the aim to include them in academic research. Nevertheless, this risk would be mitigated by either one of two strategies. First, case studies and/or interviews will be used to validate intermediate results. For example, this project proposes both organizational capabilities and strategies for enhancing organizational cyber resilience. Using interviews, feedback will be obtained on the proposed capabilities and strategies, while exemplary case studies will be used to illustrate the appropriateness of them.

Secondly, case studies and interviews will be used to supplement or illustrate specific aspects. For example, it has been shown in the literature review that management awareness is an important precondition for organizational cyber resilience. Also, depending on organizational characteristics, it has been shown that specific IT management and governance processes are more important. Based on these or other examples, a single exemplary case study, or a multiple extreme case study design can be used to illustrate those difference and their impact.

3.5 Bringing it all together: Design science taxonomy development

Based on the combination of the aforementioned research initiatives, a taxonomy will be developed for organizational cyber resilience strategies depending on organizational characteristics. For this, the design science research guidelines of Hevner et al. (2004) will be followed, combined with the taxonomy development method of Nickerson et al. (2013) as shown in Figure 3. This combination ensures the taxonomy of strategies are developed iteratively by incorporating continuous feedback when additional data is gathered and analyzed. While the design science

research approach ensures the proposed taxonomy to be both relevant for practice and academically rigorous, the taxonomy development model outlines the process of continuously integrating both theory and empirical findings.

That way this PhD project hopes to present strategies that are both grounded in theory and empirically validated, as compared to existing literature. Also, these insights should enable practitioners to tailor the strategies to the organization's needs based on specific characteristics.

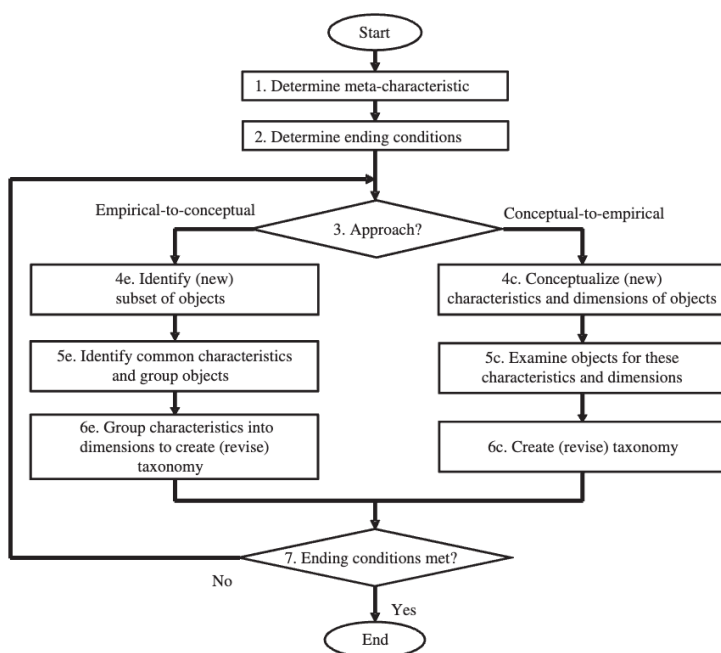


Figure 3: The taxonomy development method of Nickerson et al. (2013)

4 Risks, challenges and opportunities

This PhD project has some obvious risks, challenges and opportunities. However, the focus here will only be on the most significant differentiator as compared to other PhD projects, the explicit collaboration with an IT consulting firm.

While this offers some unique opportunities in terms of valorization and access to empirical data there are some risks and challenges that need to be discussed. Firstly, as already stated, the PhD researcher is employed by the consulting firm, which introduces the risk of conflict of interest. This risk is acknowledged and will be mitigated using the aforementioned strategies. Next, the involvement of the consulting firm could potentially hinder any alternative collaborations outside the consulting firm's network. On the contrary, when other organizations might be approached, the consulting firm might be interested from a commercial point of view, rather than an academic one. Additionally, approaching potential organizations or professionals to participate in the research becomes increasingly more difficult when an IT consulting firm is involved.

5 Conclusion

To conclude, the preliminary findings based on existing academic literature, systems-theory and two international surveys offers a good foundation for the future development of this project. Nevertheless, this project is characterized by some unique challenges and opportunities due to the involvement of an IT consulting firm as the main sponsor of this PhD project.

References

- Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32. <https://doi.org/10.1080/1364557032000119616>
- Beer, S. (1979). *The heart of enterprise*. Wiley.
- Beer, S. (1981). *Brain of the firm* (Second ed.). John Wiley & Sons Ltd.
- Beer, S. (1985). *Diagnosing the system for organizations*. Wiley.
- Chakravarty, A., Grewal, R., & Sambamurthy, V. (2013). Information Technology Competencies, Organizational Agility, and Firm Performance: Enabling and Facilitating Roles. *Information Systems Research*, 24(4), 976-997. <https://doi.org/10.1287/isre.2013.0500>
- Espejo, R., & Reyes, A. (2011). *Organizational Systems: Managing Complexity with the Viable System Model*. Springer.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in Information Systems research [Review]. *Mis Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- Levac, D., Colquhoun, H., & O'Brien, K. (2010). Scoping studies: advancing the methodology [Article]. *Implementation Science*, 5, Article 69. <https://doi.org/10.1186/1748-5908-5-69>
- Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(1). <https://doi.org/10.1186/s12874-018-0611-x>

- 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. <https://doi.org/10.1057/ejis.2012.26>
- Nolan, R., & McFarlan, F. (2005). Information technology and the board of directors. *Harvard Business Review*, 83(10), 96-106, 157. <https://www.ncbi.nlm.nih.gov/pubmed/16250628>
- Peters, M., Marnie, C., Tricco, A., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C., & Khalil, H. (2021). Updated methodological guidance for the conduct of scoping reviews [Article]. *JBIEVIDENCE IMPLEMENTATION*, 19, 3-10. <https://doi.org/10.1097/XEB.0000000000000277>
- Tallon, P. P., & Pinsonneault, A. (2011). Competing Perspectives on the Link between Strategic Information Technology Alignment and Organizational Agility: Insights from a Mediation Model. *Mis Quarterly*, 35(2), 463-486. <Go to ISI>://WOS:000290842900011
- Zupic, I., & Cater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods*, 18(3), 429-472. <https://doi.org/10.1177/1094428114562629>

38TH BLED ECONFERENCE

EMPOWERING TRANSFORMATION: SHAPING DIGITAL FUTURES FOR ALL: CONFERENCE PROCEEDINGS

ANDREJA PUCIHAR (ET AL.)

University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
andreja.pucihar@um.si

The Bled eConference, organised by the University of Maribor, Faculty of Organizational Sciences, has been shaping electronic interactions since 1988. The theme of the 38th conference is "Empowering transformation: Shaping digital futures for all". The central theme highlights the need for a digital transformation that is focused on the well-being of individuals and society as a whole. It emphasizes the importance of inclusive and sustainable development that enables equal participation of diverse stakeholders in shaping a digital future for all—one that is grounded in ethical, social, and sustainable principles. Within this framework, digital technologies are viewed as enablers for the transformation of existing systems and solutions in a human-centered way. The goal is to foster innovation, strengthen resilience, and promote long-term societal and economic progress. The papers in this conference proceedings explore a range of topics including the opportunities and challenges of the twin transition, emerging technologies, artificial intelligence and data science, decision analytics for business and societal changes, digital innovation and business models, restructured work and the future workplace, digital health, digital ethics, digital education, smart sustainable cities, digital consumers, and the digital transformation of the public sector.

DOI
[https://doi.org/
10.18690/um.fov.x.2025](https://doi.org/10.18690/um.fov.x.2025)

ISBN
978-961-286-998-4

Keywords:

digital transformation,
digital innovation,
twin transition,
sustainability,
digital transition,
green transition,
digital future,
digital technologies



University of Maribor Press



University of Maribor

Faculty of Organizational Sciences

