BLED eCONFERENCE 2021

UNI BIZ TECH GOV

34ᵀᴴ BLED eCONFERENCE

DIGITAL SUPPORT FROM CRISIS TO PROGRESSIVE CHANGE

June 27 – 30, 2021, Online
Conference Proceedings

Andreja PUCIHAR
Mirjana KLJAJIĆ BORŠTNAR
Roger BONS

CRIPPS Helen
SHEOMBAR Anand
VIDMAR Doroteja

EDITORS

University of Maribor Press

UNI BIZ TECH GOV

University of Maribor

Faculty of Organizational Sciences

34th Bled eConference Digital Support from Crisis to Progressive Change

June 27 – 30, 2021, Online

Conference Proceedings

Editors
Andreja Pucihar
Mirjana Kljajić Borštnar
Roger Bons
Helen Cripps
Anand Sheombar
Doroteja Vidmar

June 2021

Title *Naslov*	**34th Bled eConference – Digital Support from Crisis to Progressive Change**
Subtitle *Podnaslov*	**June 27 – 30, 2021, Online Conference Proceedings**
Editors *Uredniki*	Andreja Pucihar (University of Maribor, Faculty of Organizational Sciences, Slovenia) Mirjana Kljajić Borštnar (University of Maribor, Faculty of Organizational Sciences, Slovenia) Roger Bons (FOM University of Applied Sciences, Germany) Helen Cripps (Edith Cowan University, Australia) Anand Sheombar (HU University of Applied Sciences Utrecht, Netherlands) Doroteja Vidmar (University of Maribor, Faculty of Organizational Sciences, Slovenia)

Plomp (Vrije Universiteit Amsterdam), Miran Pobar (University of Rijeka), Andreja Pucihar (University of Maribor, Faculty of organizational sciences), Uroš Rajkovič (University of Maribor, Faculty of organizational sciences), Pascal Ravesteyn (HU University of Applied Sciences Utrecht), Liana Razmerita (Copenhagen Business School), Christopher Reichstein (Baden-Wuerttemberg Cooperative State University), Marcel Sailer (DHBW Heidenheim), Navin Sewberath Misser (HU University of Applied Sciences Utrecht), Anand Sheombar (HU University of Applied Sciences Utrecht), Muhammad Shuakat (Epworth Healthcare), Luuk Simons (Delft University of Technology), Koen Smit (HU University of Applied Sciences Utrecht), Ulrik Söderström (Digital Media Lab, Applied physics and electronics), Ronald Spanjers (Roessingh Rehabilitation), Susanne Strahringer (TU Dresden), Christian Thiel (OST Fachhochschule Ostschweiz), Edvard Tijan (University of Rijeka), Nalika Ulapane (University of Melbourne), Marko Urh (University of Maribor, Faculty of organizational sciences), Esther van der Stappen Avans University of Applied Sciences), Stan van Ginkel (HU University of Applied Sciences Utrecht), Doroteja Vidmar (University of Maribor, Faculty of organizational sciences), Doug Vogel (Harbin Institute of Technology), Pirkko Walden (Institute for Advanced Management Systems Research and Åbo Akademi University), Thorsten Weber (UCAM Universidad Católica San Antonio de Murcia), Markus Westner (OTH Regensburg), Gunilla Widén (Åbo Akademi University), Thomas Wozniak (Lucerne University of Applied Sciences and Arts), John Zelcer (Swinburne University Of Technology), Hans-Dieter Zimmermann (Eastern Switzerland University of Applied Sciences, Campus St. Gallen)

<table>
<tr><td>**Technical editors**
Tehnična urednika</td><td>Aljaž Murko
(University of Maribor, Faculty of Organizational Sciences)

Jan Perša
(University of Maribor, University Press)</td></tr>
<tr><td>**Cover designer**
Oblikovanje ovitka</td><td>Jan Perša
(University of Maribor, University Press)</td></tr>
<tr><td>**Graphics material**
Grafične priloge</td><td>Authors</td></tr>
<tr><td>**Conference**
Konferenca</td><td>34th Bled eConference - Digital Support from Crisis to Progressive Change</td></tr>
<tr><td>**Location and date**
Kraj in datum</td><td>Online, June 27 –30, 2021</td></tr>
<tr><td>**Organizing committee**
Organiazacijski odbor</td><td>Matthias Baldauf (Eastern Switzerland University of Applied Sciences, Campus St. Gallen), Roger Bons (FOM Hochschule), Christer Carlsson (IAMSR/Abo Akademi University), Helen Cripps (Edith Cowan University), Matt Glowatz (University College Dublin), Christian Kittl (evolaris next level Research Centre), Mirjana Kljajić Borštnar (University of Maribor, Faculty of organizational sciences), Gregor Lenart (University of Maribor, Faculty of organizational sciences), Marjeta Marolt (University of Maribor, Faculty of organizational sciences), Sanda Martinčić Ipšić (University of Rijeka), Guido Ongena (HU University of Applied Sciences Utrecht), Marijn Plomp (Vrije Universiteit Amsterdam), Andreja Pucihar (University of Maribor, Faculty of organizational sciences), Pascal Ravesteyn (HU University of Applied Sciences Utrecht), Juergen Seitz (Baden-Wuerttemberg Cooperative State University), Anand Sheombar (HU University of Applied Sciences Utrecht), Koen Smit (HU University of Applied Sciences Utrecht), Esther van der Stappen (Avans University of Applied Sciences), Johan Versendaal (HU University of Applied Sciences Utrecht), Doroteja Vidmar (University of Maribor, Faculty of organizational sciences), Doug Vogel (Harbin Institute of Technology), Pirkko Walden (Institute for</td></tr>
</table>

Advanced Management Systems Research and Åbo Akademi University), Nilmini Wickramasinghe (Swinburne University Of Technology / Epworth Healthcare), Hans-Dieter Zimmermann (Eastern Switzerland University of Applied Sciences, Campus St. Gallen)

34th Bled eConference
Digital Support from Crisis to Progressive Change
June 27 – 30, 2021 (online), conference proceedings
A. Pucihar (ed. et al.)

University of Maribor Press

Table of Contents

A SIMULATION MODEL FOR COVID-19 PUBLIC HEALTH MANAGEMENT: DESIGN AND PRELIMINARY EVALUATION

ROGER CLARKE

Xamax Consultancy Pty Ltd, Canberra, Computer Science, ANU, UNSW Law, Sydney, Australia; e-mail: Roger.Clarke@xamax.com.au

Abstract The COVID-19 pandemic has presented governments with challenges not only in relation to bio-medical understanding, medical treatment and health facility operations, but also the management of public health, public behaviour and the economy. In the area of public health management, discrete event simulation modelling is capable of providing considerable assistance to decision-makers. In April 2020, on the basis of publicly available information about the virus and its impacts, an analysis was undertaken of the needs of public health policy-makers, and a 16-state / 40-flow model was postulated. The model was revisited in December 2020, and experiences around the world applied in order to evaluate the model's apparent usefulness. This resulted in improved appreciation of its applicability and limitations, a revised model, and plans for further evaluation and application.

1   Introduction

From time to time, viral epidemics within individual countries threaten the health and lives of that country's inhabitants, and may wreak havoc on social and economic activities. Once the threat has passed, recovery may be quite brisk, provided that the country is sufficiently economically open. Global pandemics, on the other hand, harbour the potential for health impacts over large regions and potentially the whole world, and may have longer-term impacts on economic wellbeing because all countries' economies have been hampered and hence drivers of recovery are in short supply.

A century after 'The Maybe-Spanish Flu' at the end of World War I, the world was subjected to 'The Maybe-Chinese Coronavirus'. Naturally, ways were sought in which information technology (IT) could play a constructive role in the public response to the pandemic. Foreground needs existed in the area of prevention and treatment of the conditions that afflict patients. The particular need that is the focus of the present paper was for assistance in public health management, which seeks to slow the spread of the virus, protect particularly vulnerable sub-populations, ensure capacity to treat sufferers, and ultimately defeat the virus, while sustaining public confidence and achieving sufficiently high levels of compliance.

Emergency funding was provided to enable experiments with medical tools (e.g. for infection-testing, antibody-testing, symptom treatments, discovery of the modes of transmission, spread-containment mechanisms and vaccination) and with computing tools (e.g. for contact-detection, proximity-monitoring, contact-tracing, data management and decision support). As is to be expected of urgent, rapidly-performed experiments with available tools and the brisk conception and development of new tools, a great many projects were ineffective and short-lived. A few, however, delivered very considerable benefits to individuals, societies and economies.

The domain addressed by this paper is support for decision-making about public health policy. There is a worldwide need for contributions to public policy in the area. Yet IT's contributions have been at best mediocre. Data gathering, reporting and graphical presentation are helpful, but far from adequate assistance to decision-makers, and in any case data collection and analysis have been haphazard and often

ill-informed. To constitute information, and to enable the people responsible for public health management to make decisions, data must have context. That context may be provided by each individual policy-maker's own mental model. However, major programmes of this nature involve many stakeholders with diverse perspectives. The context is therefore multi-dimensional, it features competition among values, and the conception of the problem-space needs to be shared rather than personal.

The most powerful form of context is provided by models that are shared, that impose some degree of formality on the problem-space, that are sufficiently graphic that all stakeholders can relate to them, and that have an associated terminology that is reasonably common among the stakeholders. Given such a model, it becomes much easier to identify data that would be valuable input to deliberations, to generate and evaluate alternative courses of action, and to assess both the potential and the actual impacts of interventions.

The focus of this article is on a particular form of modelling tool, commonly referred to as 'discrete-event simulation' (DES). DES modelling enables expression of a model that represents a set of COVID-19 states that individuals may pass through, the conditions that determine the paths they follow, and key characteristics of both the states and the transitions. Research was conducted whose purpose was to postulate a model, apply it, assess its efficacy and improve it, in order to support mind-experiments and conversations about the real world into which public policy-makers were injecting successive interventions and refinements of interventions. As the purpose was the creation of an artefact, the appropriate approach was design science.

A DES model is a socio-technical artefact as that term is used in design science (Niedermann & March 2012, Gregor & Hevner 2013, p.337). The process described by Peffers' Design Science Research Methodology (DSRM) commences with problem identification and definition of objectives, which are followed by design and development (Peffers et al. 2007). Peffers et al. distinguish two related phases towards the end of the design research approach. 'Demonstration' of the use of the artifact solves one or more instances of the problem (e.g. by means of experimentation, simulation or case study), whereas 'evaluation' involves more formal observation and measurement of the new artefact's effectiveness in

addressing stated objectives. This research included Peffers phases of Problem Definition, Objectives Definition, Design and Expression, and a Demonstration step.

The paper commences by briefly summarising key features of the COVID-19 pandemic during the period March to December 2020. An outline is provided of the scope for modelling to assist, at various levels of investigation and decision-making. It is argued that discrete-event simulation modelling has a good fit to the needs of public health management. A model is presented which was devised in April 2020 on the basis of then-available information about the pandemic and government responses to it. Developments in the field during the following 8 months are identified, and their implications for that model are investigated. It is concluded that such a model can provide an effective contribution by IT to the decision processes of public health policy-makers.

2   The COVID-19 Pandemic

The new virus first came to public notice in the form of an epidemic in the Chinese city of Wuhan beginning in December 2019. Unsurprisingly, it took some time to be recognised and then accepted as a serious threat to public health. On 11 March 2020, based on rapid growth in detected case-numbers in northern Italy, Iran and South Korea, the World Health Organisation (WHO) declared a pandemic. By the end of March 2020, it had exploded in the USA, Spain, Germany and France, with rapid spread emergent in many other countries.

In most countries, there was an early peak of infections lasting 2-4 months with deaths following after a 1-3 week lag, then a lull, then 6-8 months later a 'second wave' in many cases worse than the first (Econ 2020). By the end of 2020, substantial second waves were infecting very large numbers of people and killing large numbers, with the cumulative (known) case-count worldwide past 80m and the death-count approaching 2m. On these measures, only two pandemics of the last century have been worse: the 'Spanish Flu' of 1918-20, and HIV/AIDS since 1980.

The cause was identified as a form of coronavirus, spread primarily by an infected person coughing or sneezing, or perhaps even speaking or breathing out, contaminated droplets (over a range of perhaps 1m), or possibly aerosols or droplet

nuclei (very small droplets, over a range of perhaps 3-4m), or by direct contact with another person, or by contaminating 'fomites', i.e. objects and surfaces in the infectee's immediate environment (WHO 2020b).

Susceptibility appeared to be quite low under 10 years of age, increasing with age, and very high for those over 70. Impacts on individuals ranged from short-term, unpleasant but variable experiences, to very serious lung malfunction and death from that or consequential causes. Over time, it became apparent that there are small but significant numbers of people who suffer impacts for an extended period after the initial (predominantly pulmonary) impact of the virus (SWPRS 2020). However, most infectees are asymptomatic, decreasing the likelihood of detection and hence increasing the likelihood of spread.

A person with the virus may be infectious from 1-3 days before symptom onset, then for a further 1-2 weeks for asymptomatic persons, up to 3 weeks in mild to moderate cases, but much longer in severe cases (WHO 2020c). There were no known treatments for the virus itself. The proportion of hospitalised patients needing admission to Intensive Care Units (ICUs) ranged from 5% to 15%. In some regions, ICU capacity proved inadequate.

Mortality was very heavily skewed towards people over 70, with the likelihood of death much higher for those with bronchial and some other relevant or otherwise debilitating conditions. Employees in hospitals and aged care homes were at risk of high viral load, and high-quality hygiene and personal protective equipment (PPE) were essential. Despite precautions, many health care workers succumbed, in the USA about 3,000, 1% of the country's more than 300,000 deaths during 2020 (Gn 2020b).

The focus of public health actions was on prevention of spread, most urgently among those at greatest risk. The public health imperative is constrained by the limitations of enforcement powers and resources, and by conflict with freedoms of action and movement, and with economic management. In some jurisdictions, those challenges were exacerbated by a lack of political will. Countries adopted varying approaches to public health management, with highly varying senses of urgency, varying levels of compliance by the public, and highly varying case-counts, fatality-counts and fatality-rates (WOM 2020).

3 Modelling

IT was applied to the gathering and publication of data, partly to inform and entertain the public, but more critically to support policy-makers in their efforts to understand the phenomenon. This paper investigates the question as to whether the seeming absence of an 'enterprise model' of the undertaking, and of 'information architecture' and 'data models' to support it, have hampered the potential IT contribution, and hence whether return on investment in IT can be improved by applying insights from modelling theory and practice.

A model is a simplified representation of a real-world system, which reflects interdependence among the relevant entities, structures and processes. Real-world socio-economic systems are open, complex and highly inter-connected. Simplification necessarily involves limiting the scope of the model, by placing the focus on one sub-system or two or more closely-related sub-systems, at one particular level of abstraction, and by excluding some factors and using proxies for others (Gault et al. 1987). A model therefore cannot replicate the real-world system (von Bertalanffy 1968). However, if key factors are appropriately reflected, experimentation with a model can deliver insights. At the very least, experimentation can suggest what data might be the most valuable to collect. In addition, participation in the modelling process may enhance observers' understanding of the world, and assist in making decisions about actions to take.

During the first quarter of 2020, it became clear that COVID-19 had a high infection-rate and was life-threatening for some categories of people. As the epidemic in Wuhan developed into a pandemic, it became increasi8ngly apparent that decision support systems (DSS), and the modelling activities intrinsic to DSS (Sprague 1980, p.1), needed to be applied.

During the pandemic, epidemiological models of the SEIR(D) family were much-discussed. These treat the population as comprising Susceptible individuals (S – those able to contract the disease), Exposed individuals (E – those who have been infected but are not yet infectious), Infective individuals (I – those capable of transmitting the disease), Recovered individuals (R – those who have become immune), and possibly Dead individuals (D).

Generally, however, models in the SEIR(D) family lack details needed for public health management purposes: "These models have [provided] information about tipping points and inform[ed] policy decisions. But ... these models are not adequate for modelling the human behavioural aspects that are important in disease transmission and epidemic dynamics" (Siebers et al. 2010, p.206); and Schipper (2020) argued that modelling us during the pandemic, because "the current epidemic model is medical, and narrowly so" (p. 7).

This paper adopts the position that appropriate support by IS and IT for public-policy decision-makers in dynamic contexts like a pandemic depends on the application of appropriate modelling techniques. They need to be instrumentalist, with a social-engineering orientation. Such models depend on careful definition of the system scope, and the level of abstraction at which the system is being observed. Key requirements of public health management are the establishment and progressive adaptation of a model that clearly distinguishes start-point(s), states, transitions, and end-point(s), and that identifies key attributes of each individual passing through the model (e.g. age-range and relevant-prior-conditions), and supports experimentation with different distributions of those variables.

The following section postulates such a model. The adequacy of the model is then tested against the phenomena and interpretations of them reported during the following 8 months of 2020, and adaptations are proposed in an endeavour to improve the model's capacity to assist policy-makers.

4   Simulation Modelling for Public Health Management

This section first discusses particular needs that arose during the COVID-19 pandemic in 2020, then outlines the relevant form of modelling, and finally describes an application of it that is argued to be of benefit to policy-makers.

4.1   The Needs of Public Policy-Makers

The focus of public health management is **"population-based health protection and promotion"** (Novick & Morrow 2008, p.60), with efforts **"organized and directed to communities rather than to individuals"**, and with the prevention and control of epidemics high on the priority-list (Novick & Mays 2008, p.3). Key

functions and practices are (Novick & Morrow 2008, pp. 40-47) are assessment, policy development and assurance.

The target-area for modelling activities in public health management is accordingly the processes of the spread of the disease, and the purpose is to deliver insights to policy-makers regarding the shape that interventions may usefully take, and their likely contribution to containing that spread.

An important distinction is made in decision theory between **factors that are strategic or controllable and those that are environmental or uncontrollable** (Peterson 2009). A further distinction is necessary between directly-controllable factors and those that can only be indirectly influenced. For example, outputs include published government advice, formal declarations and laws, whereas outcomes comprise the acts of individuals, which are only influenced, not determined, by advice, declarations and laws. The extent to which public behaviour is compliant with the intentions of public health managers depends on controllable factors such as expression, channels of communication and timing, and on factors that are far less controllable, such as attitudes to authority, perceptions of the health threat, and prior experience of government actions.

Two of the key weapons in fighting epidemics are **quarantine and isolation**. The term 'quarantine' applies to people who have been, or are suspected to have been, exposed to an infectious disease, but who are not at that stage known to be infected. The term 'isolation', on the other hand, is applied to people known to be infected. However, supervision of suspect-quarantine and infectee-isolation may be dependent on inadequately-trained staff, contractors or military personnel. Travel restrictions are difficult to police. Records of attendance at venues are maintained by individuals and venue-operators, and assurance of data quality and data-compatibility is challenging. The implementation of border restrictions may be haphazard where multiple agencies are involved.

Public health activities inherently involve a very broad range of stakeholders, and great **diversity among perspectives and values** spanning the social, economic and psychological dimensions. As a result, decisions are actively contested, and the decision-making processes complex and at best only modestly well-structured. The Vroom-Yetton-Jago Decision Model identifies five decision-making

implementation styles (Vroom & Yetton 1973, Vroom & Jago 1988). For decisions that have significant impact and require input and 'buy-in' from many participants, the relevant two of the five are consultative (group-based but leader-decided) and collaborative (group-based and group-decided).

Because many stakeholders are involved, policy-makers are confronted by diverse views and a rich choice of experts, of approaches to models, of assumptions inherent within them, and hence of the findings presented by the modellers. For consultative and collaborative processes to be effective, participants need to have a shared understanding of the relevant domain and of the terminology used to describe it.

Models can contribute to that understanding by reflecting the key features of real-world systems that policy-makers seek to influence. Further, because pandemics develop in unpredictable ways, and new information and insights become available, policy-makers' appreciation of the context is adaptive. It is therefore crucial that policy-makers develop a degree of clarity about the context in which they are working, communicate that to modellers, and update modellers on changes in their perceptions of the relevant systems.

The most effective way in which modellers can contribute is to start with an appreciation of the relevant domain, to become familiar with the policy-makers' initial mental models, and to be sufficiently 'embedded' to detect changes in their thinking. Further, modellers must convey enough information about their purposes, their assumptions, the capabilities and limitations of their methods, the nature, quality and quantum of the data that they are using, and the extent to which it has and has not been feasible to test findings against the real world. Without great care, there is a high probability of misunderstandings, and of policy-makers being misled.

The following sub-section considers how a particular form of modelling can be used to address these needs.

4.2   The Modelling Method

Multiple forms of modelling exist. At the strategic level, for example, system dynamics is appropriate (Brailsford & Hilton 2000, Brailsford et al. 2014). A particular modelling approach that matches well to the needs of public health management during a pandemic is discrete-event simulation (DES) (Allen et al.

2015). DES modelling involves the identification of the various states that an entity (in this case a person) may be in, their transitions or flows from one state to another, and the factors that determine when transitions occur. A systematic review of publications on DES in health care, in Zhang (2018), concluded that DES has "rich potential ... to provide a broader picture of ... health care systems behavior" (p.9).

Currie et al. (2020), in discussing the application of simulation modelling technique to the COVID-19 pandemic, describe DES models as being "typically used to model the operation of systems over time, where entities (people, parts, tasks, messages) flow through a number of queues and activities. They are generally suitable for determining the impact of resource availability (doctors; nurses), on waiting times and the number of entities waiting in the queues or going through the system" (p.85). The Currie article identifies a range of potential applications of DES in the context of the pandemic. In Wood et al. (2020), a report is provided of a DES model "designed to capture the key dynamics of the intensive care admissions process for COVID-19 patients" (p.1). Beyond health facility management, Bolla & Sarl (2020) model flows of COVID patients in Switzerland from home to hospital to ICU and beyond and Jalayer et al. (2020) model "citizens living, working, pursuing their needs and travelling inside a geographical environment" (p.3).

Rhodes et al. (2020) perceive models for policy to "blend various heterogeneous data (quantitative, qualitative, abstract, empirical) from various diverse contexts (different viruses, countries, localities, studies, historical periods) ... to enable a decision" (p.2). The authors discuss an "approach to the modelling of pandemics which envisages the model as an intervention of deliberation in situations of evolving uncertainty" (p.1). "The model, precisely because it has latitude as a space of triangulation and speculation, potentiates a working relationship, in which dialogue is made possible" (p.6).

Many researchers assume that DES models have to be fed quantitative data, and that the calculations are what matters. For example, the text elided from the p.2 quotation from Rhodes et al. (2020) in the previous paragraph is '[blend] into a single calculative process". This ignores the considerable limits on the usefulness of quantitative analysis in such circumstances, whether conducted mathematically or numerically by experimentation. For example, a comparison across four models of the path of COVID-19 infections in South Africa (Chi et al. 2020) found very wide

variation in the models' predictions of case-counts and death-counts, highlighting the folly of reliance on any of them.

The Rhodes et al. article overlooks the fact that the 'blending', the 'deliberation', the 'speculation', the 'working relationship' and the 'dialogue' are all highly valuable in their own right, and may offer far better value to policy-makers than unverified rules applied to mediocre-quality data in a 'calculative process'. Table 1 identifies ways in which a suitable DES model can be applied in the style of a decision support system to enable Vroom-Yetton-Jago consultative or collaborative policy decision-making.

To the extent that the model is adequately articulated, tested for logical completeness, and checked against real-world activities, it is also capable of being used to simulate flows of people through the system, and 'stocks' of people currently in each state. This approach would need to be supplemented by a segmentation analysis, distinguishing in particular:

- high-risk-of-mortality categories, e.g. based on vulnerability (age, prior disposing conditions) and on intensity of exposure (health care staff);
- high-risk-of-being-highly-infective categories ('super-spreaders');
- high-risk-of-highly-infective-circumstances ('super-spreader events').

Table 1: Benefits of DES Modelling for Policy Decision Support

• The activity of building and reviewing a DES model can assist participants in gaining an **understanding** of the main states, and the main factors determining transitions between states, that are relevant to decisions about interventions and policy-settings

• That activity, through the conduct of workshops involving the relevant stakeholders, can assist in identifying differences among the stakeholders' mental models, and in achieving **commonality of understanding** of:
 	
    o relevant states, and transitions among them
    o relevant attributes of each state and each transition; and
    o key terms, arising from words being used within a structured context rather than loosely in unstructured conversations

• Consideration of particular actions in light of the model enables **sharing of information and insights** into the dynamics of the real-world system(s) that policy-makers seek to influence
• When alternative strategies are being considered, the model may assist stakeholders in the **identification of key data** needed to support evaluation of

	the strategies, and hence where resources can be usefully invested in data gathering and data quality assurance, and in finding proxies for data that is too difficult or too expensive to collect
•	The model can assist in **experimentation** with the impacts that different data-values might have on the evaluation process and inferences drawn from it
•	The model can be used as a basis for pre-implentation review of conclusions reached, and pre-publication review of public statements and data, to check that important factors have not been overlooked

By manipulating key parameters (such as detection-rate; the proportion of infectees needing admission to hospital and to ICU; hospital- and ICU-capacity; treatment-periods; and mortality-rates), estimates can be made of the limits to the ability of health facilities to cope, and the extent to which urgent investment in additional facilities might be necessary.

4.3   The Postulated Model

The purpose of this research was to investigate the extent to which a DES model could support public health management in the context of an rapidly-developing epidemic. During March-April 2020, I postulated a state-transition model, intended to represent the population of a jurisdiction, and the flow of individual members of it through various states associated with infection, hospitalisation, to recovery or death. The intention was to commence with the minimum complexity, in terms of the number of states, flows, and data about each, based on the available information about the challenges that public policy-makers were addressing. The model could then be experimented with, and expanded to the extent necessary to embody a sufficiently rich understanding of the public policy problem-space. Based on government publications and media reports, and taking account of previous SEIR models, it appeared that the model would need to incorporate about a dozen states, 30-40 flows, and data-items representing the key attributes of the people passing through the system.

Although a DES model can be applied computationally, that was not the intention, because the complexities and dynamism of the relevant part of the real world are such that the results would inevitably be spurious. The model is a framing tool for the problem-space, intended to help policy-makers formalise their own mental models, appreciate and resolve differences among those models, experiment with the model, and draw inferences relevant to the many decisions they needed to make during the weeks and months of the epidemic.

The first iteration of the model, of April 2020, is in Figure 1. A textual description of aspects of the model was also developed. Broadly, individuals were conceived as beginning as Uninfected, with a proportion passing through Infection, possibly via Hospitalisation, and on to Immunity or Death. Each of the four broad domains was conceived as encompassing a number of states, such as being in hospital, or in ICU, or in a queue to get into one of them. Various aspects of each state required some articulation, and so did the conditions under which transitions occur between states.

The following section outlines the steps undertaken in order to assess the potential of this model to support public health policy decision-makers.

5   Model Testing and Articulation

During the process of postulating the model, a variety of design issues arose. Some were formal questions, such as whether and on what basis some of the state-transitions could logically arise, and could be appropriately represented.

For example, it quickly became apparent that the representation of Testing as a state was inappropriate. A more useful approach was to specify attributes of each individual, which travelled with them as they passed through the network. The key attributes appeared to be tested-awaiting-result, tested-negative, and tested-positive. That removed one state and four flows, with no loss of model richness. For ease of reading, the state Quarantined was re-numbered as (2).

Many other issues, however, were concerned with the appropriate representation of real-world states and processes. In April 2020, it was too early to resolve those issues, with the result that the model depicted in Figure 1 was provisional, even tentative.
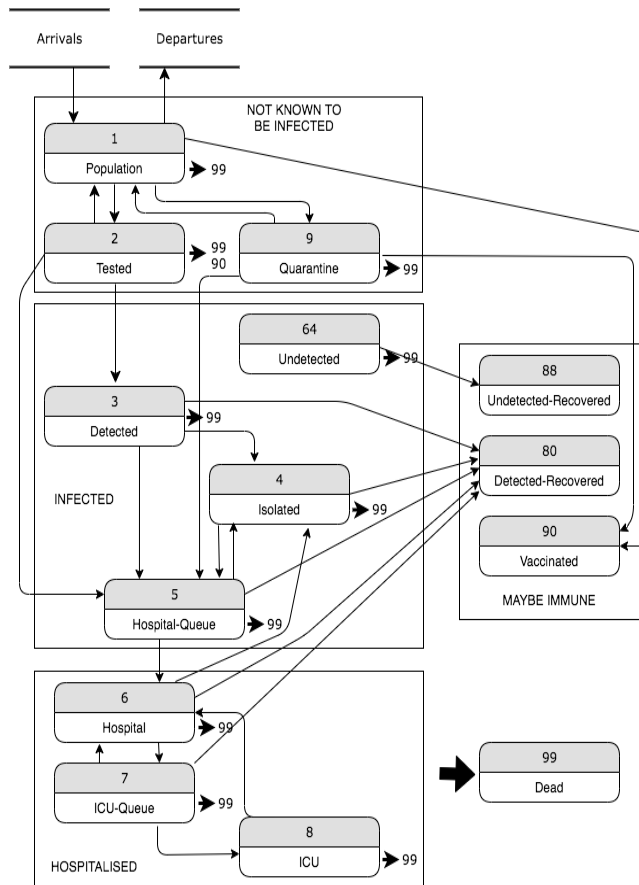
Figure 1: The Model Postulated in April 2020

In order to implement the demonstration phase of Peffers' design science method, experience needed to be gained concerning the relevant real-world systems and the appropriateness or otherwise of the representation of them in the model. A conventional way to gain such experience is through case studies. However, as the COVID-19 pandemic was only just developing, no directly-relevant case studies were available. Although case studies of other pandemics could have been sought out, it was already clear that there were distinct differences between the COVID-19 pandemic and other well-documented events, even other coronavirus events.

A more appropriate method might therefore have been to conduct a contemporaneous field study, seeking embedment within some particular jurisdiction's public health policy apparatus, and preparing a longitudinal case study of that jurisdiction's path, including uncontrolled events, interventions, and subsequent experiences. However, such access would have been very difficult to negotiate (not least in a context in which physical distancing was being imposed). It would also have to a considerable extent limited the testing and articulation process to the factors that arose in a specific jurisdiction. Each country has its own cultural context, and the events, the details of the interventions, the sequences of events, and the timings of events, varied greatly among different jurisdictions. It would be very challenging to try to draw generically useful inferences from such a field study.

An alternative approach was accordingly formulated. Monitoring was undertaken of the ongoing reporting of developments, interventions and experiences in countries worldwide. These reports provided a wide range of circumstances against which the efficacy of the model could be reviewed. This section summarises information about interventions, and identifies and briefly discusses some key themes that emerged.

Governments around the world responded to COVID-19 with a wide range of interventions intended to protect public health. A scan was undertaken of documents published by relevant international and national government agencies, including WHO (2020d) and ICAO (2020), supplemented by academic articles and media reports. Table 2 identifies mainstream public health interventions, clustered into six groups. It is important that IS and IT be brought to bear to assist policy-makers to judge the likely effectiveness of these actions in particular contexts, to design interventions, and to time and manage their implementation, adaptation and eventual withdrawal.

Because the infection-vector appeared to be primarily brief, airborne transmission from infectees to those close by, physical separation among people generally (often referred to using the misleading term 'social distancing') was widely adopted. Physical separation is also the objective of quarantine and isolation. Despite the differences in meaning, and the clear explanations provided by a range of national health agencies (e.g. CDC 2020), some agencies were not consistent in their uses of the terms. As a result, media reports evidenced considerable confusion, and it is

very likely that many people were unclear about the obligations of suspects, contacts and infectees.

Table 2: The Primary Public Health Interventions

Case Discovery, Management	Facility Restrictions, Closedown
- Identification of suspects - Quarantine of suspects - Testing of suspects - Isolation of infectees - Contact-tracing of infectees - Location of and communication with contacts	- Hospitals - Aged care facilities - Institutions, e.g. prisons - Group accommodation, e.g. backpacker dormitories - Face-to-face businesses (shops, personal services, gyms) - Workplaces - Entertainment venues - Public gatherings - Geographical areas (*cordon sanitaire*) - Pre-schools, schools, tertiary educational institutions
Personal Protective Measures	**Environmental Measures**
- Hand hygiene - Respiratory etiquette (sneeze/cough protection) - Avoidance of surfaces - Face-masks - Clinical Personal Protective Equipment (PPE) in hospitals and aged-care facilities	Cleaning of surfaces
Physical Distancing Measures	**Travel-Related Measures**
- Physical distancing in public places (1.5m / 4sqm) - Quarantining of suspects - Isolation of infectees 'Work-at-Home' Recommendations to Employers 'Stay-At-Home' Recommendations for at-risk segments	- Border restrictions - Border screening and testing - Border closure - Stay-at-home, work-at-home - Domestic movement restrictions - Public transport - Private vehicles - Walking

Countries around the world adopted vastly different approaches to interventions, recognised different triggering events, and timed their interventions differently. Most also changed their approaches over time. Despite the enormous differences in context, some comparisons are feasible, such as among Scandinavian countries. Sweden implemented only limited actions (physical distancing, bans on large gatherings, and travel restrictions), whereas its neighbours used additional and stronger interventions to reduce the opportunity for the virus to spread, including closedown of many more categories of venue, curfews and border closures. The outcome was a death-rate per capita in Sweden during 2020 that was 4-9 times those of its neighbours (Barrett 2020).

There may also be lessons to be learnt from the juxtaposition of the apparently worst examples of mismanagement and/or outcomes (in particular, the UK, the USA, Belgium, Brazil) and the most successful (e.g. China, Singapore, New Zealand, Australia). The UK flirted with a no-action 'strategy' rationalised as striving for herd immunity, overrode professional advice, lacked coherent and consistent leadership, reacted slowly to new information, and continually changed tack in a haphazard manner (Minghella 2020). A wide range of media reports and some semi-formal reviews gave rise to the list of public health management behaviours associated with serious failure in Table 3.

On the other hand, the most successful countries acted quickly and decisively. A range of interventions, and characteristics of interventions, were associated with success in preventing spread and/or reining in spread that had already begun (e.g. OxCGRT 2020).

Table 3: Behaviours Associated with Serious Failure

- Data suppression (e.g. during the first few weeks in Wuhan)
- Disparagement of the epidemic's seriousness by national leaders (e.g. USA, Brazil)
- Disregard for public health policy advice (e.g. USA)
- Denial of the efficacy of key interventions and/or support for 'quackery' (e.g. USA)
- Delay in the implementation of constraints (e.g. Belgium, UK, Sweden)
- Inaction justified as a means to rapidly attain 'herd immunity' (e.g. UK, Sweden)

- Weak enforcement of constraints (e.g. many countries, particularly in the early stages)
- Premature easing of effective constraints (e.g. many countries, particularly after the first wave)

The list of actions in Table 4 was prepared on the basis of reports about actions and outcomes in China (BBC 2020), New Zealand (Baker et al. 2020, Jefferies et al. 2020), Melbourne (Gn 2020a), and Vietnam and Taiwan (Whitworth 2020).

Table 4: Actions Associated with Success

Known-Infectee Control Measures

- Detect **infectees** early
- Isolate **infectees** immediately, and perhaps household members, especially partners
- Trace **close contacts of infectees** quickly
- Quarantine **close contacts of infectees**
- Impose closedown in and near **infection hot-spots**

Community-Spread Control Measures

- Suspend or dilute **large-scale events** in which people are closely-packed, including live entertainment, bars, clubs, churches, rallies, public transport services
- Suspend **sustained-contact circumstances**, including face-to-face retail, personal services and workplaces in which physical distancing cannot be achieved

High-Risk-Segment Protection Measures

- Shield **high-risk groups**, in particular through lockdown of health and aged care facilities against non-essential entry, application in those facilities of rigid hygiene, and provision to frontline health care staff of fresh, clinical-grade personal protection equipment (PPE)
- Quarantine **new arrivals into the jurisdiction**, at least until a test for infection returns a negative result

Policy-makers needed to make judgements about which interventions were needed, the specifics of their design and application, and the timing of introduction, easing, and suspension. Their judgments were affected by a great many factors. The efficacy of some actions was apparent from the beginning of the pandemic, whereas the value of other actions emerged slowly, with the gradual accretion of understanding.

In order to balance health safety against social and economic disruption, it was vital to be able to judge **the appropriate length of time for quarantine of suspects and isolation of infectees**. That depended on the ability to make reasonable determinations about the required period of confinement (requiring an estimate of when infection occurred) and the circumstances under which shorter and longer periods may be appropriate.

A guideline for discontinuing transmission-based precautions that was available in mid-2020 was that patients could be released 10 days after symptom onset plus 3-4 days without symptoms, or, in asymptomatic cases, 10 days after a positive test (WHO 2020a). In some countries, that was later variously extended to 14 days, or adjusted to permit discharge as soon as a negative test result was received. This reflected judgements made about the balance between the risk of transmission and the risk of reduced public support and hence compliance levels.

An important aspect of the public health problem is advance warning about the capacity of health facilities to cope with demand. It would be a valuable contribution if the model were able to assist in projecting **demand for and supply of hospital-beds and ICU beds**, in total, and by geographical area and hospital. This is likely to be dependent on recent testing rates, positive-result rates, hospitalisation-rates of positive cases, the proportion admitted to ICU, and the periods patients spend in those facilities. The source-data would need to be collected on an ongoing basis, in order to ensure that current indicators were readily to hand. This draws to attention another important attribute of individuals: **non-COVID admissions** to hospital queues and onwards. Factoring that in enables total demand for hospital and ICU beds to be modelled, and avoids confusing non-COVID-related transitions with those arising from the epidemic.

One of the challenging questions was the extent to which all infectees have much the same **degree of infectiveness**, or whether there are 'super-spreaders' who are much more prone to infecting other people and/or contaminating surfaces. If there is considerable variability, effort could be valuably invested in determining what infectee-attributes are associated with 'super-spreaders' and whether it is possible to focus available tracing and quarantine resources on people with those attributes. A complication arose to the estimation of infectiveness when, at the end of 2020, strains of the virus emerged that appeared to be substantially more infective.

The importance of the category of people in the the state **Undetected-Infected (64)** became apparent as the epidemic unfolded in each country, because undetected infectees are a primary source of virus-spread. In order to gain an insight into the overall progress of the epidemic at a population level, an estimate is needed of the Undetected-Infected status, for example by means of adequate random-sample testing of the public for the virus. Strategies are needed to find more of the people who are in that state, so that they can be requested or required to shift state to Isolated (4). Possibilities include extensions to contact-tracing, suspect-definition based on locations and time-periods, infection-testing in the vicinity of hotspots, and random infection-testing. It may be possible to estimate the scale of the count in Undetected (64), by random-testing for antibodies in order to develop estimates of the cumulative count in Undetected-Recovered (88), and to then reason back from there to the scale of current Undetected-Infecteds.

To reflect the uncertainties, there are benefits in using ghostly outlines to represent both Undetected-Infected (64) and Undetected-Recovered (88), and inflows to those two states. On the other hand, transitions are visible when an individual moves from Undetected (64) to Detected (3) or Hospital-Queue (5).

Considerable discussion arose about **the 'excess mortality' statistic**, and the ways in which COVID-19 affected that measure. It became clear that the terminal state Dead (99) needed to be categorised more finely. Cause of death needed to distinguish cases where COVID-19 was the cause of death, or was a significant factor in the death because it compounded prior conditions (99A) from all other causes of death, including not only where COVID-19 was not present, but also where infection was, or was assumed to be, present at death but was not listed as a cause (99B).

This phase of the research stimulated reconsideration of many aspects of the April 2020 model in Figure 1, in order to ensure that it could assist policy-makers in navigating their juridiction's particular maze. The result was the revised model in Figure 2, incorporating many adaptations reflecting the insights arising from 8 months of vicarious learning from many different jurisdictions. Further detail on the revised model is provided in the underlying Working Paper (Clarke 2021).

6   Conclusions

Despite the great contemporary enthusiasm for IT, it delivered relatively little value during the COVID-19 pandemic of 2020. This appears to have been attributable to an 'applied' approach, 'throwing technology at the problem' and at worst matching the caricature of 'when you have a hammer in your hand, everything looks like a nail'.

It appears more likely that IT can deliver for society and the economy if the approach adopted is both more strategic in nature, and 'instrumentalist' / problem-oriented rather than 'applied' / tool-oriented. That means standing far enough back to be able to identify and describe the problem-space, and then modelling the key aspects of that space. On that base, architectures, process models and data models can emerge and be refined, that will much better serve the needs of decision-makers.

The scale of activities, even within a single jurisdiction, has been so great that a detailed assessment of the models used during 2020 is difficult to assemble. The research reported in this paper comprised a mixture of thought-experiment, abstract design, and testing and adaptation of the initial model against information arising from experience across the world during the period May to December 2020.
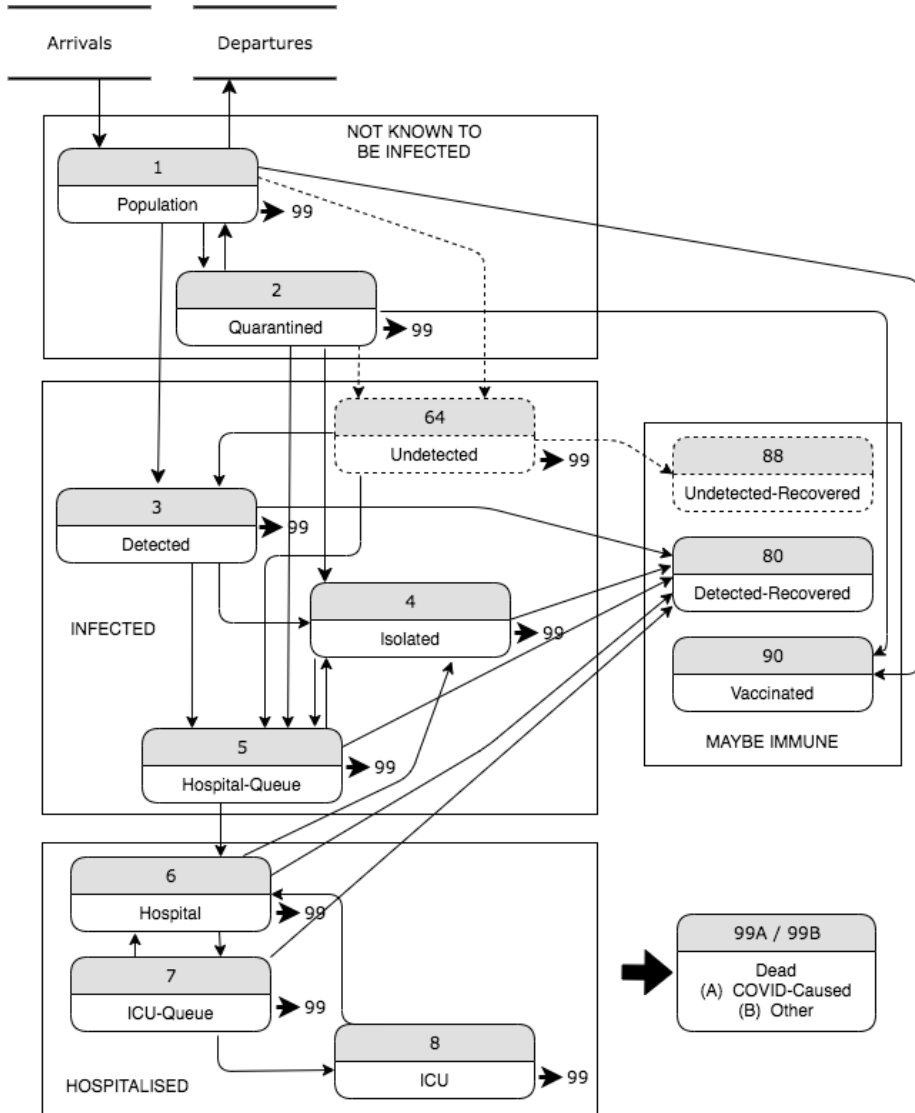
Figure 2: The Revised Model

A model was initially postulated that was envisaged as being suitable as a supporting tool for public policy decisions in relation to public health management. On the basis of new information streaming in during the subsequent months, the need for refinements was apparent. The revised model is capable of further articulation,

through specification of data models in support of states and flows, and alternative processing rules for state transitions.

An example of the kind of development that requires rapid adaptation was the mid-January 2021 revelation that the hospital readmission rate for people who had recovered from COVID-19, in the UK, during January to September 2020, was 30% – 3.5 times that for the population generally – and 12% died following discharge – 7 to 8 times that for matched control groups Ayoubkhani et al. 2021). This suggests that the incidence of 'Long COVID' or 'post-COVID syndrome (PCS)' may be much higher than previously thought. It also raises questions about the extent to which people in the Recovered (80, 88) and even Vaccinated (90) states are immune to COVID-19, and even whether they are incapable of becoming infective again.

The work reported here opens up a variety of possibilities for further research, including:

- case studies of individual jurisdictions' decision processes;
- articulation of the model in specific jurisdictions;
- presentation of the model to, and workshopping with, policy-makers, in order to gain further understanding of its usefulness;
- animation of the model, in order to assist in its communication and use;
- expression of the model in a DES language such as GPSS;
- experimentation with a mix of actual data and synthetic data.

The model proposed and tested in this paper has, by its nature, limited focus. Its target-area is expressly public health management, although it has application also to the adjacent level of health facility management, and implications for population management. It is not suggested that this model subsumes or replaces models at other levels of abstraction. It is contended, however, that public health policy-makers, and government ministers and their advisers, can greatly benefit from the development, articulation and ongoing adaptation of a model of this nature. It enables processes to be better understood, strategies considered, and implications of possible interventions thought through. This information can then be combined with that arising from work at other levels of abstraction (bio-medical and medical, on the one hand, and public behaviour management and economic management on the other).

A very large number of articles have been published during 2020 in the medical, public health management and social policy literatures. In the IS discipline on the other hand, significant contributions have been few and far between. However, see Thomas et al. (2020) and Trang et al. (2020). It is contended that, unless the IS discipline adopts considered, strategic approaches to public policy needs, proponents of IT will be dismissed as 'technological solutionists' and even 'delusionists', and IT and IS will lose their lustre.

References

Allen M., Spencer A., Gibson A., Matthews J., Allwood A., Prosser S. & Pitt M. (2015) 'What is discrete event simulation, and why use it?' Chapter 5 of 'Right cot, right place, right time: improving the design and organisation of neonatal care networks – a computer simulation study' NIHR Journals Library, May 2015, at https://www.ncbi.nlm.nih.gov/books/NBK293948/

Aron J.L. & Schwartz I.B. (1984) 'Seasonality and period-doubling bifurcations in an epidemic model' J. Theor. Biol. 110:665-679, 1984, at https://www.researchgate.net/profile/Ira_Schwartz/publication/16673793_Seasonality_and_period-doubling_bifurcations_in_an_epidemic_model/links/59d7a186a6fdcc2aad0645fc/Seasonality-and-period-doubling-bifurcations-in-an-epidemic-model.pdf

Ayoubkhani D., Khunti K., Nafilyan V., Maddox T., Humberstone B., Diamond I. & Banerjee A. (2021) 'Epidemiology of post-COVID syndrome following hospitalisation with coronavirus: a retrospective cohort study' medRxiv, 15 January 2021, at https://www.medrxiv.org/content/10.1101/2021.01.15.21249885v1

Baker M.G., Kvalsvig A. & Verrall A.J. (2020) 'New Zealand's COVID-19 elimination strategy' Med J Aust 213, 5 (September 2020) 198-200, at https://www.mja.com.au/journal/2020/213/5/new-zealands-covid-19-elimination-strategy

Barrett M. (2020) 'What have Norway, Finland and Denmark got right on Covid-19?' New Statesman, 22 December 2020, at https://www.newstatesman.com/world/europe/2020/12/what-have-norway-finland-and-denmark-got-right-covid-19

BBC (2020) 'Covid-19: China's Qingdao to test nine million in five days' BBC News, 12 October 2020, at https://www.bbc.com/news/world-asia-54504785

von Bertalanffy L. (1968) 'General System Theory: Foundations, Development, Applications' George Braziller, 1968

Bolla P. & Sarl A. (2020) 'Using Discrete Event Simulation model to simulate COVID19 epidemics in Switzerland' ALASS, August 2020, at 14/7/21https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7341703/

Brailsford S., Churilov L. & Dangerfield B. (eds.) (2014) 'Discrete-Event Simulation and System Dynamics for Management Decision Making' Wiley, 2014

Brailsford S. & Hilton NB. (2000) 'A Comparison of Discrete Event Simulation and System Dynamics for Modelling Healthcare Systems' Proc. ORAHS 2000, Glasgow, pp. 18–39, at https://eprints.soton.ac.uk/35689/1/glasgow_paper.pdf

Brauer F. (2008) 'Compartmental Models in Epidemiology' Springer, Lecture Notes in Mathematics, volume 1945, pp. 19-79, at https://link.springer.com/content/pdf/10.1007%2F978-3-540-78911-6_2.pdf

Budd J., Miller B.S., Manning E.M., et al. (2020) 'Digital technologies in the public-health response to COVID-19' Nature Medicine 26 (August 2020) 1183–1192, at https://www.nature.com/articles/s41591-020-1011-4

CDC (2020) 'Legal Authorities for Isolation and Quarantine' Center for Disease Control, Washington DC, rev. February 2020, at https://www.cdc.gov/quarantine/aboutlawsregulationsquarantineisolation.html

Chi Y.-L., Chalkidou K. & Walker D. (2020) 'What Can Policymakers Learn about COVID-19 from Looking at Different Model Estimates?' Center for Global Development, June 2020, at https://www.cgdev.org/blog/what-can-policymakers-learn-about-covid-19-looking-different-model-estimates

Clarke R. (2021) 'Simulation Modelling for Public Health Management during the COVID-19 Pandemic' Xamax Consultancy Pty Ltd, January 2021, at at http://rogerclarke.com/EC/CVMB.html

Currie C.S.M., Fowler J.W., Kotiadis K., Monks T., Onggo B.S., Robertson D.A. & Tako A.A. (2020) 'How simulation modelling can help reduce the impact of COVID-19' Journal of Simulation 14:2 (2020) 83-97, DOI: 10.1080/17477778.2020.1751570

Econ (2020) 'Tracking the coronavirus across Europe' The Economist, 3 July 2020, at https://www.economist.com/graphic-detail/2020/07/03/tracking-the-coronavirus-across-europe

Gault F.D., Hamilton K.E., Hoffman R.B. & McInnis B.C. (1987) 'The design approach to socio-economic modelling' Futures 19, 1 (February 1987) 3-25, at https://www.peakoil.net/files/THE_DESIGN_APPROACH_TO_SOCIO_ECONOMIC_MODELLING.pdf

Giordano G., Blanchini F., Bruno R., Colaneri P., Di Filippo A., Di Matteo A. & Colaneri M. (2020) 'Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy' Nature Medicine 26 (April 2020) 855–860, at https://www.nature.com/articles/s41591-020-0883-7

Gn (2020a) 'Melbourne Covid lockdown rules and coronavirus restrictions explained' The Guardian, 9 November 2020, at https://www.theguardian.com/australia-news/2020/nov/09/melbourne-restrictions-victoria-lockdown-rules-covid-19-metropolitan-metro-explained-what-you-need-to-know

Gn (2020b) 'Our key findings about US healthcare worker deaths to date' The Guardian, 23 Dec 2020, at https://www.theguardian.com/us-news/ng-interactive/2020/dec/22/lost-on-the-frontline-our-findings-to-date

Gregor S. & Hevner A. (2013) 'Positioning Design Science Research for Maximum Impact' MIS Quarterly 37, 2 (June 2013) 337-355, at https://ai.arizona.edu/sites/ai/files/MIS611D/gregor-2013-positioning-presenting-design-science-research.pdf

Gumel A.B., McCluskey C.C. & Watmough J. (2006) 'An SVEIR Model for Assessing Potential Impact of an Imperfect Anti-SARS Vaccine' Mathematical Biosciences And Engineering 3, 3 (July 2006) 485–512

ICAO (2020) 'Public Health Risk Mitigation Measures' International Civil Aviation Organisation, 2020, at https://www.icao.int/covid/cart/Pages/Public-Health-Risk-Mitigation Measures.aspx

Jalayer M., Orsenigo C. & Vercellis C. (2020) 'CoV-ABM: A stochastic discrete-event agent-based framework to simulate spatiotemporal dynamics of COVID-19' arXiv (Jul 2020), at 14/7/21https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7341703/

Jayadevan R. (2020) 'The 6 most successful anti-Covid strategies so far, and the reasons why they worked' Onmanorama, 3 June 2020, at https://www.onmanorama.com/lifestyle/health/2020/06/03/most-successful-anti-covid-strategies-reasons-why-they-worked.html

Jefferies S., French N., Gilkison C., Graham G., Hope V., Marshall J., McElnay C., McNeill A., Muellner P., Paine S., Prasad N., Scott J., Sherwood J., Yang L. & Priest P. (2020) 'COVID-

19 in New Zealand and the impact of the national response: a descriptive epidemiological study'
Lancet Public Health 5 (October 2020) e612–23, at
https://www.thelancet.com/action/showPdf?pii=S2468-2667%2820%2930225-5

Kermack W.O. & McKendrick A.G. (1927) 'A Contribution to the Mathematical Theory of Epidemics'
Proc. Roy. Soc. Lond. A 115, 700-721, 1927, at
http://www.math.utah.edu/~bkohler/Journalclub/kermack1927.pdf

Minghella D. (2020) 'The UK and Covid-19: the tragedy of the road not taken' New Statesman, 22
December 2020, at
https://www.newstatesman.com/politics/uk/2020/12/uk-and-covid-19-tragedy-road-not-
taken

Moss R., Wood J., Brown D., Shearer F., Black A.J., Cheng A.C., McCaw J.M. & McVernon J. (2020)
'Modelling the impact of COVID-19 in Australia to inform transmission reducing measures
and health system preparedness' medRxiv, 11 April 220, at
http://dx.doi.org/10.1101/2020.04.07.20056184

Niederman F. & March S. (2012) 'Design Science and the Accumulation of Knowledge in the
Information Systems Discipline' ACM Transactions on MIS 3, 1 (2012), 1

Peffers K., Tuunanen T., Rothenberger M. & Chatterjee S. (2007) 'A design science research
methodology for information systems research' Journal of Management Information Systems,
24, 3 (Winter 2007) 45-77

Nikulin I.I., Woolverton K.E. & Withington M.J. (2020) 'Using Dynamic Models and Empirical
COVID-19 Data to Showcase Effective Pandemic Prevention Measures' Worcester
Polytechnic Institute, at
https://digitalcommons.wpi.edu/iqp-all/5759

Novick L.F. & Mays G.P. (eds.) (2008) 'Public Health Administration: Principles for Population-based
Management' Jones & Bartlett, 2008

Novick L.F. & Morrow C.B. (2008) 'A Framework for Public Health Administration and Practice' Ch.
2 in Novick & Mays (2008), pp. 35-68

OxCGRT (2020) 'Codebook' Oxford Covid-19 Government Response Tracker, v.2.6, 09 December
2020, at
https://github.com/OxCGRT/covid-policy-
tracker/blob/master/documentation/codebook.md

Peterson M. (2009) 'An Introduction to Decision Theory' Cambridge University Press, 2009

Rhodes T., Lancaster K., Lees S. & Parker M. (2020) 'Modelling the pandemic: attuning models to
their contexts' BMJ Global Health (June 2020), doi:10.1136/ bmjgh-2020-002914, at
https://gh.bmj.com/content/bmjgh/5/6/e002914.full.pdf

Schipper H. (2020) 'How pandemic modelling failed policy-makers, and how to do better' Macdonald-
Laurier Institute, July 2020, at
https://macdonaldlaurier.ca/files/pdf/July_Commentary_Schipper_FWeb.pdf

Siebers P.O., Macal C.M., Garnett J., Buxton D. & Pidd M. (2010) 'Discrete-Event Simulation is Dead,
Long Live Agent-Based Simulation!' Journal of Simulation 4, 3 (2010) 204-210, at
https://link.springer.com/content/pdf/10.1057/jos.2010.14.pdf

Silva C.L.S, Batista P.V.C., Lima H.S., Alves M.A., Guimarães F.G. & Silva R.C.P. (2020) 'COVID-
ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects
of social distancing interventions' Chaos, Solitons & Fractals 139 (October 2020) 110088, at
https://doi.org/10.1016/j.chaos.2020.110088

Sprague R.H. (1980) 'A Framework for the Development of Decision Support Systems' MIS Qtly 4,
4 (December 1980) 1-26

SWPRS (2020) 'Post-Acute Covid ('Long Covid')' Swiss Policy Research, August 2020, at
https://swprs.org/post-acute-covid-long-covid/

Thomas O., Hagen S., Frank U., Recker J., Wessel L., Kammler F., Zarvic N. & Timm I. (2020) 'Global
Crises and the Role of BISE' Business & Information Systems Engineering 62, 4 (June 2020)
385–396, at https://link.springer.com/article/10.1007/s12599-020-00657-w

Trang S., Trenz M., Weiger W.H., Tarafdar M. & Cheung C.M.K. (2020) 'One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps' Euro. J. of Infor. Syst. 29, 4 (Jun 2020) 415-428, at
https://orsociety.tandfonline.com/doi/full/10.1080/0960085X.2020.1784046

Vroom V.H. & Jago A.G. (1988) ''The New Leadership: Managing participation in organizations' Prentice Hall, 1988

Vroom V.H. & Yetton P.W. (1973) 'Leadership and Decision Making' University of Pittsburgh Press, 1973

Weitz J.S. & Dushoff J. (2015) 'Modeling Post-death Transmission of Ebola: Challenges for Inference and Opportunities for Control' Sci Rep 5, 8751 (March 2015), at
https://www.nature.com/articles/srep08751

Whitworth J. (2020) 'Lessons from around the world on fighting COVID's second wave' The Conversation, 21 November 2020, at
https://theconversation.com/lessons-from-around-the-world-on-fighting-covids-second-wave-150432

WHO (2020a) 'Criteria for releasing COVID-19 patients from isolation: Scientific
Brief' 17 June 2020, at
https://www.who.int/news-room/commentaries/detail/criteria-for-releasing-covid-19-patients-from-isolation

WHO (2020b) 'Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations' World Health Organisation, Scientific brief, 29 March 2020, updated 9 July 2020, at https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations

WHO (2020c) ' Transmission of SARS-CoV-2: implications for infection prevention precautions: Scientific Brief' World Health Organisation, Interim guidance, July 2020, at
https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions#:~:text='The%20incubation%20period%20of%20COVID,to%20a%20confirmed%20case.

WHO (2020d) 'Considerations in adjusting public health and social measures in the context of COVID-19' World Health Organisation, Interim guidance, April 2020, rev. November 2020, at
https://www.who.int/publications/i/item/considerations-in-adjusting-public-health-and-social-measures-in-the-context-of-covid-19-interim-guidance

WOM (2020) 'Reported Cases and Deaths by Country, Territory, or Conveyance' Worldometer, December 2020, at
https://www.worldometers.info/coronavirus/#countries

Wood R.M., McWilliams C.J., Thomas M.J., Bourdeaux C.P. & Vasilakis C. (2020) 'COVID-19 scenario modelling for the mitigation of capacity-dependent deaths in intensive care' Health Care Manag Sci. 2020 8 (Jul 2020) 1–10, doi: 10.1007/s10729-020-09511-7, at
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7341703/

Zhang X. (2018) 'Application of discrete event simulation in health care: a systematic review' BMC Health Services Research volume 18, Article number: 687 (2018), at
https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-018-3456-4

Factors Contributing to the Business Digital Divide: A Systematic Literature Review

Ali Acilar, Dag Håkon Olsen, Niels Frederik Garmann-Johnsen & Tom Roar Eikebrokk

University of Agder, Faculty of Social Science, Kristiansand, Norway; e-mail: ali.acilar@uia.no, dag.h.olsen@uia.no, niels.f.garmann-johnsen@uia.no, tom.eikebrokk@uia.no

Abstract The main aim of this study is to review the literature relating to the factors that contribute to the business digital divide. A systematic literature review was conducted using two databases (Scopus and Web of Science). A total of 28 articles were selected and analyzed. The selected studies are conducted in various developing and developed countries, including all firm sizes and different sectors, and cover several different digital technologies. Identified factors determining the business digital divide are categorized as technological, organizational, and environmental factors. The discussion and the potentials for further research are also presented.

1   Introduction

The rapid and continuous developments of ICTs facilitate access and process data and improve the inter and intra-organizational integration of companies, but at the same time, these technological developments bring a new type of exclusion, the digital divide (Souza, Siqueira, & Reinhard, 2017). A significant number of businesses, especially SMEs, tend to be on the wrong side of the digital divide, and therefore do not benefit from the potential advantages of ICTs. Even though digitalization provides new opportunities for SMEs to benefit from the global economy, significant numbers of SMEs lag behind in the digital transition (North, Aramburu, & Lorenzo, 2019; OECD, 2017).

The digital divide can be defined as "the gap between individuals, households, businesses and geographic areas at different socio-economic levels with regard both to their opportunities to access ICTs and to their use of the Internet for a wide variety of activities" (OECD, 2001, p. 5). The digital divide can emerge from individual, organizational, and global levels (Dewan & Riggins, 2005). Unequal access and use of ICT are the main issues of the digital divide. Castells (2002, p. 270) describes the digital divide as "the divide created between those individuals, firms, institutions, regions, and societies that have the material and cultural conditions to operate in the digital world, and those who cannot, or cannot adapt to the speed of change." As among people, the digital divide also exists among businesses and refers to ICT access and the ability of appropriate use of the technology (Wielicki & Arendt, 2010). In addition to preventing access to ICT, the digital divide prevents commercial applications of these technologies, such as e-business (Di. Gregorio, Kassicieh, & De Gouvea Neto, 2005).

Several academic disciplines, from sociology and political science to business and information systems, have been involved in research about the digital divide; and most of these research studies focus on the individual or societal level (Wielicki & Arendt, 2010). The business digital divide is not discussed in the literature as much as the digital divide among people or organizations (Souza et al., 2017). We focus on the digital divide among businesses in this study. It is important to understand the business digital divide since it significantly affects how firms compete in the global market, how they communicate with their customers and business partners, and how they formulate their strategies for e-commerce (Dewan & Riggins, 2005; Wielicki &

A. Acilar, D. Håkon Olsen, N. Frederik Garmann-Johnsen & T. Roar Eikebrokk:
Factors Contributing to the Business Digital Divide: A Systematic Literature Review

31

Arendt, 2010). This study systematically reviews the literature with the aim of understanding the factors contributing to the digital divide among businesses. The literature review was driven by the following research question:

What are the determinant factors of the digital divide among businesses?

2   Methodology

In this study, a systematic literature review was conducted. The systematic literature review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) approach (Moher, Liberati, Tetzlaff, Altman, & Group, 2009). PRISMA is well accepted and used in a broad range of academic disciplines in the literature.

2.1   Inclusion and Exclusion Criteria

The search process was conducted using two scientific databases: Scopus and Web of Science. These two databases are "two world-leading and competing citation databases" (Zhu & Liu, 2020). We conducted the search with the following keywords: ("digital divide" OR "digital gap") AND (busines* OR firm* OR compan* OR corporate OR corporation* OR "small and medium size* enterpris*" OR SME* OR enterpris*) in "title, abstract, keywords" search fields. After the initial search, search results were restricted to journal articles from 2000 to 2019 in the English language for both databases. Only journal articles were included in this literature review.

2.2   Data Collection

The searches of the two databases resulted in 712 records. After 155 duplicate articles were removed, 557 articles remained for further screening. At this stage of the study, articles were excluded on the basis of irrelevant titles or abstracts. After the title and abstract screening process, 71 articles were selected for further full-text analysis. Nine articles could not be obtained from the databases. A total of 62 articles were accessed for full-text screening. Among them, two articles were excluded because they were written in Spanish. Even though database searches were limited based on language, these articles were listed by databases. Twenty-eight articles were

selected after the full-text screening. In order to code the selected articles, a spreadsheet was created. Full-text articles were excluded, with the following reasons: theoretical, not empirical, data collection methods, and out of focus of this study. The selected articles were coded with the following data: authors' names, article title, publication year, source title, technology, sample country, data source, data collection method, sample size, firm size, sector, methodology, and determinant factors. The steps of the systematic literature review are shown in Figure 1.
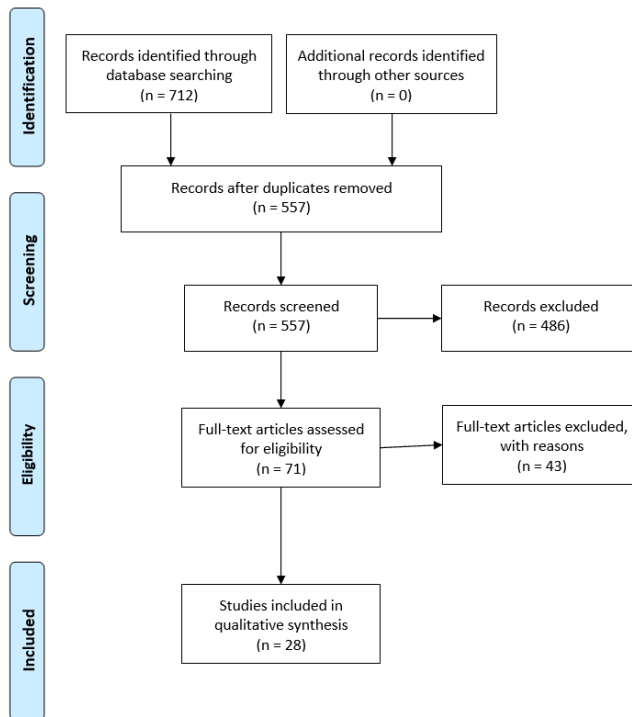
Figure 1: Flow diagram for selection of articles, based on PRISMA.

2.3 Systematic Literature Review Results

The selected studies in this literature review have been published using samples from countries around the world, including six continents, but mainly from Europe (17 studies) (Table 1).

Table 1: Selected articles, country of sample, technology

Article	no.	Country of sample	Technology
Duncombe and Heeks (2002)	A1	Botswana	ICT
Moodley (2003)	A2	South Africa	B2B e-commerce
Forman (2005)	A3	USA	Internet
Gengatharen and Standing (2005)	A4	Australia	e-marketplaces
Sun and Wang (2005)	A5	China	Internet access & use
Arbore and Ordanini (2006)	A6	Italy	Broadband
Hinson and Sorensen (2006)	A7	Ghana	E-business
Labrianidis and Kalogeressis (2006)	A8	Europe	A list of ICTs
Pighin and Marzona (2008)	A9	Italy	ICT use and process automation
Atzeni and Carboni (2008)	A10	Italy	ICT
Billon, Ezcurra, and Lera-López (2009)	A11	Europe	Website
Karen L. Middleton and Chambers (2010)	A12	USA	Wifi
Galve-Górriz and Gargallo Castel (2010)	A13	Spain	ICT
Rodríguez-Ardura and Meseguer-Artola (2010)	A14	Spain	E-commerce
Wielicki and Arendt (2010)	A15	USA, Spain, Portugal, Poland	ICT-based solutions
K. L. Middleton and Byus (2011)	A16	USA	ICT
Chang, Wu, and Cho (2011)	A17	Taiwan	ICT
Oni (2013)	A18	Nigeria	Applic. of ICT tools
Bach (2014)	A19	Europe	ICT indicators
Oliveira and Dhillon (2015)	A20	Europe	B2B e-commerce
Billon, Lera-Lopez, and Marco (2016)	A21	Europe	ICT
Doherty, Ramsey, Harrigan, and Ibbotson (2016)	A22	Ireland	Broadband technologies
Billon, Marco, and Lera-Lopez (2017a)	A23	Europe	ICT
Billon, Marco, and Lera-Lopez (2017b)	A24	Europe	ICT
Ayinla and Adamu (2018)	A25	Global	BIM technology
Ruiz-Rodríguez, Lucendo-Monedero, and González-Relaño (2018)	A26	Europe	ICT
Jordá Borrell, López Otero, and Contreras Cabrera (2018)	A27	Global	ICT
Bowen and Morris (2019)	A28	United Kingdom	Broadband, website, social media

Data in two studies are collected on a global scale. Various technologies are subject to the articles as indicators of the digital divide, such as the Internet, broadband, e-business, e-marketplace, website, social media, wifi, e-commerce, and B2B e-commerce. Some studies did not indicate the specific technology; instead, they used

the general term ICT. The sample sizes of the selected studies vary from 5 to more than 40,000 enterprises. Firm sizes in the studies are also various. Samples include enterprises with different sizes, from micro-enterprises to large-size enterprises, in sectors including manufacturing, finance, service, construction, and food. Four studies (A1, A5, A15, and A20) have samples in more than five sectors. Based on the level of the study, there are two main groups of articles: country or region level and firm-level articles. Country-level articles (A8, A11, A19, A21, A23, A24, A27) mainly used secondary data and applied econometric statistical analyses. Data of the selected studies come mainly from surveys. Almost half of the studies used secondary data. The selected studies used various quantitative methods for analyzing their data, such as the Chi-square test, regression analysis, correlation analysis, ANOVA, MANOVA, factor analysis, cluster analysis, spatial data analysis, logit analysis, structural equation modeling.

It is found that there are 54 different factors identified in the selected literature. Based on the selected articles in this literature review, we categorized the factors determining the digital divide as technological (Table 3), organizational (Table 4), and environmental factors (Table 5), using Technology-Organizational-Environmental (TOE) framework (Tornatzky & Fleischer, 1990). Table 3 presents factors related to technology. The most common technological factors reported in the selected articles are identified as perceived usefulness, cost, degree of ICT readiness, and relative advantage.

Table 3: Technological Factors

Factors	Article No.
Relative advantage	A4, A22
Perceived benefits	A4
Perceived usefulness	A4, A7, A22
Perceived impact on the image of the firm	A22
Perceived need	A1
Cost	A1, A2, A25
Digital awareness	A18
The degree of ICT readiness	A14, A15, A20
Prior IT investment	A3
Technology/interoperability	A25
Technology integration	A20
Innovation target (technology) to be used	A9

A. Acilar, D. Håkon Olsen, N. Frederik Garmann-Johnsen & T. Roar Eikebrokk:
Factors Contributing to the Business Digital Divide: A Systematic Literature Review

35

Table 4 shows organizational factors. The most common organizational factor is firm size. As it is directly related to both firms' financial ability to acquire and human resources to use, firm size is a prominent factor in adopting the technology. Small businesses with limited financial and human resources struggle with following technological developments. In addition to firm size, several other organizational factors are identified in the articles, such as factors related to human resources (employees´ education, expertise, training, investment per employee), owner's characteristics, internalization, organizational culture, and firm's age.

Table 4: Organizational Factors

Factors	Article No.
Firm size	A3, A5, A6, A8, A20
Firm's age	A5, A10
Organizational Culture	A9, A25
Ethnicity	A12, A16
Owners age	A8, A12
Owner innovativeness	A4
Owners Education level	A8
Outsourcing strategy	A6
Financial constraints	A10
Lack of resources	A28
Reorganization	A10
Internationalization	A14, A28
R&D (Innovation capacity)	A10, A23
Employees´ education	A13, A20, A24
Labor composition	A10
Skills and capabilities	A2, A25
Training	A13, A25
Individual growth ability of employees	A9
Geographic dispersion of employees	A3
Perceived obstacles	A20

Table 5 presents the identified environmental factors in the selected articles, which are related to the environment of the firm. The most commonly reported environmental factor is location. After that, sector, customers, firms' pressure, and financial support are other significant environmental factors reported by the researchers. The location of the firm is an important factor in the adoption of technology. The urban and rural divide still exists for businesses. Also, businesses in less developed countries or regions tend to be on the wrong side of the digital divide.

Table 5: Environmental Factors

Factors	Article No.
Location	A5, A6, A8, A21, A28
Sector	A5, A8, A21, A24
Sectoral composition	A11
Customers	A14, A22, A25
Pressure of firms	A3, A14, A20
Network intensity	A8
Trading partner collaboration	A20
Financial support, subsidies, government support	A8, A10, A17
Government policy	A2
Legal requirements	A25
REM ownership structure and governance	A4
Critical mass	A4
Infrastructure	A2
Innovation performance of the country	A19
GDP per capita	A11
Fiscal decentralization	A21
Population density	A11
The extent of countries' globalization	A27
Digital development of the country	A26
Accessibility to ICT capabilities of the country	A27
Technological readiness of market forces	A14
The educational level of the region	A11, A21

3   Discussion

This study has presented the results of a systematic literature review of studies on the digital divide among businesses, published between 2000 and 2019. The business digital divide phenomenon has been investigated in developing and developed countries, particularly in Europe. There is a relatively small number of studies in developing economies. The divide has been approached with different digital technologies, involving adoption and use. The studies suggest that the digital divide exists among businesses in different sizes, sectors, and countries. Identified factors determining the business digital divide are categorized as technological, organizational, and environmental factors. The most commonly reported factors in the articles are identified as firm size, human resources, location, sector, customers, the pressure of firms, financial support, perceived usefulness, cost, and the degree of ICT readiness.

A. Acilar, D. Håkon Olsen, N. Frederik Garmann-Johnsen & T. Roar Eikebrokk:
Factors Contributing to the Business Digital Divide: A Systematic Literature Review

37

The literature has increasingly emphasized digitalization as an important vehicle for generating value from information technology for society, industry, and enterprises (Reis, Amorim, Melão, Cohen, & Rodrigues, 2019). In order to significantly benefit from digitalization, extensive changes are required in the organization. Digitalization implies significant changes for businesses, including strategy and business models, internal and external processes, organizational culture, etc. (Parviainen, Tihinen, Kääriäinen, & Teppola, 2017), a digital transformation. We found that the literature on the digital divide has barely addressed the digital transformation issues.

The size and pace of the digital transformation make investments in digitalization for businesses of all sizes and in all industries inevitable to ensure success and survival (Hossnofsky & Junge, 2019). "Digitalisation is feared as a source of disruption, with the risk that only a few firm will emerge as winners while many firms and workers lose out, leading to a more polarised economic structure" (Veugelers, Rückert, & Weiss, 2019). Therefore, digitalization involves internal and external challenges for businesses, particularly SMEs. With limited financial and human resources, digitalization is a real threat for many SMEs and can widen the digital divide between SMEs and large businesses. Firms need dynamic capabilities to cope with the digital transformation and to adapt to the changing environment. However, it is a challenge for businesses to design mechanisms that enable repeatable, continuous adaptation (Vial, 2019). Besides, it is challenging for businesses to grasp how digitalization can be leveraged to transform their business models to achieve sustainable benefits (Parida, Sjödin, & Reim, 2019). Businesses need to understand how they can continuously derive and leverage value through developing their IT capabilities (Eikebrokk & Olsen, 2007). We, therefore, argue that further studies should explore the digitalization divide, focusing on factors causing the divide in digitalization processes and digitalization capabilities.

The business digital divide studies mainly focus on the adoption and use of ICTs, and there are not many studies about outcomes of ICT usage (third-level digital divide) in businesses. For example, there is not much evidence that the digitalization of the business causes a significant productivity boost (Veugelers et al., 2019). Future research can also aim to investigate the divide between businesses in terms of outcomes and benefits of using ICTs.

4   Conclusion

The digital divide is a global phenomenon that affects people, organizations, and countries around the world. This study provides a systematic literature review about the factors that contribute to the business digital divide. The review is conducted by using two databases: Scopus and Web of Science. Twenty-eight journal articles published between 2000 and 2019 made up the sample of this study and were analyzed in the review. We investigated the characteristics of the business digital divide research and summarized the research distribution in terms of sample characteristics, methodological approaches, and the digital divide determinants. The digital divide exists among businesses in different sizes, sectors, and countries. Identified factors determining the digital divide are categorized as technological, organizational, and environmental factors.

The main limitations of the study can be summarized as follows: Only two databases (Scopus and Web of Science) were used in this study. This review is based on only journal articles written in English. There are certainly other types of publications and studies in different languages, which address the business digital divide. Lastly, it is possible to have different search results using different search strings.

References

Arbore, A., & Ordanini, A. (2006). Broadband divide among SMEs: The role of size, location and outsourcing strategies. International Small Business Journal, 24(1), 83-99.

Atzeni, G. E., & Carboni, O. A. (2008). The effects of grant policy on technology investment in Italy. Journal of Policy Modeling, 30(3), 381-399.

Ayinla, K. O., & Adamu, Z. (2018). Bridging the digital divide gap in BIM technology adoption. Engineering, Construction and Architectural Management, 25(10), 1398-1416.

Bach, M. P. (2014). Exploring information and communications technology adoption in enterprises and its impact on innovation performance of European countries. Ekonomicky casopis, 62(04), 335-362.

Billon, M., Ezcurra, R., & Lera-López, F. (2009). Spatial effects in website adoption by firms in European regions. Growth and change, 40(1), 54-84.

Billon, M., Lera-Lopez, F., & Marco, R. (2016). ICT use by households and firms in the EU: links and determinants from a multivariate perspective. Review of World Economics, 152(4), 629-654.

Billon, M., Marco, R., & Lera-Lopez, F. (2017a). Innovation and ICT use by firms and households in the EU: A multivariate analysis of regional disparities. Information Technology and People, 30(2), 424-448.

Billon, M., Marco, R., & Lera-Lopez, F. (2017b). Innovation and ICT use in the EU: an analysis of regional drivers. Empirical Economics, 53(3), 1083-1108.

A. Acilar, D. Håkon Olsen, N. Frederik Garmann-Johnsen & T. Roar Eikebrokk:
Factors Contributing to the Business Digital Divide: A Systematic Literature Review

39

Bowen, R., & Morris, W. (2019). The digital divide: Implications for agribusiness and entrepreneurship. Lessons from Wales. Journal of Rural Studies, 72, 75-84.

Castells, M. (2002). The Internet galaxy: Reflections on the Internet, business, and society: Oxford University Press on Demand.

Chang, S.-I., Wu, H.-C., & Cho, C.-M. (2011). The development of digital divide assessment mechanism for SMEs: A perspective from the Taiwan manufacturing industry. Journal of Global Information Technology Management, 14(1), 6-34.

Dewan, S., & Riggins, F. J. (2005). The digital divide: Current and future research directions. Journal of the Association for information systems, 6(12), 298-337.

Di. Gregorio, D., Kassicieh, S. K., & De Gouvea Neto, R. (2005). Drivers of e-business activity in developed and emerging markets. IEEE Transactions on Engineering Management, 52(2), 155-166.

Doherty, E., Ramsey, E., Harrigan, P., & Ibbotson, P. (2016). Impact of broadband Internet technologies on business performance of Irish SMEs. Strategic Change, 25(6), 693-716.

Duncombe, R., & Heeks, R. (2002). Enterprise across the digital divide: information systems and rural microenterprise in Botswana. Journal of International Development, 14(1), 61-74.

Eikebrokk, T. R., & Olsen, D. H. (2007). An empirical investigation of competency factors affecting e-business success in European SMEs. Information & Management, 44(4), 364-383.

Forman, C. (2005). The Corporate Digital Divide: Determinants of Internet Adoption. Management Science, 51(4), 641-654.

Galve-Górriz, C., & Gargallo Castel, A. (2010). The relationship between human resources and information and communication technologies: Spanish firm-level evidence. Journal of theoretical and applied electronic commerce research, 5(1), 11-24.

Gengatharen, D. E., & Standing, C. (2005). A framework to assess the factors affecting success or failure of the implementation of government-supported regional e-marketplaces for SMEs. European Journal of Information Systems, 14(4), 417-433.

Hinson, R., & Sorensen, O. (2006). E-business and small Ghanaian exporters: Preliminary micro firm explorations in the light of a digital divide. Online Information Review, 30(2), 116-138.

Hossnofsky, V., & Junge, S. (2019). Does the market reward digitalization efforts? Evidence from securities analysts' investment recommendations. Journal of Business Economics, 89(8), 965-994.

Jordá Borrell, R. M., López Otero, J., & Contreras Cabrera, G. A. (2018). Factors that influence ICT adoption at the country level. PLS-SEM Modelling. Revista De Economia Mundial, 50, 153-

Labrianidis, L., & Kalogeressis, T. (2006). The digital divide in Europe's rural enterprises. European Planning Studies, 14(1), 23-39.

Middleton, K. L., & Byus, K. (2011). Information and communications technology adoption and use in small and medium businesses: The influence of Hispanic ethnicity. Management Research Review, 34(1), 98-110.

Middleton, K. L., & Chambers, V. (2010). Approaching digital equity: is wifi the new leveler? Information Technology & People, 23(1), 4-22.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS med, 6(7), 264-270.

Moodley, S. (2003). Whither business-to-business electronic commerce in developing economies? The case of the South African manufacturing sector. Information Technology for Development, 10(1), 25-40.

North, K., Aramburu, N., & Lorenzo, O. J. (2019). Promoting digitally enabled growth in SMEs: a framework proposal. Journal of Enterprise Information Management, 33(1), 238-262.

OECD. (2001). Understanding the Digital Divide.

OECD. (2017). Enhancing the contributions of SMEs in a global and digitalised economy.

Oliveira, T., & Dhillon, G. (2015). From Adoption to Routinization of B2B e-Commerce. Journal of Global Information Management, 23(1), 24-43.

Oni, A. O. (2013). Digital divide - a challenge to the real estate practice in Nigeria? Property Management, 31(1), 22-38.

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

Parviainen, P., Tihinen, M., Kääriäinen, J., & Teppola, S. (2017). Tackling the digitalization challenge: how to benefit from digitalization in practice. International journal of information systems and project management, 5(1), 63-77.

Pighin, M., & Marzona, A. (2008). An Uncontrolled Item of Enterprise ICT Innovation: the High Level Digital Divide. Journal of Information and Organizational Sciences, 32(1), 51-66.

Reis, J., Amorim, M., Melão, N., Cohen, Y., & Rodrigues, M. (2019). Digitalization: a literature review and research agenda. Paper presented at the International Joint Conference on Industrial Engineering and Operations Management.

Rodríguez-Ardura, I., & Meseguer-Artola, A. (2010). Toward a Longitudinal Model of e-Commerce: Environmental, Technological, and Organizational Drivers of B2C Adoption. The Information Society, 26(3), 209-227.

Ruiz-Rodríguez, F., Lucendo-Monedero, A. L., & González-Relaño, R. (2018). Measurement and characterisation of the Digital Divide of Spanish regions at enterprise level. A comparative analysis with the European context. Telecommunications Policy, 42(3), 187-211.

Souza, C. A. d., Siqueira, É. S., & Reinhard, N. (2017). Digital divide of small and medium-sized enterprises: An analysis of influencing factors using The TOE theory. RAM. Revista de Administração Mackenzie, 18(2), 15-48.

Sun, Y., & Wang, H. (2005). Does Internet access matter for rural industry? A case study of Jiangsu, China. Journal of Rural Studies, 21(2), 247-258.

Tornatzky, L. G., & Fleischer, M. (1990). The Processes of Technological Innovation. Lexington, MA: Lexington Books.

Veugelers, R., Rückert, D., & Weiss, C. (2019). Bridging the divide: new evidence about firms and digitalisation. Bruegel Policy Contribution Issue# 17 December 2019.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), 118-144.

Wielicki, T., & Arendt, L. (2010). A knowledge-driven shift in perception of ICT implementation barriers: Comparative study of US and European SMEs. Journal of Information Science, 36(2), 162-174.

Zhu, J., & Liu, W. (2020). A tale of two databases: the use of Web of Science and Scopus in academic papers. Scientometrics, 123, 321-335.

PORT COMMUNITY SYSTEM BUSINESS MODELS

MARIJA JOVIĆ,[1] SAŠA AKSENTIJEVIĆ,[2] BORNA PLENTAJ[3] &
EDVARD TIJAN[1]

[1] University of Rijeka, Faculty of Maritime Studies, Rijeka, Croatia; e-mail:
jovic@pfri.hr, etijan@pfri.hr
[2] Aksentijević Forensics and Consulting, Viškovo, Croatia; e-mail:
sasa.aksentijevic@gmail.com
[3] Actual d.o.o., Žminj, Croatia, e-mail: bornnaplentaj7@gmail.com

Abstract Port Community Systems have become a staple
technological platform used to exchange information between
the public and private agents and entities involved in ship and
cargo services within seaports. In this paper, the theoretical
background of the Port Community System is provided,
emphasizing the importance of its implementation, and
stakeholder collaboration. Different models of introducing an
integrated Port Community System in seaports are analysed using
literature review and actual cases in some of the most prominent
seaports.

1 Introduction

Seaports, defined as a geographical location where cargo changes its transport mode (one of these being a seagoing vessel), are important drivers of the regional economy (Hintjens, Hassel, Vanelslander, & Voorde, 2020). The seaport's competitiveness depends not only on exceptional geographical position, related to closeness of important markets and connection to seaport hinterland, but also on cost, efficiency, reliability, accessibility, safety, as well as quality of various services it offers, including transportation services, auxiliary services and added value logistic services. Seaports have to continuously improve their operations, both commercial and administrative in order to stay competitive (Tijan, Jović, & Karanikić, 2019).

Numerous seaports have already designed and implemented the Port Community System (PCS). PCS allows the users to make service requests and input their information directly into the port's information system (Keceli, 2011), and enables the intelligent and protected exchange of information between involved public and private port users (Simoni, Schiavone, Risitano, Leone, & Chen, 2020). The higher the level of collaboration and integration between the port and supply chain stakeholders, the greater the sustainability of both the overall supply chain and the port (Tijan, Agatić, Jović, & Aksentijević, 2019). Seaport stakeholders have their own distinctive interests, which may minimize the ability to incorporate the PCS into seaport operations. Nonetheless, numerous seaports have recognized the benefits that PCS brings, and are utilizing it to assist with everyday operations.

There is no universal approach to PCS introduction and exploitation, or a universal applicable business model, due to the heterogenous nature of global seaports and their management. Furthermore, the majority of research in this area is focused on digitalization of processes and unification of underlying procedures and document flow, and not on relevant PCS introduction and exploitation models. By working not only on the theoretical aspects of PCS systems, but also in their envisaging and implementation, the authors have identified the lack of applicable knowledge in this area and therefore wanted to verify the pragmatic and empiric findings by cross checking them using scientific resources. To overcome this research gap, the authors have conducted the review of available literature and sources. The goal of the research is to analyze the various PCS business models, given the existence of various stakeholders who have their own particular interests and preferences. Given

the fact that the transparency and easy access to data are the basis for successful transport business, the research problem stems from unnecessary costs and lost time due to the outdated business procedures and inadequate execution and monitoring of business processes in transport, which can largely be remedied by introducing a PCS. This paper presents a review of research papers and other sources (such as official webpages of seaports, seaport stakeholders and maritime transport enterprises), ultimately providing a better understanding of PCS business models.

2   Theoretical background

Port Community System is the technological platform that enables networking between the public and private agents and entities involved in the ship and cargo services offered by ports (Caldeirinha, Felício, Salvador, Nabais, & Pinho, 2020), through a single point of data entry (Aloini, Benevento, Stefanini, & Zerbino, 2020).

Two main values are co-created by the interrelated organizations operating within a PCS: the movement of goods and human beings and enforcing the law, public order, and safety. (Nota, Bisogno & Saccomanno, 2018). Generally accepted guidelines for development of a Port Community System require that PCS is formed by the community for the community and that the community are, in general, shareholders of the PCS Operator. (European Port Community Systems Association, 2011) However, in real PCS scenarios, in many cases, the community are not participants or owners of the PCS Operator, instead, it is often mandated by the governmental body (for example, port authorities) or maintains a mixed management and ownership structure. This presents one of the major issues in PCS building and operations and makes it difficult for PCS to facilitate smooth flow of electronic data and reduce inefficiencies in port business processes. Therefore, the selection of a proper operating model is crucial for the success of every PCS project. While individual business information systems that relate to individual stakeholders process and store only data and messages that are relevant for them, PCS can exchange data that is useful for a wider number of users (Tijan, Aksentijević, & Čišić, 2014). PCS exists in a dynamic network consisting of a significant number of stakeholders (as shown in Figure 1 with different business processes, technologies and roles (Bezić, Tijan, & Aksentijević, 2011).
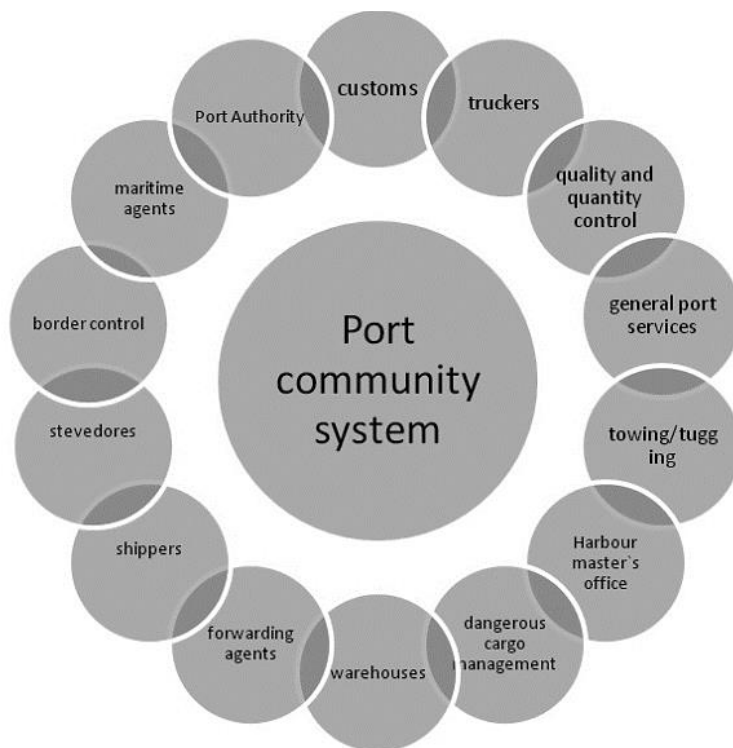
Figure 1: Port community system and involved stakeholders
Source: (Tijan, Kos, & Ogrizović, 2009)

PCS is largely based on a strong collaboration between all the involved public and private organizations (Baron & Mathieu, 2013), establishing a link between different types of technologies, processes, people, and standards (Rodon & Ramis-Pujol, 2006). Regional or even global PCSs might be designed (Jović, Tijan, Žgaljić, & Karanikić, 2020), helping to enhance the overall PCS performance in both local and foreign trade activities (Moros-Daza, Amaya-Mier, Garcia-Llinas, & Voß, 2019). In both cases a further standardization of interfaces and processes would be required. The requirements of and benefits for each company would have to be outlined and agreed on in advance (Treppte, 2011).

Bringing all users together, PCS enhances the efficiency of the physical flow of freight, drives economic growth, and as a secondary result, assists in reducing externalities such as pollution, congestion, and land use impacts (Irannezhad, Hickman, & Prato, 2017). According to (Zerbino, Aloini, Dulmin, & Mininno,

2019), one of the reasons for PCS development is the possibility to reduce the average time frame of port procedures, and to enhance information exchange, consequently improving overall port competitiveness.

Going beyond traditional function of PCS to share information, a PCS can offer modules to support a variety of activities (Baalen, Zuidwijk, & Nunen, 2009). The recent versions of PCS include the cloud services, which is becoming a significant factor in the historical development of information technology outsourcing (Johansson & Muhic, 2017).

Although a PCS connects multiple systems operated by a variety of organizations that make up a seaport community (IPCSA, 2020), it should be noted that for each port region, a PCS can take different forms in response to various physical, modal, jurisdictional, and operational characteristics (Tsamboulas, Moraiti, & Lekka, 2012). PCS functions may be divided into three categories: port management functions (documents provided to port authorities or terminal operators), customs functions (documents needed for customs clearance) and online platforms for electronic commerce between port users (Keceli, 2011).

3   Methodology

The literature review was conducted in order to research the theoretical foundations of models of port community systems. The authors opted to perform the search using only resources in English language and started with the inclusion criteria by using a combination of keyword "Port Community System models" and "Port Community System" (title, abstract and keywords). Web of Science, Google Scholar, ResearchGate databases were mainly used for this purpose. Due to the previously identified lack of the research dealing on this topic, the search for articles was not limited to a specific period, and mostly includes journal articles and conference papers. To ensure that possible useful findings from various fields were not excluded, the authors did not limit the queries to a specific field or index. Furthermore, due to aforementioned lack of research dealing with PCS models, the PCS models are further analysed by means of several real-life implementations such as: Port of Hamburg, The Port Authority of Valencia, Port of Rotterdam, The Antwerp PCS, etc. Using described methodology, a total of 36 resources have been

identified and used in the description of development and exploitation models of a PCS that follows.

4   Development and exploitation models of a Port Community System

Seaport stakeholders have their own preferences, which can decrease the willingness of certain members of the port community to incorporate PCS. The PCS should therefore enable the promotion of autonomy of all participants and at the same time include and support activities in various business processes in relation to seaports. In this respect, such a system does not only deal with internal needs of each individual company, but also with needs of other seaport stakeholders. In particular, the port authority attempts (or should attempt) to optimize the impact of the seaport's activities on the territory in terms of value added (local employment and incomes); on the other hand, the port operators should attempt to maximize the value for the final customer (De Martino, Errichiello, Marasco, & Morvillo, 2013).

Depending on a type of stakeholder's organization and its objectives, ownership model can be private, public or mixed public-private (PPP) (Marek, 2017). If the ownership model is of a private kind, the so-called bottom-up approach would be implemented in the system implementation. In this way, it is expected that the stakeholders (shipping companies, shipping agents, brokers, etc...) will support the work with the PCS since it is accepted by the operators themselves. Ports such as port of Singapore, Hamburg, Felixstowe belong to that kind of PCS model. For example, the Port Community System for the Port of Hamburg is operated by DAKOSY, one of the leading platform and software providers for logistics (IPCSA, 2021a). The PCS connects all stakeholders involved in cargo handling to perform fast, efficient and largely-automated processes in seaports and enables integrated intermodal hinterland handling of all modes of transport (Dakosy, 2021).

The top-down approach would be implemented if the ownership model is more similar to the public style, with a focus on port authorities and public bodies as the key stakeholders who determine the speed of implementation and set targets in the development of the PCS system (Marek, 2017). Ports such as Port of Valencia, Port of Rotterdam and Amsterdam belong to public PCS model.

The port authority plays an important role in implementing and creating the port development strategy and in coordinating the port community as a whole (João, Batista, Ayala Botto, & Cordón Lagares, 2018). The port authority is responsible for secure, sustainable and competitive port growth and may be a key factor in the implementation of the PCS (Tijan, Agatić, & Hlača, 2012). The implementation of the PCS may allow port authorities to coordinate port activities, monitor the activities of port operators and control port operations more easily (Carlan, Sys, & Vanelslander, 2016). For example, via Valenciaport PCS, the Port Authority of Valencia offers e-commerce solutions that make it easier for goods to move through the ports of Valencia, Sagunto and Gandía, adding a clearly perceptible value to the consumers and port users ("Port Authority of Valencia," 2021).

(Chandra & Hillegersberg, 2017) have conducted a Port of Rotterdam case study in which the importance of cooperation between port authorities and other stakeholders involved in the implementation of the PCS is visible. According to the study (Chandra & Hillegersberg, 2017), due to dissatisfaction with the Port of Rotterdam's information system, Port Infolink B.V. was established in 2002 (as a separated governance entity). Initiated by the Port of Rotterdam Authority, the pre-partnership cooperation process started by defining the most important issues that hinders the efficient flow of goods through the port. The Port Authority was the sole owner of Port Infolink, meaning that it was responsible for the initial investment in the development of the information system. This initiative included other stakeholders in the partnership program delivery phase (e.g. Customs). In early 2009, the next governance life cycle was marked by the merger of Port Infolink in Rotterdam and PortNET in Amsterdam, which provided the Ports of Rotterdam and Amsterdam with a single PCS (Chandra & Hillegersberg, 2017).

PPP is, in essence, a mixture of the two previously described ownership models. The aim of this ownership model is either to achieve complete acceptance of the PCS system or active role of private corporations in implementing the PCS system through a top-down approach (Marek, 2017). According to (Klievink, 2015), in *public-private collaboration* PCS design, data are handed over to the PCS but are still owned by the individual actors submitting the data. This allows government to access the data and allows the PCS to optimize port operations by enabling companies operating in the port data sharing without losing control. Ports such as Port of Barcelona and Antwerp belong to this type of PCS model. PORTIC is the

Port Community System operator in Barcelona and a private-public partnership between the Port Community of Barcelona, Port Authority of Barcelona, Financial Institutions (La Caixa, Banc Sabadell) and the Chamber of Commerce of Barcelona (IPCSA, 2021b). The Antwerp PCS is a cooperation between Antwerp Port Authority and Alfaport Antwerpen -Federation of Port Companies and Logistic Service Providers - private IT-sector (Descartes – Porthus) (Waterschoot, n.d.).

(Mendes Constante, 2019), outlines the features of business models based on combinations of PCS ownership and operational models. In the scenario where both the PCS ownership and the operational model are private, active engagement by the public sector is required in order to successfully implement complete integration and interoperability between all stakeholders involved. In the scenario where the PCS ownership is public but the operating model is private, private company operates the PCS on a commercial basis whereas public bodies play a crucial role in ensuring that services are provided fairly and neutrally to all stakeholders involved (Mendes Constante, 2019).

To summarize, the above points to the fact that when it comes to PCS business objectives, the main goal is to add value and improve the quality of port operations, logistics and the transport chain while at the same time reducing operating costs. It is also important to remember that during development it is extremely important to take care of the selection of PCS model because it will determine the specific financial model and goals that PCS as a project aspires to.

5   Discussion and conclusion

Seaports, as complex systems, are of vital importance for global trade activities because the most important international transport corridors and cargo flows pass through them and dictate global trends of economic development.

Daily port operations highly depend on information technologies and information systems. They have become irreplaceable elements in numerous seaports where they play an important role in port's overall business success. Information and communication systems such as PCS have become the staple technical ingredients used in optimal flow of information and provision of quality and efficient transport service and flexible and efficient functioning of port system as an important link in the transport chain.

Collaboration between stakeholders is a very important factor that enhances port system functioning. Utilizing coordination with other systems and technologies, they form an entity that significantly affects port system operations efficiency and coordination. Familiar expression stating that the chain is only as strong as its weakest link is certainly applicable to this concept too. If a port is recognizable on the global market, it attracts the largest ships and therefore the largest companies in international shipping industry. On the other hand, if the service it provides is not at an equally high level, or it is provided in a way that stakeholders offering port services are not interconnected and harmonized, the whole chain, including the port itself leave an impression of inconsistency.

The ownership and control over the PCS system are often overlooked parameters during the PCS inception phase. A PCS is a constantly developing system that needs to reflect every change in the port's environment, underlying technology, business processes, legal framework and all stakeholders. As it requires a significant coordination effort for proper functioning, it is very important to involve all stakeholders to provide a meaningful input to this process, reserve proper funding and ensure stakeholders' collaboration in order to achieve the goal of PCS's introduction.

The limitation of this research is primarily the fact that only English resources were used. PCS systems are adopted world-wide, and it is possible that the research base would be wider if other languages were included too, but that would lead the research outside of the applicable scope and format. Furthermore, PCS implementation is a highly operative endeavour, and many lessons learned are not published in a form of a scientific research, and therefore they cannot be used. Primary research hypothesis was confirmed, as not many quality resources are dealing with the selection of PCS model being a crucial factor in its successful implementation. Most researches are focused to project implementation phases, project management and encompassing all processes as success factors, but take PCS business model as something that is predetermined and not questioned. Additional authors' finding is that the selection of a proper PCS business model is a prerequisite for its successful implementation and operation, as only the appropriate PCS model can guarantee resource savings typically tied to PCS, as opposed to manual administrative processing or disjointed and heterogeneous port IT systems.

The implementation of a PCS helps improve the efficiency in communication among port community members, avoids duplicate data entry, optimizes flow and timely exchange of information, and increases protection from unauthorized access. It should enable electronic business operations which should result in a more efficient service, better mutual coordination, decrease in operative expenses and, finally, a more competitive port. Leading international seaports have recognized the importance and advantages of modern technologies in providing high quality services in ports.

The research can have many potential new venues and possibilities. PCS systems in the future will have to be highly flexible and interconnected, especially with introduction and absorption of novel technologies like Internet of Things, entire information platforms being delivered using cloud approach, active and passive tags and globally recognized cargo ledgers with distributed and transparent proof of authenticity. Integration of all these technologies will be a challenge for all PCS operators and provide new possibilities for research as those models that are successful now might not be suitable. One realistic possibility of a future research in this area might be analysis of impact of new information technologies on selected PCS business models.

Acknowledgements

References

Aloini, D., Benevento, E., Stefanini, A., & Zerbino, P. (2020). Process fragmentation and port performance: Merging SNA and text mining. International Journal of Information Management, 51(C).

Baalen, P. J. Van, Zuidwijk, R., & Nunen, J. van. (2009). Port Inter-Organizational Information Systems: Capabilities to Service Global Supply Chains. Foundations and Trends® in Technology, Information and Operations Management, 2(2–3), 81–241.

Baron, M., & Mathieu, H. (2013). PCS interoperability in Europe: a market for PCS operators? The International Journal of Logistics Management, 24(1), 117–129.

Bezić, H., Tijan, E., & Aksentijević, S. (2011). Port Community Systems - Economic Feasibility Evaluation. Ekonomski Vjesnik / Econviews – Review of Contemporary Entrepreneurship, Business, and Economic Issues, (2), 247–256.

Caldeirinha, V., Felício, J. A., Salvador, A. S., Nabais, J., & Pinho, T. (2020). The impact of port community systems (PCS) characteristics on performance. Research in Transportation Economics, 80(C).

Carlan, V., Sys, C., & Vanelslander, T. (2016). How port community systems can contribute to port competitiveness: Developing a cost-benefit framework. Research in Transportation Business & Management, 19, 51–64.

Chandra, D. R., & Hillegersberg, J. van. (2017). Governance lifecycles of inter-organizational collaboration: A case study of the Port of Rotterdam. Procedia Computer Science, 121, 656–663.

Dakosy. (2021). Port Community System. Retrieved February 3, 2021, from https://www.dakosy.de/en/solutions/cargo-communications/port-community-system

De Martino, M., Errichiello, L., Marasco, A., & Morvillo, A. (2013). Logistics innovation in Seaports: An inter-organizational perspective. Research in Transportation Business and Management, 8, 123–133. https://doi.org/10.1016/j.rtbm.2013.05.001

Hintjens, J., Hassel, E. van, Vanelslander, T., & Voorde, E. Van de. (2020). Port Cooperation and Bundling: A Way to Reduce the External Costs of Hinterland Transport. Sustainability, 12(23).

European Port Community Systems Association (2011.) How to develop a Port Community System. .Retrieved from http:// https://unece.org/.

IPCSA. (2020). Port Community Systems. Retrieved from https://ipcsa.international/pcs

IPCSA. (2021a). DAKOSY, Hamburg, Germany. Retrieved February 3, 2021, from https://ipcsa.international/about/members/members-europe-and-north-america/dakosy/

IPCSA. (2021b). PORTIC, Barcelona, Spain.

Irannezhad, E., Hickman, M., & Prato, C. G. (2017). Modeling the Efficiency of a Port Community System as an Agent-based Process. Procedia Computer Science, 109, 917–922. https://doi.org/10.1016/j.procs.2017.05.422

João, L. N., Batista, C., Ayala Botto, M., & Cordón Lagares, E. (2018). Computational Framework for Port Community Systems Towards Synchromodal Freight Networks. Estudios de Economia Aplicada, 36–3, 691–714.

Johansson, B., & Muhic, M. (2017). Relativism in the Cloud: Cloud Sourcing in virtue of IS Development Outsourcing - A literature review. International Journal of Information Systems and Project Management, 5(4), 55–65.

Jović, M., Tijan, E., Žgaljić, D., & Karanikić, P. (2020). SWOT analysis of selected digital technologies in transport economics. In 43rd International Convention on Information, Communication and Electronic Technology (pp. 1635–1640). Opatija, Croatia. Retrieved from https://www.researchgate.net/publication/344440791_SWOT_analysis_of_selected_digital_technologies_in_transport_economics

Keceli, Y. (2011). A proposed innovation strategy for Turkish port administration policy via information technology. Maritime Policy & Management: The Flagship Journal of International Shipping and Port Research, 38(2), 151–167.

Klievink, B. (2015). Understanding Public-Private Collaboration Configurations for International Information Infrastructures. In 14th International Conference on Electronic Government (EGOV), (pp. 170–180). Thessaloniki, Greece.

Marek, R. (2017). The Role and place of Customs in the port community system experiances from Poland. In The 17th International Scientific Conferences Business Logistics in Moderen Management (pp. 451–469).

Mendes Constante, J. (2019). International case studies and good practices for implementing Port Community Systems. Retrieved January 15, 2021, from https://publications.iadb.org/publications/english/document/International_Case_Studies_and_Good_Practices_for_Implementing_Port_Community_Systems.pdf

Moros-Daza, A., Amaya-Mier, R., Garcia-Llinas, G., & Voß, S. (2019). Port Community System Adoption: Game Theoretic Framework for an Emerging Economy Case Study. In Computational Logistics (pp. 136–153).

Nota, G., Bisogno, M., & Saccomano, A (2018.) A service-oriented approach to modeling and performance analysis of Port Community Systems. In International Journal of Engineering Business Management (pp. 1-17).

Port Authority of Valencia. (2021). Retrieved January 29, 2021, from https://www.valenciaportpcs.com/en/community/port-authority/

Rodon, J., & Ramis-Pujol, J. (2006). Exploring the intricacies of integrating with a Port Community System. In 19th Bled eConferences eValues (pp. 1–15).

Simoni, M., Schiavone, F., Risitano, M., Leone, D., & Chen, J. (2020). Group-specific business process improvements via a port community system: the case of Rotterdam. Retrieved January 26, 2021, from https://www.tandfonline.com/doi/abs/10.1080/09537287.2020.1824029?journalCode=tppc 20

Tijan, E., Agatić, A., & Hlača, B. (2012). The Necessity of Port Community System Implementation in the Croatian Seaports. PROMET - Traffic&Transportation, 24(4), 305–315. https://doi.org/10.7307/ptt.v24i4.444

Tijan, E., Agatić, A., Jović, M., & Aksentijević, S. (2019). Maritime National Single Window — A Prerequisite for Sustainable Seaport Business. MDPI Sustainability, 11(17), 1–21.

Tijan, E., Aksentijević, S., & Čišić, D. (2014). Seaport cluster information systems - a foundation for Port Community Systems' architecture. In Digital Economy - MIPRO 2014 37th International Convention Proceedings (pp. 1774–1779).

Tijan, E., Jović, M., & Karanikić, P. (2019). Economic and ecological aspects of electronic Transportation Management Systems in seaports. Retrieved December 31, 2019, from https://www.bib.irb.hr/1003853

Tijan, E., Kos, S., & Ogrizović, D. (2009). Disaster recovery and business continuity in port community systems. Scientific Journal of Maritime Research, 23(1), 243–260.

Treppte, S. (2011). The Role and Scope of Port Community Systems in Providing Data that Enhances Supply Chain Risk Management - A Case Study for Freight Forwarders in the Port of Rotterdam. Retrieved November 14, 2019, from https://www.erim.eur.nl/fileadmin/default/content/erim/research/centres/smart_port/ad min/c_news/110911_st_master thesis_pcss and risk management.pdf

Tsamboulas, D., Moraiti, P., & Lekka, A. M. (2012). Performance Evaluation for Implementation of Port Community System. Transportation Research Record: Journal of the Transportation Research Board, (2273), 29–37.

Waterschoot, K. (n.d.). Antwerp Port System.

Zerbino, P., Aloini, D., Dulmin, R., & Mininno, V. (2019). Towards Analytics-Enabled Effciency Improvements in Maritime Transportation: A Case Study in a Mediterranean Port. Sustainability, 11(16).

Clinical Tele-Assessment: The Missing Piece in Health Care Pathways for Orthopaedics

Oren Tirosh,[2] Muhammad Nadeem Shuakat,[1]
John Zelcer[2] & Nilmini Wickramasinghe[1,2]

[1] Epworth Healthcare, Australia; e-mail: drnadeem.work@gmail.com, nilmini.work@gmail.com
[2] Swinburne University Of Technology, Australia; e-mail: otirosh@swin.edu.au, john.zelcer@gmail.com

Abstract An aging population coupled with longer life expectancy has resulted in an exponential growth in total hip and total knee replacements (THR) (TKR). Especially during the 2020 COVID-19 pandemic, support for patients recovering form THR and TKR was difficult due to reduction in face-to-face visits. To address this and enable Australians to have a better patient experience, the following proffers a tele-assessment solution, ARIADNE (Assist foR hIp AnD kNEe), that can provide high quality care, with access for all and support for high value outcomes.

1   Introduction

In Australia, the exponential growth of joint replacements, in particular total hip (THR) and total knee replacements (TKR), is projected to reach an unsustainable burden by 2030 (Ackerman et al., 2019), which has many severe and far-reaching implications for healthcare delivery and for the demand on public and private hospitals. Given several key contributing factors, most notably an aging population and longer life expectancy (Nutt & Solan, 2017), the most prudent way to address this is to leverage technology solutions that can support cost-effective, efficient and effective care delivery post-surgery. We proffer tele-assessment, a noted void in current telemedicine solutions for orthopaedic care, as such a solution.

A key bottleneck in the recovery from THR and TKR is the return to appropriate postural and functional control (Russell, Buttrum, Wootton, & Jull, 2011). The current standard clinical pathway involves 12 to 60 face to face visits over a period of three months (The Brigham and Women's Hospital, 2010). This is not only costly and difficult to manage, especially for isolated and disadvantaged populations (The Brigham and Women's Hospital, 2010), but if not done successfully leads to poor clinical outcomes and low patient satisfaction (Jansson, Harjumaa, Puhto, & Pikkarainen, 2020). Moreover, clinical best practice notes that this 3-month window post-surgery is imperative for optimal recovery and best results (Aasvang, Luna, & Kehlet, 2015). To address this critical aspect on the THR and TKR patient journey, and support quality clinical outcomes and patient satisfaction as well as ease the burden for our healthcare system, we design, develop and test ARIADNE (Assist foR hIp AnD kNEe), a pervasive tele-assessment solution that can support clinical tele-assessment to assess postural and functional control to support post-surgery THR and TKR recovery. ARIADNE will enable objective, remote examination and monitoring of patient functional performance during their typically long rehabilitation journey, something that to date is missing from current telemedicine solutions especially in orthopaedic care. By implementing such a pervasive tele-assessment solution within traditional practice, we have the potential to: a) improve existing practice patterns, b) shorten the recovery trajectory, c) increase the likelihood for optimal clinical outcomes, and d) support a superior patient experience.

Background

Total hip or knee replacement is a common surgical intervention for treating advanced hip/knee Osteoarthritis (OA). As a strategy to address the burden of disease of OA in Victoria and optimally align health services to consumers' needs and evidence, the Department of Health and Human Services (Victoria) commissioned the development of a Model of Care (MoC) for Osteoarthritis of the Hip and Knee. A MoC is an evidence and consultation-informed framework that describes what and how health services and other resources should be delivered locally to people who live with specific health conditions. In 2018 (Victorian Musculoskeletal Clinical Leadership Group, 2018), the MoC recommended the "Innovation in service delivery model". The model was designed to establish: 1) telehealth services to improve consumers' access to specialist clinics for the purposes of clinical assessment, management planning and treatment, and 2) web-based and smartphone app tools that deliver accurate health information and support behaviour change to consumers and care providers. The development of ARIADNE is designed to addresses the above acknowledgement of the importance of a telemedicine platform to improve health care services for rehabilitation following THR and TKR. (Department of Health, 2010; Victorian Musculoskeletal Clinical Leadership Group, 2018).

Tele-rehabilitation via online video communication is an emerging area attracting increased attention as a potential alternative to conventional, face to face rehabilitation, suggested to be an option for people located remotely to reduce the need for frequent travel (Russell et al., 2011). A recent systematic review concluded that tele-rehabilitation can lead to better healthcare at lower costs (Klaassen, van Beijnum, & Hermens, 2016). An example is the tele-rehabilitation system eHAB (NeoRehab, Brisbane, Australia) that enables real-time video conferencing to the patient's home and includes features such as recording instruction and exercises (Richardson, Truter, Blumke, & Russell, 2017). Similarly, MyRehab offers a tele-rehabilitation communication system via text or voice messages and video-conference, evaluated in a RCT with THR and TKR patients (Eichler et al., 2017).

Indeed, tele-rehabilitation partially addresses some of the requirements of the MoC. However, a critical missing element in current solutions is tele-assessment which supports an objective remote postural and functional assessment integrated with

web-based management and planning capabilities. ARIADNE addresses this key void by being able to transform standard care with a face to face assessment, mostly available only in major cities with experts, to provide remote assessment access and quality of care to a wider and remote community. Thus, ARIADNE, will significantly enhance Australian health care services, ensuring objective postural and functional examination can be performed. It will provide the foundation for future telemedicine platforms for clinical trials and treatment monitoring.

To further improve upon the Australian telehealth system the following objectives need to be addressed:

1. An <u>examination</u> into the measurement consistency and agreement of a newly established tele-assessment system with respect to a face to face clinical based reference condition.
2. A <u>determination</u> of the feasibility and the extent to which the tele-assessment can be used by clinicians and patients to achieve effectiveness (accuracy and completeness), efficiency (resources needed for effectiveness) and satisfaction (comfort and acceptability).
3. An <u>assessment</u> of the cost-effectiveness associated with tele-assessment including those related to healthcare, purchase of equipment, mobile phone data usage, and costs associated with establishing and delivering the service, and analysing the results.

Overview of ARIADNE

In orthopaedics, performance measures following THR and TKR are required to identify patient functional competency and physical progress. In existing clinical practice, these postural and functional measures include: 1) range of motion, 2) postural balance, 3) chair rise, 4) 40 meter fast paced walk, and 5) timed up and go (TUG) (Victorian Musculoskeletal Clinical Leadership Group, 2018), and are executed face to face while the clinician manually records the duration and number of repetitions to complete the task. A more robust objective, but to date only used in research and not in clinical settings due to availability and accessibility, is quantifying performance using Inertial Measuring Units (IMU) motion sensors comprising accelerometers and gyroscopes (Galan-Mercant, Baron-Lopez, Labajos-Manzanares, & Cuesta-Vargas, 2014; Pham et al., 2018; Witchel et al., 2018) to

O. Tirosh, M. Nadeem Shuakat, J. Zelcer & N. Wickramasinghe:
Clinical Tele-Assessment: The Missing Piece in Health Care Pathways for Orthopaedics

57

measure linear acceleration and angular velocity, respectively. Once the raw data is captured, the level of performance is quantified by a further well-defined signal processing methods (Adusumilli et al., 2017; Galan-Mercant et al., 2014; Pham et al., 2018; Steinberg, Adams, Waddington, Karin, & Tirosh, 2017; Steinberg, Tirosh, Adams, Karin, & Waddington, 2017).

ARIADNE has been developed and designed to support the above requirements and is built from previous work by one of the authors around web based repository applications Gaitabase and PROMsBase (Tirosh, Baker, & McGinley, 2010; Tirosh, Tran, et al., 2019), and leverages his research on the use of IMU motion sensor signals to capture, process, and interpret postural and functional performance (Kuo, Culhane, Thomason, Tirosh, & Baker, 2009; Steinberg, Adams, et al., 2017; Steinberg, Tirosh, et al., 2017; Tirosh, Orland, Eliakim, Nemet, & Steinberg, 2017, 2019, 2020). Gaitabase has been used by world leading gait laboratories for clinical gait analysis, having 22 different centres in 8 different countries on four different continents. PROMsBase is routinely used at Western Health in Victoria to collect patients' satisfaction and wellbeing data pre and post joint replacement procedures with over 8,000 questionnaires from over 10,000 surgery procedures now collected.

In order to be clinically useful as a tele-assessment platform, we extended the technology with unique integration methods of the web based repository system coupled with the motion sensor IMU data captured from a mobile phone. During assessment, the clinician remotely connects to the motion capture app installed on the patient's mobile phone that is strapped above the ankle (to measure joint angle) or at the lower back (to measure postural control) using an ankle or waist strap, respectively. Once connected, the clinician remotely operates the app while the patient performs the specific functional task as instructed by the clinician. Once the task is completed the clinician remotely saves the mobile sensors data that is automatically uploaded to the web-based application for further processing and analysis. Both the clinician and the patient can login to the web application and generate performance and progress reports. Figure 1 shows an example of the motion sensors data analysis during the clinical tasks.
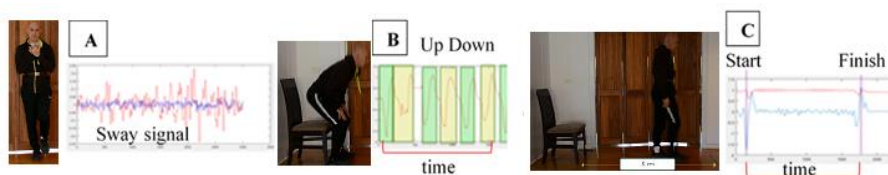
Figure 1. Example of motion sensor data during Tele-assessment of (A) postural balance, (B) sit stand, and (C) TUG tests.

Methodology

To test the proposed tele-assessment solution this section outlines the research plan. Tele-assessment is the missing piece in telemedicine care for orthopaedic rehabilitation. Unique aspects of ARIADNE include that it: a) can provide remote, quantified, and postural and function control, and b) provide early detection of deviation, problems and potential complications. Our pilot study will serve to incorporate key co-design principles to ensure clinician and patient input in the design and development of ARIADNE, and then test the definitive solution in terms of: (a) desirability (patients and clinicians), and clinician and patient usability and acceptance, (b) reliability (ability to deliver consistently on key clinical outcomes) and "fit for purpose", and (c) cost-effectiveness.

The tele-assessment platform ARIADNE is very simple to use as it integrates a web-based database and interface platform and motion sensor data that is captured remotely from the patient's mobile phone while the patient performs their essential postural and functional measures, including (Victorian Musculoskeletal Clinical Leadership Group, 2018): 1) range of motion, 2) postural balance, 3) chair rise, 4) 40 meter fast paced walk, and 5) timed up and go (TUG). The motion sensor data is processed to objectively quantify the patient's performance level.

Overview of design

To ensure a robust solution it is essential to conduct a feasibility pilot study to measure the desirability, usability, reliability, Minimal Detectable Change (MDC), fit for purpose, and cost-effectiveness, and better access, quality and value of the tele-assessment platform for THR and TKR patients. The tele-assessment will be utilised during the conventional THR and TKR standard care pathway.

O. Tirosh, M. Nadeem Shuakat, J. Zelcer & N. Wickramasinghe:
Clinical Tele-Assessment: The Missing Piece in Health Care Pathways for Orthopaedics

59

Protocol

We have secured HREC (Human Research Ethics Committee) approval and now planning for final ARIADNE <u>co-design</u> session with 2 clinicians and 4 patients (2 TKR and 2 THR). Prior to tele-assessment at **T0**, clinicians and patients will participate in an educational focus group session. In this session participants will be educated on the use of ARIADNE with preparation for their joint replacement journey.

Tele-assessments will be performed on 10 occasions including; base-line pre surgery-1 (**T1**), pre surgery-2 (**T2**), and at 1 (**T3**), 2 (**T4**), 3 (**T5**), 4 (**T6**), 5 (**T7**), 6 (**T8**), 9 (**T9**), and 12 (**T10**) weeks post-surgery. The duration of each tele-assessment session is 20 minutes. In each session, the patient will start the app that automatically connects to the clinician web portal. The patient will then insert the mobile phone in the waist pouch and attach it around their waist. The clinician will instruct the patient to perform the tasks (balance, TUG, chair rise, range of motion) while the mobile app captures the motion data and automatically uploads it to the web portal for storage and further analysis. At **T2** and **T10** patient questionnaires will be administered to assess usability of ARIADNE. In addition, at **T10** a clinician focus group will be conducted.

Outcome and data analysis: At the completion of the study the following aims will be achieved:

Primary Aim (a): assess desirability and usability of ARIADNE:

The Unified Theory of Acceptance and Use of Technology (UTAUT) will be used as the theoretical framework underpinning the research, to understand and empirically test the factors that influence the end-users' acceptance and adoption of the tele-assessment services in THR and TKR patients. The UTAUT includes four determinants; performance expectancy (PE), effort expectancy (EE), social influence (SI), and facilitating conditions (FCs), which can explain 70% of the variance of behavioural intentions (Hoque & Sorwar, 2017), while other models explained approximately 40% of technology acceptance (Lim, 2003). It has been widely adopted in different areas including mobile health such as to investigate the intention to use Physical Activity Apps (Liu et al., 2019), but to our knowledge there

is no published study applying UTAUT to the investigation of tele-assessment use intention in joint replacement patients. To specify, PE refers to the services of the clinician in the clinician–patient interaction in tele-assessment. EE refers to patients' perceived ease of interacting with physicians in tele-assessment. SI refers to the impact of other people's feelings, views, and behaviours on the behavioural intention of patients interacting with clinicians in telehealth. FC refers to users' perceptions of their ability to perform the behaviour and measure the degree to which the tele-assessment fits with their existing values, previous experiences and current needs. Each determinant includes related questions with a 5-point Likert-type response format that ranges from "strongly disagree" (1) to "strongly agree" (5) to measure each construct covering the variables in the research model. The UTAUT will be completed at **T1**, **T8** and **T10**. Separate analyses will be conducted for older and younger groups of participants in order to test usability and efficacy across all age levels.

Primary Aim (b): assess reliability using MDC:

The outcome measure for each postural and functional task is calculated by processing the motion sensor signal from the mobile device. The outcome measures are the hip and knee joint maximum range of motion, the magnitude of body sway for the postural balance test, the time taken to complete five repetitions of the sit-to-stand manoeuvre test, and time taken to complete the TUG and 40 meter fast paced walk tests.

Reliability and MDC will be analysed from the tele-assessment outcome measures (described above) when comparing outcomes between the pre surgery-1 and the pre surgery-2 sessions. The reliability will be estimated using repeated-measures analysis of variance and the intra-class correlation coefficient. Standard error of measure (SEM) and Minimal Detectable Change (MDC) will be calculated utilising the same methodology as our previous study using IMU measurements in gait (Tirosh, Orland, et al., 2019). SEM will be calculated as SD x $\sqrt{1 - ICC}$, where SD is the standard deviation of all scores from the participants. SEM is also presented as a SEM% by dividing the SEM with the average of the test and retest values. The MDC will be calculated as SEM x 1.96 x $\sqrt{2}$ to construct 95% CI. Multi-level analysis will be used to test for significant differences across younger and older age groups separately for THR and TKR patients. This is an intention-to-treat method which

O. Tirosh, M. Nadeem Shuakat, J. Zelcer & N. Wickramasinghe:
Clinical Tele-Assessment: The Missing Piece in Health Care Pathways for Orthopaedics

61

allows for irregular assessment measures, allowing for the control of other variables as needed.

Primary Aim (c): *evaluate if ARIADNE is fit for purpose*:

Empirical evidence shows that users will not simply accept and use the technology if it does not fit their needs and improve their performance (Gebauer & Ginsburg, 2009; Junglas & Watson, 2008). Hence, to assess the fit for purpose of ARIADNE, we apply *task-technology fit*, the degree to which a technology assists an individual in performing his or her task. To measure the level of "fit for purpose" of ARIADNE, during the focus group session conducted at **T10**, we will ask clinicians to compare their experience with ARIADNE to that without tele-assessment; i.e. as per their normal delivery of care. This will enable us to understand key task characteristics, how the technology supported those tasks, and whether clinician users perceived ARIADNE to perform better, as good as, or worse than face to face assessments.

Identified Risks

Risk #1 – Poor desirability and reliability of the tele-assessment platform: The tele-assessment platform is built for the purpose of assessing joint replacement intervention postural and functional control outcomes. Thus, it is important for the platform to be desirable and reliable. This risk is mitigated from findings of previous studies on the desirability and reliability of using IMU in measuring postural and functional performance of the proposed tasks in a range of population types including: healthy older adults (Chan, Keung, Lui, & Cheung, 2016; Keogh et al., 2019), Parkinson's disease (Pham et al., 2018), multiple sclerosis (Witchel et al., 2018), and TKR (Huang, Liu, Hsu, Lai, & Lee, 2020), with very good to excellent desirability and reliability. These studies and a systematic review (Keogh et al., 2019) used mobile phone motion sensors, suggesting that mobile phones are non-inferior compared to the other postural and functional measurement techniques. Furthermore, our unpublished preliminary experiment validating our tele-assessment knee range with the gold standard video analysis showed excellent correlation ($r = 0.98$) and very good agreement with clinically acceptable bias of 5.4 degrees with 17.3 and -6.4 degrees for upper and lower 95% confidence bounds respectively.

Risk #2 – Security of the data: to increase data security the data will be stored on Nectar cloud and will be backed up daily. Nectar is an online infrastructure for researchers to store, access, and analyse data remotely, and is managed and funded by the Australian Government through the National Collaborative Research Infrastructure Strategy.

Risk #3 – Participants can feel directionless and overwhelmed with the technology: to reduce participant anxiety with the digital technology, the patient advocate team member will provide coaching and mentoring support to participants, enabling them to optimise self-management.

Risk #4 – Falls, limited space, and poor environmental set up at home or difficulties to attach the phone to the ankle and/or waist: patients' home visits will occur to inspect and set up the testing environment, and reduce potential hazards.

Risk #5 - Poor Internet connection: the necessary internet connection will utilise the mobile device internet access provider. Data collection is performed using the app installed on the patient's mobile device, thus a poor network will not disrupt data capture and quality.

Results to Date:

The assessment of the fidelity, efficacy and fit for purpose of the developed ARIADNE solution requires many stages and is thus a longitudinal study. On the receipt of ethics approval, initial phases of the design science research methodology have been conducted with a small group of patients and clinicians respectively to fine tune the solution. This is an important key step to ensure high clinician and patient use as well as ensure the developed solution will support the required needs for rehab of THR and TKR patients. The ARIADNE solution now has patient and clinician approval and based on a small pilot study demonstrated ease of use and fit for purpose. While not statistically significant, this directional data provides support to progress to the next phase with confidence. The next key step is to conduct a large scale clinical trial to capture key data around the impact of the solution to support THR and TKR patients in their rehabilitation. Once the clinical trial is

O. Tirosh, M. Nadeem Shuakat, J. Zelcer & N. Wickramasinghe:
Clinical Tele-Assessment: The Missing Piece in Health Care Pathways for Orthopaedics

63

concluded it will then be possible to address issues around deployment of the solution into appropriate clinical contexts.

Discussion:

ARIADNE is designed to remotely capture, analyse, and interpret body motion using the accelerometer and gyroscope motion sensors embedded in today's mobile phones. The mobile phones we have today have 3-axial accelerometer and gyroscope components. The accelerometer allows the measurement of linear acceleration in three orthogonal directions (x, y, and z) and the gyroscope allows the measurement of angular velocity in the x, y, and z axes. The linear acceleration and angular velocity signals can be processed and used to analyse body motion and further provide interpretation of the movement quantity and quality, such as level of stability during quiet standing. The ability to remotely connect to a patient's mobile phone and capture accelerometer and gyroscope data creates new opportunities for clinicians, sport trainers, and engineers to remotely quantify and analyse the performance level of any posture and/or movement task.

Conclusions:

Our designed solution, ARIADNE, represents a novel and unique approach to telehealth rehabilitation in orthopaedic care for THR and TKR patients. To date, current telehealth solutions in this space do not address tele-assessment, which means that there is a significant limitation in the current post-operative critical 12-week period for THR and TKR patients. Hence, ARIADNE not only addresses this key void, but it serves to also potentially help to address a major conundrum facing healthcare delivery around THR and TKR. If the results of the clinical trial provide a positive endorsement for ARIADNE, then we would have successfully developed a unique, COVID-19 safe, tele-assessment solution that addresses a key gap in post-surgical recovery for THR and TKR patients.

References

Aasvang, E. K., Luna, I. E., & Kehlet, H. (2015). Challenges in postdischarge function and recovery: the case of fast-track hip and knee arthroplasty. Br J Anaesth, 115(6), 861-866. doi: 10.1093/bja/aev257

Ackerman, I. N., Bohensky, M. A., Zomer, E., Tacey, M., Gorelik, A., Brand, C. A., & de Steiger, R. (2019). The projected burden of primary total knee and hip replacement for osteoarthritis in Australia to the year 2030. BMC Musculoskelet Disord, 20(1), 90. doi: 10.1186/s12891-019-2411-9

Adusumilli, G., Joseph, S. E., Samaan, M. A., Schultz, B., Popovic, T., Souza, R. B., & Majumdar, S. (2017). iPhone Sensors in Tracking Outcome Variables of the 30-Second Chair Stand Test and Stair Climb Test to Evaluate Disability: Cross-Sectional Pilot Study. JMIR Mhealth Uhealth, 5(10), e166. doi: 10.2196/mhealth.8656

Chan, M. H. M., Keung, D. T. F., Lui, S. Y. T., & Cheung, R. T. H. (2016). A validation study of a smartphone application for functional mobility assessment of the elderly. Hong Kong Physiother J, 35, 1-4. doi: 10.1016/j.hkpj.2015.11.001

Department of Health, W. A. (2010). Elective Joint Replacement Service Model of Care. Western Australia.

Eichler, S., Rabe, S., Salzwedel, A., Muller, S., Stoll, J., Tilgner, N., . . . ReMove-It study, g. (2017). Effectiveness of an interactive telerehabilitation system with home-based exercise training in patients after total hip or knee replacement: study protocol for a multicenter, superiority, no-blinded randomized controlled trial. Trials, 18(1), 438. doi: 10.1186/s13063-017-2173-3

Galan-Mercant, A., Baron-Lopez, F. J., Labajos-Manzanares, M. T., & Cuesta-Vargas, A. I. (2014). Reliability and criterion-related validity with a smartphone used in timed-up-and-go test. Biomed Eng Online, 13, 156. doi: 10.1186/1475-925X-13-156

Gebauer, J., & Ginsburg, M. (2009). Exploring the black box of task-technology fit. Communications of the ACM, 52(1), 130-135.

Hoque, R., & Sorwar, G. (2017). Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. Int J Med Inform, 101, 75-84. doi: 10.1016/j.ijmedinf.2017.02.002

Huang, Y. P., Liu, Y. Y., Hsu, W. H., Lai, L. J., & Lee, M. S. (2020). Progress on Range of Motion After Total Knee Replacement by Sensor-Based System. Sensors (Basel), 20(6). doi: 10.3390/s20061703

Jansson, M. M., Harjumaa, M., Puhto, A. P., & Pikkarainen, M. (2020). Patients' satisfaction and experiences during elective primary fast-track total hip and knee arthroplasty journey: A qualitative study. J Clin Nurs, 29(3-4), 567-582. doi: 10.1111/jocn.15121

Junglas, I., & Watson, R. (2008). Location-based services. Communications of the ACM, 51(3), 65-69.

Keogh, J. W. L., Cox, A., Anderson, S., Liew, B., Olsen, A., Schram, B., & Furness, J. (2019). Reliability and validity of clinically accessible smartphone applications to measure joint range of motion: A systematic review. PLoS One, 14(5), e0215806. doi: 10.1371/journal.pone.0215806

Klaassen, B., van Beijnum, B. J., & Hermens, H. J. (2016). Usability in telemedicine systems-A literature survey. Int J Med Inform, 93, 57-69. doi: 10.1016/j.ijmedinf.2016.06.004

Kuo, Y. L., Culhane, K. M., Thomason, P., Tirosh, O., & Baker, R. (2009). Measuring distance walked and step count in children with cerebral palsy: an evaluation of two portable activity monitors. Gait Posture, 29(2), 304-310. doi: 10.1016/j.gaitpost.2008.09.014

Lim, N. (2003). Consumers' perceived risk: sources versus consequences. Electronic Commerce Research and Applications, 2(3), 216-228.

Liu, D., Maimaitijiang, R., Gu, J., Zhong, S., Zhou, M., Wu, Z., . . . Hao, Y. (2019). Using the Unified Theory of Acceptance and Use of Technology (UTAUT) to Investigate the Intention to Use Physical Activity Apps: Cross-Sectional Survey. JMIR Mhealth Uhealth, 7(9), e13127. doi: 10.2196/13127

Nutt, J. G., & Solan, M. C. (2017). Ageing and orthopaedics. Orthopaedics and Trauma, 31(5), 321-325.

O. Tirosh, M. Nadeem Shuakat, J. Zelcer & N. Wickramasinghe:
Clinical Tele-Assessment: The Missing Piece in Health Care Pathways for Orthopaedics

65

Pham, M. H., Warmerdam, E., Elshehabi, M., Schlenstedt, C., Bergeest, L. M., Heller, M., . . . Maetzler, W. (2018). Validation of a Lower Back "Wearable"-Based Sit-to-Stand and Stand-to-Sit Algorithm for Patients With Parkinson's Disease and Older Adults in a Home-Like Environment. Front Neurol, 9, 652. doi: 10.3389/fneur.2018.00652

Richardson, B. R., Truter, P., Blumke, R., & Russell, T. G. (2017). Physiotherapy assessment and diagnosis of musculoskeletal disorders of the knee via telerehabilitation. J Telemed Telecare, 23(1), 88-95. doi: 10.1177/1357633X15627237

Russell, T. G., Buttrum, P., Wootton, R., & Jull, G. A. (2011). Internet-based outpatient telerehabilitation for patients following total knee arthroplasty: a randomized controlled trial. J Bone Joint Surg Am, 93(2), 113-120. doi: 10.2106/JBJS.I.01375

Steinberg, N., Adams, R., Waddington, G., Karin, J., & Tirosh, O. (2017). Is There a Correlation Between Static and Dynamic Postural Balance Among Young Male and Female Dancers? J Mot Behav, 49(2), 163-171. doi: 10.1080/00222895.2016.1161595

Steinberg, N., Tirosh, O., Adams, R., Karin, J., & Waddington, G. (2017). Influence of Textured Insoles on Dynamic Postural Balance of Young Dancers. Med Probl Perform Art, 32(2), 63-70. doi: 10.21091/mppa.2017.2012

The Brigham and Women's Hospital, I., Department of Rehabilitation Services. (2010). Standard of Care: Total Hip Replacement.

Tirosh, O., Baker, R., & McGinley, J. (2010). GaitaBase: Web-based repository system for gait analysis. Comput Biol Med, 40(2), 201-207. doi: 10.1016/j.compbiomed.2009.11.016

Tirosh, O., Orland, G., Eliakim, A., Nemet, D., & Steinberg, N. (2017). Tibial impact accelerations in gait of primary school children: The effect of age and speed. Gait Posture, 57, 265-269. doi: 10.1016/j.gaitpost.2017.06.270

Tirosh, O., Orland, G., Eliakim, A., Nemet, D., & Steinberg, N. (2019). Repeatability of tibial acceleration measurements made on children during walking and running. J Sci Med Sport, 22(1), 91-95. doi: 10.1016/j.jsams.2018.04.006

Tirosh, O., Orland, G., Eliakim, A., Nemet, D., & Steinberg, N. (2020). Attenuation of Lower Body Acceleration in Overweight and Healthy-Weight Children During Running. J Appl Biomech, 1-6. doi: 10.1123/jab.2019-0138

Tirosh, O., Tran, P., Renouf, J., Pergaminelis, N., Purdie, C. N., Ho, A., & Gibbens, A. (2019). PROMsBase: Web-based repository portal for patient-reported outcome measures in orthopaedics. Health Informatics J, 25(3), 867-877. doi: 10.1177/1460458217725904

Victorian Musculoskeletal Clinical Leadership Group. (2018). Victorian Model of Care for Osteoarthritis of the Hip and Knee. Melbourne: MOVE muscle, bone & joint health.

Witchel, H. J., Oberndorfer, C., Needham, R., Healy, A., Westling, C. E. I., Guppy, J. H., . . . Klucken, J. (2018). Thigh-Derived Inertial Sensor Metrics to Assess the Sit-to-Stand and Stand-to-Sit Transitions in the Timed Up and Go (TUG) Task for Quantifying Mobility Impairment in Multiple Sclerosis. Front Neurol, 9, 684. doi: 10.3389/fneur.2018.00684

MAKE US SMILE! AI AND THE VIOLATION OF HUMAN INTENTIONS

CHRISTOF WOLF-BRENNER

Know-Center GmbH, Graz, Austria; e-mail: christof.brenner@gmx.at

Abstract In his book Superintelligence, Nick Bostrom points to several ways the development of Artificial Intelligence (AI) might fail, turn out to be malignant or even induce an existential catastrophe. He describes 'Perverse Instantiations' (PI) as cases, in which AI figures out how to satisfy some goal through unintended ways. For instance, AI could attempt to paralyze human facial muscles into constant smiles to achieve the goal of making humans smile. According to Bostrom, cases like this ought to be avoided since they include a violation of human designer's intentions. However, AI finding solutions that its designers have not yet thought of and therefore could also not have intended is arguably one of the main reasons why we are so eager to use it on a variety of problems. In this paper, I aim to show that the concept of PI is quite vague, mostly due to ambiguities surrounding the term 'intention'. Ultimately, this text aims to serve as a starting point for a further discussion of the research topic, the development of a research agenda and future improvement of the terminology.

1   Introduction

Recently, the importance of being in control of what Artificial Intelligence (AI) does has moved to the center of attention. There appears to be a broad consensus that AI should not do what we homo sapiens do not want it to do. The concept of Perverse Instantiations (PI) consequently describes cases where AI succeeds to achieve goals but does so in violation with our intentions. For instance, best-selling author and philosopher Nick Bostrom brought forward the thought experiment of AI choosing to paralyze human facial muscles into constant smiles to achieve the goal of making humans smile. Of course, similar more or less realistic cases can be constructed for various domains. AI might for instance attempt to achieve the goal of reducing maternal mortality by sterilizing all male homo sapiens, or to improve schoolchildren's grades by providing them with the answers to the next test beforehand. While ultimately, the goal is achieved in each scenario, the way it was achieved was not intended.

According to Bostrom, PI ought to be avoided because of the violation of human designer's intentions. In this paper, I aim to show that the current concept of PI is quite vague, mostly due to inaccuracies surrounding the term 'intention'. The prevailing terminology, if we took it seriously, would force us to label most ways of achieving goals that were uncovered by AI as unintended and consequently, as PI. Ultimately, this text serves as a starting point for a further discussion of the research topic and aims to provide reasoning why we should look deeper into the matter at hand.

2   The Meaning of 'Intention'

At first glance, Bostrom's definition of PI as AI "discovering some way of satisfying the criteria of its final goal that violates the intentions of the programmers who defined the goal"[1] sounds reasonable. In essence, he claims that AI going against the intentions of its designers might lead to undesirable outcomes. But what is the real meaning of 'intention' that he has in mind when framing his notion of PI? Revisiting the thought experiment of producing smiles, Bostrom explicates that violating intentions is to be understood as "not to do what the programmers meant when

---

[1] Bostrom 2017, p. 146

they wrote the code that represents this goal"[2]. But what did they mean when they tasked AI with coming up with ways to make humans smile? To answer that, we will take a short detour into the realm of Reinforcement Learning (RL) as a state-of-the-art example and the technologically simplest instance of AI that Bostrom himself uses when describing PI.

RL is a subfield of AI and Machine Learning that is focusing on automatically learning optimal decisions over time. For simplicity's sake, imagine a mouse-in-a-labyrinth kind of experiment. Within the maze, the designer randomly places non-deadly traps and delicious food. A mouse, conceptually referred to as an RL agent[3], is placed in the maze and can perceive its environment through its senses. Of course, its objective is to try and obtain as much food as possible without getting hurt by the traps.[4]

To achieve its goal, the digital rodent can mix and match actions from a finite list of actions called the action space, i.e., turn around, move, wait, gnaw, jump etc. RL can then be understood as repeatedly putting the same mouse into very many mazes to finally make it learn to automatically choose the optimal combination of actions from its action space to maximize the aggregate reward, i.e., eat the maximum amount of food while stepping into the fewest traps.[5]

3   Perverse Instantiations Emerging from Underspecified Goals

If we were to transfer the concept of RL to Bostrom's example of tasking AI to make us smile, we encounter a few challenges. Initially, we would need to adequately represent the goal in a way so that we can give our agent feedback on how well it has done, which includes translating 'make us smile' into a form that the AI can understand. First, it is up for interpretation what counts as a smile. Second, it is unclear whether the AI is meant to achieve the maximum number, duration, intensity etc. of smiles. Third, the term 'us' is ambiguous and contextual, potentially leaving us with an AI that might produce a lot of smiling corpses.

---

[2] Bostrom 2017, p. 147
[3] I use the term 'agent' here in the sense in which it is used in the domain of Reinforcement Learning.
[4] See Lapan 2020, pp. 1–5
[5] Ibid.

Making humans smile by paralyzing facial muscles might not be intended, but if what is intended is not made explicit in the goals, a RL agent will not be able to take it into account, as it is only interested in maximizing goal achievement, thus reward gain.

As shown with Bostrom's own example, it is in fact not the case that AI discovered a loophole to perversely instantiate the sought-after goals. The problem appears to be the programmers' vague formulation of goals.

Yet, if there is so much room for interpretation, how can we then speak of a violation of intentions? If I kindly ask you to prepare a sandwich for me, and you go ahead and fix up a ham and cheese sandwich, can we really speak of you violating my intentions because unbeknownst to you, I am a vegetarian? It is my firm believe that we should not put the blame on the sandwich-maker if he only knew half of what was truly expected of him.

4   Perverse Instantiations Emerging from the Way Goals Are Achieved

Let us, however, assume, that there are cases in which the goals are in fact stated in a way such that the intentions of the programmers are absolutely and unmistakably clear. Could there still be PI? Bostrom's examples strongly imply that even clear-cut goals can be perversely instantiated by employing unintended ways to reach them. Reconsidering our mouse-in-a-labyrinth scenario, the finite list of actions that the mouse can take to interact with the maze is called the action space. It can be understood as a kind of toolbox that programmers infuse their RL agents with to act in the environment.

However, there are two ways I can think of by which a finite list of actions might still circumvent the designer's intentions. First, knowing all individual actions might simply not be sufficient to check for conformance. For instance, if our mouse in the maze would be able to act in three distinct ways, a solution to optimize reward gain might include chaining together these actions up to ten times, which already amounts to well over 500 possible combinations. There is a significant chance for combinations that the programmers never would have thought of, and therefore, could not have been intended by them.

C. Wolf-Brenner:
Make Us Smile! AI and the Violation of Human Intentions

71

Second, unexpected interactions between agents and complex environments in RL leave ample room for PI. Researchers tasked RL agents to play hide and seek against each other, and the emerging collaborative strategies far exceeded what was initially anticipated. To win, seekers learned to surf on crates by exploiting the way movement was implemented. By doing so, they were able to overcome the shelters that hiders had built as part of their defensive strategy[6]. The sheer range of things that agents could do in the game world was simply ungraspable for humans beforehand, which lead to unintended solutions, even though the goal was never unclear or misunderstood.

5   The Intention of Violating Intentions

By design, RL scenarios aim to produce optimal solutions for gaining some reward with minimal or no explicit instructions on how to do so. In a way, RL agents having the freedom to experiment within some boundaries is exactly what we intend to do when employing that kind of AI. So, following our arguments from the previous sections, should we label every instance of AI not doing what the programmer meant as PI?

When Lee Sedol, one of the best Go players on the planet, sat down to play against AlphaGo, an AI trained by way of RL, the human champion lost four out of five rounds. Move 37, which the AI came up with in the second game, stunned Sedol. In thousands of years of humans playing Go, nobody had ever come up with something as inhuman, unique, or creative.[7] But what were the intentions of the designers of AlphaGo? Clearly, they meant to design an AI that can play and excel at Go. Obviously, Move 37 was intended insofar as it is in accordance with the goal, which I loosely interpret as win at Go by playing the game by the rules. But did the programmers intend Move 37? Arguebly, they did not. Move 37 was unexpected for the opponent, the spectators and even more so for the creators of the AI. AlphaGo itself estimated that a human player would have played this move with a probability of one in 10,000 but decided to go for it anyway[8].

---

[6] See Baker et al. 2020, p. 6
[7] See Holcomb et al. 2018, p. 68
[8] See Holcomb et al. 2018, p. 70

Bostrom does slightly hint at his account of intention being tied to the way goals are achieved in the sense of a method, and not the goals themselves[9]. If this is the case, labeling cases as PI boils down to the question whether we can at the same time intend to search for the optimal way of satisfying a goal and already know the result of the search. Searching for a solution and already intending it seems contradictory. However, if we tasked AI to make us smile, play Go or eat food and avoid traps, any case of it achieving these goals through unintended ways would still have to labeled as PI. It appears that simply reducing PI to unintended ways to achieve goals is not enough to clearly explain the problem arising from applying AI in such scenarios.

6   Conclusion

In this paper, I have shown that the concept of PI hinges on the definition of 'intention' and consequently, what the intention latches on to. I have concluded that PI arising from underspecified goals are simply a matter of intentions not being laid out to AI in an understandable, unmistakable, and complete manner. By no means would we be justified to declare our creation at fault, because everything it does is directed at what it knows about our intentions, expressed through goals, and achieving them.

Additionally, I have discussed that even for cases in which intentions are perfectly clear to AI, there still is room for PI based on actions and interactions with the environment not being fully understood by designers beforehand. It appears contradictory to claim to have intended a particular way of achieving a goal which AI just now discovered. In contrast, not every time some goal is achieved in a way that designers were previously unaware of is unintended, and thus a case of PI.

It is clear that the definition of 'intention' plays a pivotal role in identifying PI. However, as I have shown, there is more work to be done on the terminology. First, a starting point for further conceptual contribution to the topic could be an investigation of a potential conflation of the terms 'unintended' and 'unanticipated' in the context of unintended consequences. This distinction could shed some light on the issue of mislabeling cases as PI.

---

[9] See Bostrom 2017, p. 147

C. Wolf-Brenner:
Make Us Smile! AI and the Violation of Human Intentions

73

Second, empirical research could be conducted with regards to what cases of AI, humans or other animals achieving goals through unintended ways people would in fact label as PI. Consequently, bias for or against our artificial creations and fellow planet dwellers could be identified and elaborated on.

References

Baker, B.; Kanitscheider, I.; Markov, T.; Wu, Y.; Powell, G.; McGrew, B.; Mordatch, I. (2020): Emergent Tool Use From Multi-Agent Autocurricula. In: 8th International Conference on Learning Representations. Addis Ababa, April 26-30, 2020.

Bostrom, N. (2017): Superintelligence. Paths, Dangers, Strategies. 2nd ed. Oxford: Oxford University Press.

Holcomb, S. D.; Porter, W. K.; Ault, S. V.; Mao, G.; Wang, J. (2018): Overview on Deepmind and its AlphaGo Zero AI. In Proceedings of the 2018 International Conference on Big Data and Education, pp. 67–71.

Lapan,, M. (2020): Deep Reinforcement Learning Hands-On. 2nd ed. Birmingham: Packt Publishing.

BUSINESS DATA SHARING THROUGH DATA MARKETPLACES: A SYSTEMATIC LITERATURE REVIEW

ANTRAGAMA EWA ABBAS,[1] WIRAWAN AGAHARI,[1] MONTIJN VAN DE VEN,[2] ANNEKE ZUIDERWIJK[1] & MARK DE REUVER[1]

[1] Delft University of Technology, Faculty of Technology, Policy and Management, Delft, The Netherlands; e-mail: a.e.abbas@tudelft.nl, w.agahari@tudelft.nl, A.M.G.Zuiderwijk-vanEijk@tudelft.nl, G.A.deReuver@tudelft.nl
[2] Eindhoven University of Technology, Department of Industrial Engineering & Innovation Sciences, Eindhoven, The Netherlands; e-mail: m.r.v.d.ven@tue.nl

Abstract Data marketplaces are expected to play a crucial role in tomorrow's data economy but hardly achieve commercial exploitation. Currently, there is no clear understanding of the knowledge gaps in data marketplace research, especially neglected research topics that may contribute to advancing data marketplaces towards commercialization. This study provides an overview of the state of the art of data marketplace research. We employ a Systematic Literature Review (SLR) approach and structure our analysis using the Service-Technology-Organization-Finance (STOF) model. We find that the extant data marketplace literature is primarily dominated by technical research, such as discussions about *computational pricing* and *architecture*. To move past the first stage of the platform's lifecycle (i.e., platform design) to the second stage (i.e., platform adoption), we call for empirical research in non-technological areas, such as *customer expected value* and *market segmentation*.

1   Introduction

Data marketplaces are expected to play a crucial role in tomorrow's data economy (European Commission, 2020). Business data sharing via data marketplaces may contribute to overall economic growth by stimulating data-driven innovation, improving the competitiveness of Small and Medium-sized Enterprises (SMEs), and opening up job markets (Virkar et al., 2019). A data marketplace can be broadly defined as a multi-sided platform that matches data providers and data buyers, and that facilitates business data exchange and financial transactions. Key actors that provide data marketplace functionalities include data marketplace owners, operators, and third-party providers (TPPs) (Fruhwirth et al., 2020; Koutroumpis et al., 2020; Spiekermann, 2019). However, despite the alleged potential, data marketplaces generally remain conceptual and hardly commercially exploited. For instance, Microsoft's Azure Data Marketplace, xDayta, and Kasabi are among out-of-market data marketplaces (Spiekermann, 2019).

From an academic perspective, recent trends in the European Union policy-making agendas towards business data sharing have led to a proliferation of data marketplaces studies, resulting in a constantly expanding yet fragmented body of literature. Recent work provides understanding of the state-of-the-art of data marketplaces in practice (e.g., Fruhwirth et al., 2020) but does not provide a comprehensive overview of data marketplaces research in academia. Consequently, we have no clear understanding of the knowledge gaps in data marketplace research. Specifically, we lack understanding of whether research is lacking on topics that would advance data marketplaces towards commercialization. As it stands, it might well be that academic research is focusing on topics that do not help resolve the standstill in data marketplace commercialization. To evaluate this assertion, this paper aims to investigate the current state of the art of data marketplace research.

We will conduct a systematic literature review on existing data marketplaces research by adopting the guideline provided by Okoli (2015). To cover the broad range of issues that plays a role in technology commercialization, we use the business model construct as a synthetic device (cf., Solaimani et al., 2015). In this way, our study will be, to our knowledge, the first to provide a comprehensive overview of data marketplace research, which will be beneficial in steering future research towards the commercialization of such marketplaces. In this paper, we consider all data

A. Ewa Abbas, W. Agahari, M. van de Ven, A. Zuiderwijk &M. de Reuver:
Business Data Sharing through Data Marketplaces: A Systematic Literature Review

77

marketplaces archetypes revealed by Fruhwirth et al. (2020): centralized, decentralized, and personal data trading. In *centralized data trading*, data marketplaces mediate data exchange from diverse domains and origins, incorporating different data types and pricing mechanisms. Advanced data marketplaces in this archetype employ smart contracts to execute transactions. *Decentralized data trading,* on the other hand, relies on a decentralized architecture to operate data marketplaces. Finally, *personal data trading* refers to a Customer-to-Business (C2B) relationship where individuals can sell their personal information to corporations.

2 Research Approach

This research employs a Systematic Literature Review (SLR) approach (Okoli, 2015). Our primary database is Scopus, which comprises a comprehensive database of many scientific research papers, including the area we are examining in this study. We selected articles based on three criteria: articles should be (1) written in English; (2) published in a peer-reviewed journal or conference proceedings; and (3) focused on data marketplaces. We use the search terms of *("data marketplace*") OR ("data market*").* The literature search was conducted on 6 July 2020 and resulted in 496 articles. We complemented these articles with nine additional papers that we consider key literature. The articles did not appear in the initial search because, for instance, they do not use the *data marketplace* term explicitly, neither in the title nor abstract. We extracted the articles' meta-data and saved it in an Excel spreadsheet (available here: https://doi.org/10.4121/14673813.v1).

Next, we analyzed the quality of the identified articles by employing a two-step screening approach. First, we looked into the title and abstract of the selected papers to assess their relevance. We discussed our assessment internally to reach a consensus, resulting in an exclusion of 225 papers. Second, we combined quantitative traditional metrics (e.g., citation numbers) and next-generation metrics (e.g., social media, usage, captures, and mentions) to further select the reviewed papers. The next-generation metrics provided by the Scopus database are known as the PlumX Metrics (Champieux, 2015). To do so, we first calculated the average number of citations from the existing 280 articles as a threshold, in which we include any articles above this average citation number (7.3). If the number of citations of an article was below the threshold, we checked whether the article was published in a high-quality journal or conference proceedings (i.e., ranked above the

50th percentile in its respective domain). For articles not meeting these criteria, we used the average number of next-generation metrics from the existing 280 articles (social media = 2.1, usage = 44.8, captures = 43.2, mentions = 0.2) to check whether policymakers or practitioners have read the article. In this way, we ensure that the inclusion of our sampling papers reflects both scientific reliability and relevance. In this stage, the number of included papers was 158.

Following Solaimani et al. (2015), we used the STOF model to structure the identified articles. It is a generic framework to reconstruct the logic of a business and its ecosystems (Bouwman et al., 2008). Thus, the STOF model enables a high-level representation of the service domain (S), technology domain (T), organization domain (O), and finance domain (F). The STOF model is suitable for our purpose since it is explicitly designed for ICT-enabled services like data marketplaces. Unlike frameworks such as business model canvas, the STOF model explicitly captures the role of technology in commercialization. Moreover, the STOF model helps understand the dynamics involved in developing successful business models, i.e., market adoption and sustainable profitability of the designed services. Due to the lack of commercialized data marketplaces, it is crucial to understand what we (do not) know about the breadth of the business models of data marketplaces, ranging from their value to how they deliver and capture value. Hence, the STOF model is highly appropriate to structure our discussion.

We then read the full text of the 158 remaining articles and classified each article into **a STOF model domain**. Furthermore, each article was further classified into **a category.** In classifying an article, we identified its main objective while paying attention to the primary unit of analysis. For example, Munoz-Arcentales et al. (2019) propose an architecture for providing data usage and access control. Since the discussion emphasizes technology needs, we classified this paper into the *architecture* category in the STOF *technology* domain. Although some articles can have multiple overlapping topics, we still attempted to assign each article into a single category. We justified this by analyzing the central theme of the discussion. Various articles were independently categorized by multiple authors to assess inter-rater reliability (see the supplementary material). In general, there was a high level of agreement between the authors. We also further excluded some irrelevant articles, including those that did not discuss data marketplaces. **Our final sample consisted of 137 articles.**

A. Ewa Abbas, W. Agahari, M. van de Ven, A. Zuiderwijk &M. de Reuver:
Business Data Sharing through Data Marketplaces: A Systematic Literature Review

79

3   Results: STOF Model Categorization

This section describes the results of our STOF model categorization. In total, we identified seventeen categories. Figure 1 provides the categorization of the articles. Given the page limitation, we only provide the topic examples for each category. For the whole topic list, please refer to the supplementary material.
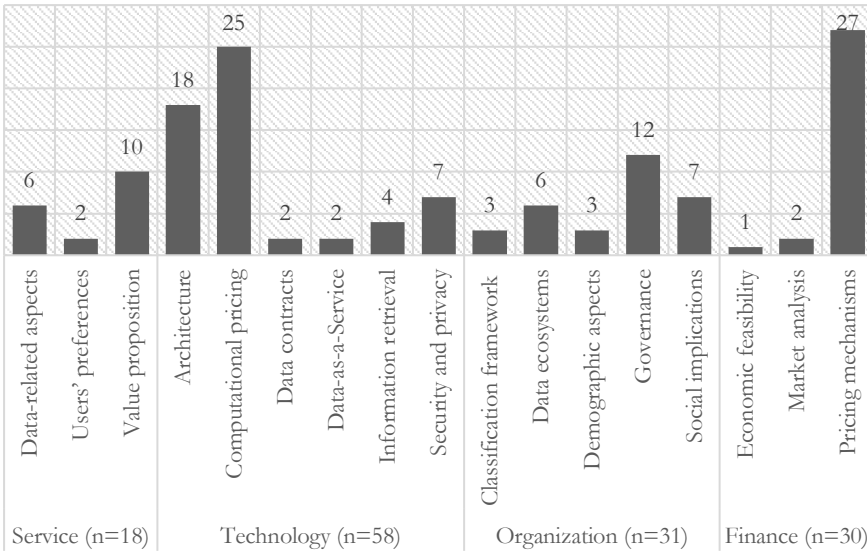
Figure 1 The selected articles categorized using the STOF model (n=137)

We identify three categories within the STOF **service domain.** The first one, the most dominant category, is the *value proposition*. The studies in this category are generally concerned with means to identify value for end-users of data marketplaces. For example, Perera et al. (2017) and Anderson et al. (2014) explore the value of trading IoT and healthcare data, respectively. An additional example is the value exploration of data marketplaces that trade anonymous personal data (Robinson, 2017). Another category that has received considerable attention from scholars concerns the *data-related aspects*. This category explores data properties as a unit of analysis, such as data characteristics as economic goods (Demchenko et al., 2018) and approaches to identify data quality problems (Zhang et al., 2019). Meanwhile, literature on the *users' preferences* discusses data providers' willingness to share data

considering aspects such as anonymity (Schomakers et al., 2020) and data ownership (Kamleitner & Mitchell, 2018).

Most publications fall within the STOF **technology domain**, which can be divided into six categories. In our sample, the most discussed category is the *computational pricing,* which focuses on technical discussions for data pricing. Publications in this category propose, for instance, online query-based mechanisms (Zheng et al., 2020), an algorithmic solution for data pricing and revenue distribution (Agarwal et al., 2019), a bi-level mathematical programming model that considers data quality (Yu & Zhang, 2017). Following this, many scholars discuss topics related to *architecture*, generally referring to the various technical components of data marketplaces. Examples of architecture in data marketplaces include blockchain-based system architectures (e.g., Jeong et al., 2020; López & Farooq, 2020), an architecture for data access and control (Munoz-Arcentales et al., 2019), and a reference architecture (Roman & Stefano, 2016). The *security and privacy* category has also gained much attention in the literature. The topics covered in this category are related to privacy-preserving technology (e.g., Niu et al., 2019; Zhao et al., 2019), property right enforcements (Sørlie & Altmann, 2019), and secure information models (Shaabany et al., 2016). *Information retrieval* topics such as schema (Hatanaka & Abe, 2015) and ontologies (Morrison et al., 2011) are also discussed in the literature. Another category examines the model for *data contracts* to explicitly specify data usage (Truong et al., 2012). Finally, the *data-as-service* category explores the structured model for services offered in the Application Programming Interface (API), e.g., Vu et al. (2012).

We identify six categories in the STOF **organization domain**. The most-discussed category is *governance*, which broadly refers to governing processes by certain actors (e.g., platform owners) via several mechanisms, such as norms or power (Bevir, 2012). Examples of governance topics include discussion about policies and strategies in data marketplaces (Tupasela et al., 2020), construction of a data property protection system (Yu & Zhao, 2019), governance mechanisms in the platform design process (Otto & Jarke, 2019), self-regulation for fairness and transparency for data sharing (Richter & Slowinski, 2019). Moreover, the *social implications* category explores topics such as ethical concerns (e.g., Ahmed & Shabani, 2019; Ishmaev, 2020), implications of data trading for social, political, economic, and cultural context (Virkar et al., 2019), and challenges faced by private data marketplaces

A. Ewa Abbas, W. Agahari, M. van de Ven, A. Zuiderwijk &M. de Reuver:
Business Data Sharing through Data Marketplaces: A Systematic Literature Review

81

(Spiekermann et al., 2015). The next category is *data ecosystems,* which covers the topics of reviewing data ecosystems (Oliveira et al., 2019; W. Thomas & Leiponen, 2016) or exploring market mechanisms for data (Koutroumpis et al., 2020). The category of *classifying data marketplaces* by developing business model taxonomies (e.g., Fruhwirth et al., 2020; Stahl et al., 2016) is also examined. Finally, the selected research discusses topics such as the composition of the actors' population (Macdonald & Frank, 2017) and the geographical distribution of victim in stolen data markets (Smirnova & Holt, 2017) is categorized in the *demographic aspects*.

Finally, we identify three categories in the STOF **finance domain**. Articles in this domain are not equally distributed across categories because most scholars are interested in the *pricing mechanism*. Unlike the computational pricing (in the STOF technology domain) that focuses on technical aspects, the pricing mechanisms here discuss mathematical or economic approaches in evaluating, valuating, or pricing data in data marketplaces. The topic examples of this category are data trading models that consider privacy valuation (Oh et al., 2020), a debt-credit mechanism for data pricing (Liu et al., 2019), auction-based query pricing (Wang et al., 2019), and pricing mechanisms that aim for revenue or profit maximization (e.g., Mao et al., 2019). Other examples explored in this category are pricing mechanisms based on the Stackelberg game approach (Shen et al., 2019) and data quality scores for data pricing (Stahl & Vossen, 2016). Other categories in the finance domain are the *market analysis* that includes topics such as estimating the economic value of stolen data market (e.g., Holt et al., 2016; Shulman, 2010), as well as the *economic feasibility* category that examines Nash equilibria in competition between actors (Guijarro et al., 2019).

4   Discussion and Conclusion

This paper aims to investigate the current state of the art of data marketplace research. As shown in figure 1, we reveal that that data marketplace research is still primarily dominated by technical literature. Based on this fact, the pattern of evolution of data marketplace research tends to follow the *technology push* (i.e., technological advancement drives innovation). We argue that one possible reason behind this may be the availability of funding and projects that are intensely focused on the technological development of data marketplaces (refer to the description of

EU-funded projects on data markets[1]). Based on the platform's lifecycle by Henfridsson and Bygstad (2013), many of these projects are still in the initial phase of the platform's lifecycle, i.e., the platform design process. This may explain why debate focuses on technical rather than non-technical aspects. This phenomenon indirectly contributes to the limited options for publication venues. The conferences and journals that publish on data marketplaces are primarily found in the technical venues, e.g., the IEEE Access and IEEE Internet of Things Journal.

As indicated in the introduction section, data marketplaces are hardly commercially exploited, even though concepts have existed for years. Apparently, they struggle to move from the initial stage into the second stage of the platform's lifecycle, i.e., the platform adoption. One possible explanation could be that previous studies have not dealt extensively with non-technical topics. Hence, contributions from the academic perspective towards data marketplace commercialization are still scant. For example, little attention has been paid to topics categorized in the service domain (this domain was covered least by our studied papers). Based on business model ideas, this domain is essential and should be the starting point for data marketplaces to be commercially exploited. The topics in the service domain are essential to design services that fulfill customers' needs. Although a few attempts have been made to discuss relevant topics such as *value proposition*, many other topics such as *customer expected value* and *market segmentation* have barely been discussed in the selected articles.

Considering the organizational domain, one crucial overlooked aspect in current literature is *value networks* that describe actors and their interactions. It is essential to understand the dynamic to align their vision by developing *organizational arrangements* to achieve the common goal. For the finance domain, current literature mainly emphasizes data pricing. Future data marketplace research should cover other essential topics in the finance domain, such as *investments* and *cost sources,* because they are essential to build operating models of data marketplaces. Moreover, data marketplace projects are often conducted in a consortium based on academia-practitioners collaborations (e.g., the EU-funded projects). Academic publications may also reflect the work conducted by practitioners. Hence, considering non-technical investigations may open opportunities to speed up the platform adoption

---

[1] https://ec.europa.eu/digital-single-market/en/programme-and-projects/eu-funded-projects-data, accessed on February 19, 2021

A. Ewa Abbas, W. Agahari, M. van de Ven, A. Zuiderwijk &M. de Reuver:
Business Data Sharing through Data Marketplaces: A Systematic Literature Review

83

process in practice. We also suggest looking into the STOF model (Bouwman et al., 2008) for inspiration during the topic exploration process.

Our additional impressions after reading and analyzing the articles are as follows. We only found a few studies, e.g., Schomakers et al. (2020), Spiekermann and Korunovska (2017), that conduct empirical investigations in non-technical literature. Moreover, the many technology-focused studies hardly consider the link between technology solution and problem, such as is common in Design Science Research (DSR) approaches (Hevner & Chatterjee, 2010). Stronger links between technical solutions and value-related problems would help focus data marketplaces research so that practical problems are being resolved. Besides, the literature hardly discusses solutions to some core non-technical challenges of data marketplaces, such as: defining data ownership (Koutroumpis et al., 2020), assessing data quality (Koutroumpis et al., 2020), lacking legal frameworks (Richter & Slowinski, 2019), lacking technical expertise and resources to operate the ecosystem (Oliveira et al., 2019), and unclear organizational structure (Oliveira et al., 2019).

A limitation of this study is that the topic identification process is subject to the researchers' knowledge and interpretations about the topic, i.e., different readers may have different judgments. However, independently categorizing the present papers by different authors showed overall alignment. Moreover, the study is limited by its scope and the number of publications included in the analysis due to our criteria, e.g., a single database, the timeframe selection, and a paper quality check. Nonetheless, we argue that we have reached a sufficient level of saturation, i.e., analyzing the last couple of papers did not lead to new categories being identified or major shifts in the distribution of papers among categories.

This study contributes to the literature by a) providing a comprehensive overview of current data marketplace research and b) identifying neglected research topics that may contribute to data marketplaces' growth towards commercialization. We set out potential research topics to shift data marketplaces from the first stage of the platform's lifecycle, i.e., the platform design, to the second stage, i.e., the platform adoption. Our research provides the essential basis for future research towards the commercialization of data marketplaces. Finally, we call for (empirical) research in non-technological domains to complement the current technology-focused data marketplace research.

Acknowledgments

References

* [2]Agarwal, A., Dahleh, M., & Sarkar, T. (2019). A marketplace for data: An algorithmic solution. Proceedings of the 2019 ACM Conference on Economics and Computation.

*Ahmed, E., & Shabani, M. (2019). DNA Data Marketplace: An Analysis of the Ethical Concerns Regarding the Participation of the Individuals. Frontiers in Genetics, 10. https://doi.org/10.3389/fgene.2019.01107

*Anderson, N. G., Pollack, J., & Williams, D. (2014). The value of healthcare data in ophthalmology. Current Opinion in Ophthalmology, 25(3), 191-194. https://doi.org/10.1097/icu.0000000000000047

*Bevir, M. (2012). Governance : a very short introduction. Oxford University Press.

Bouwman, H., Faber, E., Haaker, T., Kijl, B., & De Reuver, M. (2008). Conceptualizing the STOF Model. In (pp. 31-70). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-79238-3_2

Champieux, R. (2015). PlumX. Journal of the Medical Library Association : JMLA, 103(1), 63-64. https://doi.org/10.3163/1536-5050.103.1.019

*Demchenko, Y., Los, W., & de Laat, C. (2018). Data as economic goods: Definitions, properties, challenges, enabling technologies for future data markets. ITUJournal-ICT Discoveries, 1(2).

European Commission. (2020). A European Strategy for Data https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN

Fruhwirth, M., Rachinger, M., & Prlja, E. (2020). Discovering Business Models of Data Marketplaces. Proceedings of the 53rd Hawaii International Conference on System Sciences.

*Guijarro, L., Pla, V., Vidal, J. R., & Naldi, M. (2019). Competition in data-based service provision: Nash equilibrium characterization. Future Generation Computer Systems, 96, 35-50. https://doi.org/10.1016/j.future.2019.01.044

*Hatanaka, H., & Abe, A. (2015). What Type of Information and Scheme does the Data Market Need? Procedia Computer Science, 60, 1309-1317. https://doi.org/10.1016/j.procs.2015.08.197

Henfridsson, O., & Bygstad, B. (2013). The generative mechanisms of digital infrastructure evolution. MIS quarterly, 907-931.

Hevner, A., & Chatterjee, S. (2010). Design science research in information systems. In Design research in information systems (pp. 9-22). Springer.

*Holt, T. J., Smirnova, O., & Chua, Y. T. (2016). Exploring and Estimating the Revenues and Profits of Participants in Stolen Data Markets. Deviant Behavior, 37(4), 353-367. https://doi.org/10.1080/01639625.2015.1026766

*Ishmaev, G. (2020). The Ethical Limits of Blockchain-Enabled Markets for Private IoT Data. Philosophy & Technology, 33(3), 411-432. https://doi.org/10.1007/s13347-019-00361-y

*Jeong, B.-G., Youn, T.-Y., Jho, N.-S., & Shin, S. U. (2020). Blockchain-Based Data Sharing and Trading Model for the Connected Car. Sensors, 20(11), 3141. https://doi.org/10.3390/s20113141

*Kamleitner, B., & Mitchell, V.-W. (2018). Can consumers experience ownership for their personal data? From issues of scope and invisibility to agents handling our digital blueprints. In Psychological ownership and consumer behavior (pp. 91-118). Springer.

---

[2] References with an asterisk indicate the inclusion of this article in the systematic literature review.

A. Ewa Abbas, W. Agahari, M. van de Ven, A. Zuiderwijk &M. de Reuver:
Business Data Sharing through Data Marketplaces: A Systematic Literature Review

85

Koutroumpis, P., Leiponen, A., & Thomas, L. D. W. (2020). Markets for data. Industrial and Corporate Change, 29(3), 645-660. https://doi.org/10.1093/icc/dtaa002

*Liu, K., Chen, W., Zheng, Z., Li, Z., & Liang, W. (2019). A Novel Debt-Credit Mechanism for Blockchain-Based Data-Trading in Internet of Vehicles. IEEE Internet of Things Journal, 6(5), 9098-9111. https://doi.org/10.1109/jiot.2019.2927682

*López, D., & Farooq, B. (2020). A multi-layered blockchain framework for smart mobility data-markets. Transportation Research Part C: Emerging Technologies, 111, 588-615. https://doi.org/10.1016/j.trc.2020.01.002

*Macdonald, M., & Frank, R. (2017). Shuffle Up and Deal: Use of a Capture–Recapture Method to Estimate the Size of Stolen Data Markets. American Behavioral Scientist, 61(11), 1313-1340. https://doi.org/10.1177/0002764217734262

*Mao, W., Zheng, Z., & Wu, F. (2019, 2019). Pricing for Revenue Maximization in IoT Data Markets: An Information Design Perspective.

*Morrison, N., Hancock, D., Hirschman, L., Dawyndt, P., Verslyppe, B., Kyrpides, N., Kottmann, R., Yilmaz, P., Glöckner, F. O., Grethe, J., Booth, T., Sterk, P., Nenadic, G., & Field, D. (2011). Data shopping in an open marketplace: Introducing the Ontogrator web application for marking up data using ontologies and browsing using facets. Standards in Genomic Sciences, 4(2), 286-292. https://doi.org/10.4056/sigs.1344279

*Munoz-Arcentales, A., López-Pernas, S., Pozo, A., Alonso, Á., Salvachúa, J., & Huecas, G. (2019). An Architecture for Providing Data Usage and Access Control in Data Sharing Ecosystems. Procedia Computer Science, 160, 590-597. https://doi.org/10.1016/j.procs.2019.11.042

*Niu, C., Zheng, Z., Wu, F., Gao, X., & Chen, G. (2019). Achieving Data Truthfulness and Privacy Preservation in Data Markets. IEEE Transactions on Knowledge and Data Engineering, 31(1), 105-119. https://doi.org/10.1109/tkde.2018.2822727

*Oh, H., Park, S., Lee, G. M., Choi, J. K., & Noh, S. (2020). Competitive Data Trading Model With Privacy Valuation for Multiple Stakeholders in IoT Data Markets. IEEE Internet of Things Journal, 7(4), 3623-3639. https://doi.org/10.1109/jiot.2020.2973662

Okoli, C. (2015). A guide to conducting a standalone systematic literature review. Communications of the Association for Information Systems, 37(1), 43.

*Oliveira, M. I. S., Lima, G. d. F. B., & Lóscio, B. F. (2019). Investigations into Data Ecosystems: a systematic mapping study. Knowledge and Information Systems, 1-42.

*Otto, B., & Jarke, M. (2019). Designing a multi-sided data platform: findings from the International Data Spaces case. Electronic Markets, 29(4), 561-580. https://doi.org/10.1007/s12525-019-00362-x

*Perera, C., Wakenshaw, S. Y., Baarslag, T., Haddadi, H., Bandara, A. K., Mortier, R., Crabtree, A., Ng, I. C., McAuley, D., & Crowcroft, J. (2017). Valorising the IoT databox: creating value for everyone. Transactions on Emerging Telecommunications Technologies, 28(1), e3125.

*Richter, H., & Slowinski, P. R. (2019). The Data Sharing Economy: On the Emergence of New Intermediaries. IIC - International Review of Intellectual Property and Competition Law, 50(1), 4-29. https://doi.org/10.1007/s40319-018-00777-7

*Robinson, S. C. (2017). What's your anonymity worth? Establishing a marketplace for the valuation and control of individuals' anonymity and personal data. Digital Policy, Regulation and Governance, 19(5), 353-366. https://doi.org/10.1108/dprg-05-2017-0018

*Roman, D., & Stefano, G. (2016, 2016). Towards a Reference Architecture for Trusted Data Marketplaces: The Credit Scoring Perspective.

*Schomakers, E.-M., Lidynia, C., & Ziefle, M. (2020). All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. Electronic Markets, 1-17.

*Shaabany, G., Grimm, M., & Anderl, R. (2016). Secure Information Model for Data Marketplaces Enabling Global Distributed Manufacturing. Procedia CIRP, 50, 360-365. https://doi.org/10.1016/j.procir.2016.05.003

*Shen, B., Shen, Y., & Ji, W. (2019). Profit optimization in service-oriented data market: A Stackelberg game approach. Future Generation Computer Systems, 95, 17-25. https://doi.org/10.1016/j.future.2018.12.072

*Shulman, A. (2010). The underground credentials market. 2010(3), 5-8. https://doi.org/10.1016/s1361-3723(10)70022-1

*Smirnova, O., & Holt, T. J. (2017). Examining the Geographic Distribution of Victim Nations in Stolen Data Markets. American Behavioral Scientist, 61(11), 1403-1426. https://doi.org/10.1177/0002764217734270

Solaimani, S., Keijzer-Broers, W., & Bouwman, H. (2015). What we do – and don't – know about the Smart Home: An analysis of the Smart Home literature. Indoor and Built Environment, 24(3), 370-383. https://doi.org/10.1177/1420326x13516350

*Sørlie, J.-T., & Altmann, J. (2019). Sensing as a Service Revisited: A Property Rights Enforcement and Pricing Model for IIoT Data Marketplaces. International Conference on the Economics of Grids, Clouds, Systems, and Services.

Spiekermann, M. (2019). Data Marketplaces: Trends and Monetisation of Data Goods. Intereconomics, 54(4), 208-216. https://doi.org/10.1007/s10272-019-0826-z

*Spiekermann, S., Acquisti, A., Böhme, R., & Hui, K.-L. (2015). The challenges of personal data markets and privacy. Electronic Markets, 25(2), 161-167. https://doi.org/10.1007/s12525-015-0191-0

*Spiekermann, S., & Korunovska, J. (2017). Towards a value theory for personal data. Journal of Information Technology, 32(1), 62-84. https://doi.org/10.1057/jit.2016.4

Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. Vietnam Journal of Computer Science, 3(3), 137-143. https://doi.org/10.1007/s40595-016-0064-2

*Stahl, F., & Vossen, G. (2016). Data Quality Scores for Pricing on Data Marketplaces. In (pp. 215-224). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49381-6_21

*Truong, H. L., Comerio, M., Paoli, F. D., Gangadharan, G. R., & Dustdar, S. (2012). Data contracts for cloud-based data marketplaces. International Journal of Computational Science and Engineering, 7(4), 280. https://doi.org/10.1504/ijcse.2012.049749

*Tupasela, A., Snell, K., & Tarkkala, H. (2020). The Nordic data imaginary. Big Data & Society, 7(1), 205395172090710. https://doi.org/10.1177/2053951720907107

Virkar, S., Viale Pereira, G., & Vignoli, M. (2019). Investigating the Social, Political, Economic and Cultural Implications of Data Trading. In (pp. 215-229). Springer International Publishing. https://doi.org/10.1007/978-3-030-27325-5_17

*Vu, Q. H., Pham, T.-V., Truong, H.-L., Dustdar, S., & Asal, R. (2012). Demods: A description model for data-as-a-service. 2012 IEEE 26th International Conference on Advanced Information Networking and Applications.

*W. Thomas, L. D., & Leiponen, A. (2016). Big data commercialization. IEEE Engineering Management Review, 44(2), 74-90. https://doi.org/10.1109/emr.2016.2568798

*Wang, X., Wei, X., Gao, S., Liu, Y., & Li, Z. (2019). A novel auction-based query pricing schema. International Journal of Parallel Programming, 1-22.

*Yu, H., & Zhang, M. (2017). Data pricing strategy based on data quality. Computers & Industrial Engineering, 112, 1-10. https://doi.org/10.1016/j.cie.2017.08.008

*Yu, X., & Zhao, Y. (2019). Dualism in data protection: Balancing the right to personal data and the data property right. Computer Law & Security Review, 35(5), 105318. https://doi.org/10.1016/j.clsr.2019.04.001

*Zhang, R., Indulska, M., & Sadiq, S. (2019). Discovering Data Quality Problems. Business & Information Systems Engineering, 61(5), 575-593. https://doi.org/10.1007/s12599-019-00608-0

*Zhao, Y., Yu, Y., Li, Y., Han, G., & Du, X. (2019). Machine learning based privacy-preserving fair data trading in big data market. Information Sciences, 478, 449-460. https://doi.org/10.1016/j.ins.2018.11.028

*Zheng, Z., Peng, Y., Wu, F., Tang, S., & Chen, G. (2020). ARETE: On Designing Joint Online Pricing and Reward Sharing Mechanisms for Mobile Data Markets. IEEE Transactions on Mobile Computing, 19(4), 769-787. https://doi.org/10.1109/tmc.2019.2900243

TELEMEDICINE IN SLOVENIA

ŽIVA RANT

National Institute of Public Health, Ljubljana, Slovenia; e-mail: Ziva.Rant@nijz.si

Abstract Telemedicine could be one of the solutions for challenges in healthcare, especially in this time of the Covid-19 pandemic. The results of the research about the state of telemedicine services in Slovenia are presented in this article. We found several telemedicine solutions in Slovenian healthcare. Metadata for them were collected. The solutions are placed in groups of telemonitoring, provision of healthcare services by remotely connecting patients with a doctor or healthcare professional and remote cooperation for the patient's treatment between doctors or healthcare professionals who are physically at different locations. The opinions of the research participants regarding the challenges associated with telemedicine services in Slovenia were also collected. They are placed in three main groups: financing, healthcare system and healthcare professionals. Telemedicine is a necessity and the future of Slovenian healthcare services. If a solution is to be applied successfully, business processes must be changed so that a practically useful service can arise from the solution.

1   Introduction

European citizens are getting older and are increasingly living with chronic diseases. Their health condition often requires enhanced medical attention. Medical support may not be available in remote areas and for certain specialities as easily or as frequently as their health condition would require. (COM(2008)689). So changes in health services are needed. On the other hand, the improvement on the information and communications technology (ICT) field is enormous. It could be successfully used for making changes in healthcare services. Some healthcare treatment can be done remotely. Telemedicine is a proper solution for these challenges, especially in this time of the Covid-19 pandemic.

Telemedicine (TM) is the provision of healthcare services, through use of ICT, in situations where the health professional and the patient (or two health professionals) are not in the same location. It involves secure transmission of medical data and information, through text, sound, images or other forms needed for the prevention, diagnosis, treatment and follow-up of patients. (COM(2008)689).

Although there are several telemedicine services available in Slovenia, no overview of them is accessible. As suggested by one of the members, the Governing board (GB) of the Slovenian Medical Informatics Association (SDMI) has decided on making the research about usage TM solutions in Slovenia.

2   Methodology

The purpose of the research was to learn about the state of telemedicine in Slovenia. The research question was: What is the state of telemedicine in Slovenia?   The research was carried out in May and June 2020, after the first wave of the Covid-19 pandemic. In the first phase members of GB SDMI were asked to find and report the TM solutions. Members of GB SDMI were selected because they are from many medical information fields: universities, institutes, healthcare providers, software developers and other experts from the medical information field. They geographically covered all Slovenia. So we could get a very broad view about the selected topic.

With the aim to collect data we prepared a structured questionnaire. In the second phase we collected key metadata for each solution: name of the solution, members of the development, contact person and contact data, target users, users amount assessment, short solution description, potential for usage on the national level. We also asked for respondents' opinions of three most important challenges and note the person who provided the data. The data was gathered and entered in the same table by the healthcare providers' members, software development companies, experts from the Slovenian universities and the national institute, and other experts from the medical information field. We published the organized table of the answers on the SDMI web page (SDMI, 2020). These answers we analyzed further and the results of the analysis are presented in this article.

3   Results

As part of the research, we collected metadata for 15 telemedicine solutions in Slovenian healthcare. We placed them in the following groups:

- telemonitoring (model Business to Customer - B2C),
- provision of healthcare services by remotely connecting patients with a doctor or healthcare professional (model Business to Customer - B2C),
- remote cooperation for the patient's treatment between doctors or healthcare professionals who are physically at different locations (model Business to Business - B2B).

A description of the solutions and their associated services is given below.

3.1   Telemonitoring

We have found five solutions for telemonitoring.

Telemonitoring is a telemedicine service aimed at monitoring the health status of patients at a distance. Data can be collected either automatically through personal health monitoring devices or through active patient collaboration (e.g. by entering weight or daily blood sugar level measurements into a web-based tool). Data, once processed and shared with relevant health professionals, may be used to optimize the patient's monitoring and treatment protocols (COM(2008)689).

Storitve centra za zdravje na daljavo CEZAR (*TM support service provided by the Regional Centre for Telehealth (CEZAR)*) (Rudel et al, 2016 Zdrav Vest). The solution was developed by General Hospital Slovenj Gradec, Healthcare Centre Ravne and MKS Electronic Systems Ltd., Ljubljana. It is made for telemonitoring and chronic disease management for patients with diabetes mellitus type 2 or chronic congestive heart failure (SDMI, 2020).

Sistem za oddaljeno spremljanje in telemedicinsko obravnavo pacientov (SOSTOP) (*The system for patient telemonitoring and telemedicine treatment*) (SOSTOP, n.d.) was developed by Ipmit d.o.o., Nova Vizija d.d., Medicina Iljaž d.o.o. and Faculty of Computer and Information Science, Ljubljana. It is made for provision of healthcare services at a distance in family medicine reference clinics for patients with chronic illnesses (SDMI, 2020).

T-MED Gluco was developed by MKS d.o.o., Ljubljana and VPD Bled d.o.o. It is made for provision of healthcare services at a distance for patients with diabetes mellitus type 2 and the diabetes clinics (SDMI, 2020).

E-zdravje Telekoma Slovenije (*eHealth Telekom Slovenia)* (Oroszy & Pustatičnik, n.d.) was developed by Telekom Slovenije, d.d., University Medical Centre Ljubljana (UMCL), University Clinic Golnik and National Institute of Public Health Slovenia (NIPH). It is made for remote monitoring of patients with chronic illnesses. It is also used by the UMCL for individual Covid-19 patients in isolation at home (The Fast Mode, n.d.) (SDMI, 2020).

E-oskrba (*E-care Service*) (Ekosmart, n.d.) was developed by Telekom Slovenije, d.d. and the Slovene Federation of Pensioners' Associations (ZDUS). It enables active, independent and safe living at home to the elderly, patients with chronic diseases, and to disabled persons (SDMI, 2020).

3.2   Remotely connecting patients with a doctor or healthcare professional

We have found four solutions for provision of healthcare services by remotely connecting patients with a doctor or healthcare professional. They are made by the software developer companies for the healthcare providers. The patient can communicate with the healthcare professional safely from a distance. The solutions

provide online scheduling of appointments, ordering electronic prescriptions, electronic referrals and electronic sick leave certificates. Some of them enable safely sharing specialist reports, radiology images and also electronic or even video consultations for distance treatments.

doZdravnika.si (doZdravnika.si, n.d.) was developed by SRC Infonet. (SDMI, 2020).

Hipokrat - eSodelovanje was developed by LIST d.o.o. (SDMI, 2020).

PriZdravniku (Modul PriZdravniku, n.d.), (Storitev eVideoPosvet, n.d.) was developed by Nova Vizija d.d. (SDMI, 2020).

Gospodar zdravja (Gospodar zdravja, n.d.) was developed by Gospodar zdravja d.o.o. (SDMI, 2020).

3.3 Remote cooperation between doctors or healthcare professionals who are physically at different locations

Seven solutions were found for remote cooperation for the patient's treatment between doctors or healthcare professionals who are physically at different locations.

Teletransfuzija (Teletransfuzija, n.d.) was developed by the Blood Transfusion Centre of Slovenia and XLAB, d.o.o. Telemedicine is used to perform mandatory pre-transfusion tests from a distance. From their own location a specialist in transfusion medicine examines, reads, orders further tests and authorises the issue of blood at other locations, where a medical laboratory scientist carries out pre-transfusion tests. The computer teleconsultation system covers the entire blood transfusion service in Slovenia since 2005 (SDMI, 2020).

TeleFarma (Telefarma, n.d.) was developed by The Faculty of Medicine at the University of Ljubljana (UL MF), SRC Infonet, General Hospital Murska Sobota and Healthcare Centre Škofja Loka. TeleFarma enables all physicians, regardless of their location, equal access to clinical pharmacists and offers better cooperation between physicians and clinical pharmacists (SDMI, 2020).

There are three telemedicine solutions in Slovenian eHealth under the governance of National Institute of Public Health (NIPH) Slovenia.

Telekap (*TeleStroke*) (Nacionalni inštitut za javno zdravje, NIJZ, n.d.) was developed by Interexport d.o.o., The Division of Neurology, UMCL, Ministry of Health of RS and NIPH Slovenia. The medical specialist from The Division of Neurology supports doctors from other hospitals if a patient is suspected of having a stroke. It takes place through a high quality video-conference, accessible 24 hours a day (SDMI, 2020).

Teleradiologija (*Teleradiology*) (NIJZ, n.d.) was developed by Interexport d.o.o., Ministry of Health of RS and NIPH Slovenia. It is a telemedicine service which involves the electronic transmission of radiographic images from one geographical location to another for the purposes of interpretation and consultation (SDMI, 2020).

ePosvet (*eConsultation*) (NIJZ, n.d.) was developed by IN2 d.o.o., Jesenice General Hospital and Osnovno zdravstvo Gorenjske. It enables consultation between general practitioners and medical specialists about a specific patient with the aim to reduce waiting periods and expedite the treatment process (SDMI, 2020).

Mobilni zdravnik (*Mobile doctor*) was developed by Nova Vizija d.d. It enables sharing medicine data between a nurse and a doctor, who is not in the clinic (SDMI, 2020).

3.4 Challenges

We also collected the opinions of the research participants regarding the challenges associated with telemedicine services in Slovenia. These opinions are analysed below.

3.4.1 Financing

Funding was the problem most commonly highlighted by respondents. First and foremost, this is funding for telemedicine services provided by healthcare providers. The service is not paid for by Health Insurance Institute of Slovenia (ZZZS), which means that providers cannot be remunerated. It is important for service provision

to be financially sustainable. The process of accepting new services for funding at the national level needs to be accelerated. Other providers need to be given funding for consultations, and funding also needs to be found for measuring devices and other equipment that the patient receives for measurement purposes. The outdated information and communication equipment in hospitals is also problematic and needs to be replaced.

3.4.2   Healthcare system

The respondents highlight the lack of awareness of the importance and potential of services on the part of decision-makers, and the rigidity of the healthcare system when it comes to accepting new services. The profile of the telemedicine system must be raised. Telemedicine services should be linked to national eHealth solutions. There must be cooperation between stakeholders in Slovenia. Interoperability and mobility must be ensured, standards introduced and materials standardised. There is also a lack of trust in the new communication methods. A telemedicine strategy is urgently required in Slovenia, additions must be made to the foundations of the system and to the systemic support for telemedicine and telecare, and standards put in place for the provision of services. The security aspect is also extremely important. New healthcare providers must be brought into the system. Equal access must be ensured for everyone who requires services.

3.4.3   Healthcare professionals

The research highlights the lack of awareness of the importance and potential of services on the part of healthcare professionals, and the readiness of healthcare professionals to accept new services. Motivation among healthcare workers is very low and they do not see use of telemedicine as a professional challenge. Healthcare professionals are often unaware that using these services can make their work significantly easier and reduce the amount of time required to treat a patient, and that freeing up telephone lines is vital. While they find the decision to start using telemedicine a difficult one, once they overcome their resistance, they never want to go back. After the first module of the solution is used, it swiftly spreads throughout the institution or into new functionalities. Healthcare professionals are afraid that patients are able to book a specific slot in their timetable without the intervention of a nurse.

4 Discussion

In our analysis of the state of telemedicine services in Slovenia, we focused on an unresearched area, which meant that quantitative empirical methods could not be used. Our assessment was that a thorough analysis is the most suitable methodological approach to researching and understanding this area.

Slovenia has a long telemedicine tradition stretching back to the 'Rdeči gumb' (Red button) system in 1992. Quite a large number of telemedicine solutions were developed (mainly as part of research projects), but not all of them remain in use. One of the reasons for this is most definitely the fact that the successful introduction of telemedicine services requires a change to business processes. (Rant, 2009) (Rant, 2010). A new service is therefore designed and then put into practice. Examples of good practice are Teletransfusion, which has been successfully in place since 2005 and enables the national remote interpretation of pre-transfusion blood tests, and the distance telemedicine monitoring of chronic heart failure and type 2 diabetes patients at the CEZAR Centre for Remote Health at Slovenj Gradec General Hospital (since 2014). The use of telemedicine services is increasing in Slovenia, which is also the result of the Covid-19 pandemic. This is also shown by the operations of the Telemedicine Centre at Ljubljana University Medical Centre, which was established during the pandemic and has provided Covid-19 patients with telemedicine treatment in cooperation with clinical departments.

Scientific and professional work in the area of telemedicine has also been under way in Slovenia for quite a few years. The Slovenian Medical Informatics Society (SDMI) organised an expert conference Telemedicina – zdravje na daljavo ('Telemedicine – Distance Health') in Ptuj in 2010. Although members of the SDMI drew up the premises for the preparation of a national distance health strategy in 2012, no strategy has yet been forthcoming.

There is scant official data on the substantive and financial benefits and weaknesses of telemedicine (Oroszy, 2020), (Rudel et al, 2016 Zdrav Vest). Telemedicine does definitely help to reduce waiting lists and times, the number of hospitalisations, the duration of hospital stays and the impact on patients, which in turn leads to savings in healthcare expenditure and increases the patient's quality of life. They return to work sooner, which also reduces expenditure on sick leave. Instead of using up

resources as a result of their absence from work, the patient is able to start creating value again. These solutions also help to relieve the burden on telephone lines and therefore make doctors more accessible. Doctors are less occupied with check-ups and therefore have more time to devote to those patients who need them more.

One of the main weaknesses is the funding and financial sustainability of telemedicine provision. However, moves are being made in this area as well, with the Health Insurance Institute of Slovenia (ZZZS) funding some services in 2020 and 2021.

Those responsible for developing the services mention the lack of awareness of the importance and potential of telemedicine services on the part of decision-makers and healthcare professionals, and the rigidity of the healthcare system when it comes to introducing new services. Positive promotion must therefore be undertaken, particularly in terms of raising awareness and providing training for healthcare professionals, especially doctors. We have to realise that telemedicine entails a fundamental change to the way we normally treat patients. Business processes often have to be changed and completely reformulated in the face of a brand new service Doctors and other healthcare professionals who wish to become involved must also be included in the process of overhauling business processes. It seems to be valued to repeat the research after the Covid-19 pandemic.

5   Conclusion

Telemedicine is a necessity and the future of the Slovenian healthcare services. It helps patients to a better quality of life and brings financial benefits to the healthcare system. Positive promotion, education and training for healthcare professionals must be undertaken. Systemic regulation is required, as is the preparation of policies, strategies and standards at national level. Financial sustainability is needed. It is also very important to take into account the security aspect, as telemedicine processes particular types of personal data. One needs to realise that information and communication technology is not enough in itself, and nor are telemedicine solutions. If a solution is to be applied successfully, business processes must be changed so that a practically useful service can arise from the solution.

References

doZdravnika.si (n.d.). doZdravnika.si. Retrieved February 2, 2021, from
    https://www.infonet.si/produkti/dozdravnika-si/,
    https://www.infonet.si/en/blog/vaccination-dozdravnika/

Ekosmart (n.d.). Program Smart System of Integrated Health Care and Home Care. Retrieved February
    2, 2021 from http://ekosmart.net/en/2018/06/13/ekosmart-program-presentation-at-the-
    hope-agora-2018-in-stockholm/

European Commission (2008). Communication from the Commission to the European Parliament,
    the Council, the European Economic and Social Committee and the Committee of the Regions
    on telemedicine for the benefit of patients, healthcare systems and society (COM(2008)689).
    Retrieved January 28, 2021 from
    https://op.europa.eu/en/publication-detail/-/publication/4e8fb922-24ab-46f6-b1a1-
    3c3b2d93d8de/language-en

Gospodar zdravja (n.d.). Retrieved January 28, 2021 from www.gospodar-zdravja.si

Oroszy, D. (2020). Telemedicinski center UKC Ljubljana (T. Marčun, Ed.). Digitalni mostovi v
    zdravstvu, pp. 21.

Oroszy D., Pustatičnik, P. (n.d.). Smart System of Integrated Health Care and Home Care. Retrieved
    February 2, 2021 from http://www.hope.be/wp-content/uploads/2019/06/PON-LD-11.30-
    Oroszy.pdf

Rant, Ž. (2009). Nove tehnologije prinašajo nove izzive tudi pri poslovnih procesih (V. Rajkovič, Ed.).
    Nove tehnologije, novi izzivi, str. 1237-1244. Kranj: Moderna organizacija.

Rant, Ž. (2010). Poslovni procesi v telemedicini = Business processes in telemedicine. Informatica
    Medica Slovenica, Vol. 15, str. 39. Maribor: Zdravstveni dom dr. Adolfa Drolca.

Rudel D, Slemenik-Pušnik C, Epšek-Lenart M, Pušnik S, Balorda Z, Lavre J. (2016). Telemedicine
    support to patients with chronic diseases for better long-term control at home. Zdrav Vestn.
    85, str. 676–85.

Nacionalni inštitut za javno zdravje - NIJZ (n.d.). eZdravje. Retrieved February 2, 2021, from
    https://www.ezdrav.si/

Modul PriZdravniku (n.d) Retrieved January 28, 2021 from https://www.vizija.si/medicinska-
    informatika/modul-prizdravniku/

SDMI, m–. (2020). Seznam telemedicinskih rešitev v slovenskem zdravstvu. Ljubljana: Slovensko
    društvo za medicinsko informatiko. Retrieved January 28, 2021 from
    https://sdmi.si/files/strokovna_srecanja/Telemedicina%20SDMI%202020.pdf

SOSTOP (n.d.). Sistem za oddaljeno spremljanje in telemedicinsko obravnavo pacientov (SOSTOP).
    Retrieved January 28, 2021 from https://www.vizija.si/medicinska-informatika/sostop/

Storitev eVideoPosvet (n.d.). Retrieved February 1, 2021, from https://www.vizija.si/medicinska-
    informatika/storitev-evideoposvet/

Telefarma (n.d.). Retrieved February 9, 2021, from https://www.infonet.si/en/blog/telefarma/

Teletransfuzija. (n.d.). Retrieved February 1, 2021, from
    http://www.ztm.si/en/research/development-education/it--projects/telemedicine-in-blood-
    transfusion/

The Fast Mode (n.d). Slovenia's Telekom Slovenije Develops Telehealth Solution to Monitor COVID-
    19 Patients Remotely. Retrieved February 2, 2021, from
    https://www.thefastmode.com/services-and-innovations/16789-slovenias-telekom-
    slovenije-develops-telehealth-solution-to-monitor-covid-19-patients-remotely

Students' Satisfaction with e-Learning Systems During the COVID-19 Pandemic—An International Comparative Study

Shahrokh Nikou[1] & Seongcheol Kim[2]

[1] Åbo Akademi University, Faculty of Social Sciences, Business and Economics, Turku, Finland, and Stockholm University, Stockholm, Sweden; e-mail: shahrokh.nikou@abo.fi
[2] Korea University, Seoul, Korea; email: hiddentrees@korea.ac.kr

Abstract In response to the global COVID-19 situation, quarantine measures have been implemented at the educational institutions around the world. This paper aims to determine the antecedent factors predicting the university students' satisfaction with e-learning systems during the COVID-19 situation. We used structural equation modelling (SEM) and evaluated a conceptual model on the basis of a sample of university students from Finland (n = 131) and South Korea (n = 114). The SEM results showed that the COVID-19 related factors, i.e., COVID-19 awareness, perceived challenges during COVID-19 and the educational institutions' preparedness indirectly influence the satisfaction with e-learning systems. Moreover, we found a statistically significant moderating effect of course design quality, and instructor's teaching style between the COVID-19 related factors and the satisfaction with e-learning systems. The implications of these results for the management of e-learning systems are discussed.

1   Introduction

As a result of the introduction of emerging digital technology, new possibilities for learning and teaching have emerged (Aavakare and Nikou, 2020). Higher education institutions use information and communications technology (ICT) to deliver contents for education and learning (Nikou and Aavakare, 2021). Due to COVID-19 pandemic, e-learning has become an emerging paradigm of modern education once again (Arafat et al., 2020; Sun et al., 2008). E-learning relies on the use of advanced digital technologies such as learning management systems (e.g., Moodle) to deliver learning materials and educational content. Given the relatively recent situation in terms of COVID-19 pandemic worldwide, e-learning and the use of learning management systems (LMS) have become increasingly important and a natural tool for providing distance learning/education (Radha et al., 2020). Literature shows that different factors impact students' satisfaction with e-learning systems. The factors include course design quality and instructors' teaching style, learning style (Lu and Chiou, 2010), content and interface of e-learning (Al-Rahmi et al., 2015), instruction medium and course content (Peng and Samah, 2006).

However, in relation to COVID-19, new factors such as COVID-19 awareness, perceived challenges during COVID-19 and educational institutions' preparedness have emerged, demonstrating a significant impact on the satisfaction of students with e-learning (Alea et al., 2020; Nikou and Maslov, 2021). Since students' satisfaction with e-learning system has a significant impact on the intention to use learning educational tools (Ramayah and Lee, 2012), the aim of this paper is to explore factors influencing the satisfaction of university students with e-learning systems. More importantly, we aim to explore factors, which are related to the COVID-19 pandemic. We argue that more research is needed to better understand the underlying impact of COVID-19 related factors in higher education. Hence, we conduct a comparative empirical research, collecting data from Finnish and South Korean university students to address this issue.

The research questions guiding this study are: "in relation to the COVID-19 pandemic, what factors are associated with the students' satisfaction with e-learning systems?" and "What are the similarities and differences between Finnish and South Korean university students in terms of their satisfaction with e-learning?" To address these RQs, we develop an integrated theoretical model that encompasses factors in

relation to COVID-19 situation and some other general factors and examine the model through structural equation modelling (SEM).

2   Literature Review and Hypotheses Development

E-learning, widely employed in many educational institutions, is an activity of self-learning and a paradigm of teaching and learning to complement face-to-face learning (Aboagye et al., 2021). It facilitates active and independent learning of the learners and provides many opportunities for self-learning. Several factors such as efficiency, reliability and quality of e-learning systems (Almaiah et al., 2020), accessibility and academic issues and the learner motivation (Aboagye et al., 2021) contribute to the usage of such systems and ultimately students' satisfaction level. Specifically, during COVID-19 pandemic where there is almost no physical presence nor social interactions between learners and instructors (Almaiah et al., 2020), the effect of such factors might become even more important. Alea et al. (2020) showed that there are multiple challenges in terms of the educational preparedness to facilitate distance and independent learning during the COVID-19. Moreover, ICT and Internet connection used by the educational institutions for e-learning may become unreliable and unable to meet the requirement for e-learning in this pandemic situation (Favale et al. 2020). Thus, instructors may have to adopt new teaching style to comply with the limitation and resections imposed by the current situation. As such, we discuss the factors associated with COVID-19 and linked to educational institutions, which are assumed to influence satisfaction with e-learning systems.

2.1   Instructors' Teaching Style

In the higher education, the instructors' teaching style is considered to be central in the success of the e-learning education. Al-Busaidi and Al-Shihi (2010, p. 1) argued that the success of e-learning relies on the instructors' acceptance of the learning management systems, which in turns promotes learners to use LMSs. Moreover, Volery and Lord (2000) argued that, for the students' satisfaction with e-learning systems, instructors must have several competences such as a good control over IT used for teaching and learning as well as to possess sufficient technical abilities to solve potential students' IT-related technical issues. Moreover, an instructor is

required to adopt interactive teaching style and encourage students to interact with their peers, hence, we make a hypothesis that:

H1: *Instructors' teaching style has a positive effect on the students' satisfaction with e-learning systems*

2.2   Course Design Quality

The quality of the course design depends on what is involved in the course, such as course data, educational goals, course layout (Wrigh, 2003). It has been found that course design quality might influence the satisfaction with the e-learning systems (Martín-Rodríguez et al., 2015). Liu and Chu (2010) argued that design quality can be used as a measure of the information quality and the course content quality. Moreover, Uppal et al. (2018) and Garavan et al. (2010) showed that the supportiveness of the overall service, information quality, system quality, content quality and learner support are different aspects of e-learning quality; thus, impacting the use of and the satisfaction with e-learning systems. Hence, we make a hypothesis that:

H2: *Course design quality has a positive effect on the students' satisfaction with e-learning systems*

2.3   COVID-19 Related Factors

Alea et al. (2020) examined the perception of teachers about the preparedness and challenges faced by higher education institutions when e-learning is implemented and found several antecedent factors including (i) COVID-19 awareness, (ii) the educational institutions preparedness to conduct distance learning, and (iii) perceived challenges during COVID-19 in distance learning education. In this study, based on literature discussions provided earlier, and following the recent study conducted by Alea et al. (2020), we use all three COVID-19 related constructs to examine the students' satisfaction with e-learning systems. In addition, in higher education, factors associated with COVID-19 are understood both as factors associated with the individual's background, requiring one to engage solely in distant e-learning, and as intermediate factors that affect how the e-learning process is conducted. Thus, we include other factors such as (i) instructors' teaching style and (ii) course design quality in our proposed conceptual model. Hence, we make the following hypotheses that:

H3a: *COVID-19 awareness has a positive effect on the instructors' teaching style*

H3b: *COVID-19 awareness has a positive effect on the course design quality*

H4a: *Perceived challenges during COVID-19 has a negative effect on the instructors' teaching style*

H4b: *Perceived challenges during COVID-19 has a negative effect on the course design quality*

H5a: *The educational institutions' preparedness has a positive effect on the instructors' teaching style*

H5b: *The educational institutions' preparedness has a positive effect on the course design quality*

However, in all the above stated hypotheses and as a null hypothesis, we assume that there is no difference between Finnish and South Korean students in any of the measured factors/constructs.

2.4   Satisfaction with e-Learning Systems

Student–student interaction, effective support, learning materials, teaching style, education and learning environment all can influence students' satisfaction with e-learning systems (Benigno and Trentin, 2000). Moreover, Almaiah et al. (2020), asserted that the provision and usage of online learning materials in e-learning system becomes the main challenge for many universities during COVID-19 pandemic. The authors further demonstrate that, there are four general challenges in relation to the use of e-learning systems and students' performance and consequently their satisfaction: technological challenges, individual challenges, cultural challenges and course challenges. In our proposed conceptual model, students' satisfaction with e-learning systems is the dependent variable (see Figure 1).
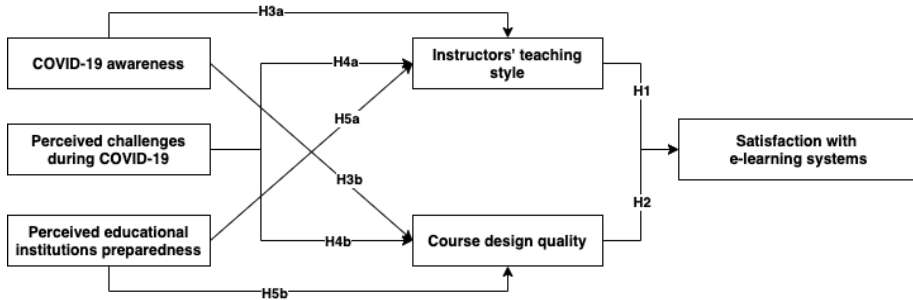
Figure 1: Conceptual Model.

3 Research Methodology

3.1 Data Collection and Instrument

This is an international comparative study and we collected data from Finland and South Korea. The main reason for selecting these two countries is due to the fact that they are frontrunner in using digital technology in their educational systems, in addition to many other similarities in the use of advanced technologies and digital infrastructures (Jang et al., 2021). We collected data only from university students, two university from Finland and one university from South Korea. Students were from difference subjects such as social sciences and natural sciences. The Finnish data was collected in August 2020 (n = 131) and The Korean data was collected in January 2021 (n = 114). We used an online survey to collect data from both countries. All survey items were derived from validated measures supported by literature.

The items have been slightly changed to fit the study context, if needed. Items for measuring COVID-19 awareness (three items), perceived challenges during COVID-19 (four items) and educational institutions preparedness to conduct distance learning (six items) all were derived from Alea et al. (2020, p. 134-136). Items for measuring course design quality (3 items) and instructors' teaching style (4 items) were derived from Wright (2003). Finally, items for measuring students' satisfaction with e-learning (four items) were derived from Arbaugh (2000, p. 41).

4 Results

4.1 Descriptive Results

The average age of the respondents was 25.8 years-old for the Finnish sample (26.55 years-old for the South Korean sample). In the Finnish sample, there were 73 females and 56 males, and two students preferred not to indicate their gender. In the South Korean sample there were 69 females and 45 males. The use of e-learning systems in the Finnish sample was (< 1 a year n = 61), (1-3 years n = 37), and (> three years n = 32) and one student indicated never used it, whereas the use of e-learning systems in the South Korean sample was (< a year n = 31), (1-3 years n = 51), and (> three years n = 32).

We used PLS-SEM to assess the path relationships proposed in our research model. PLS-SEM results showed that all factor loadings (except for few items) were above the recommended value of .70. In total, we used 24 items to measure the six constructs and retained 19 items for further analysis. All internal reliability and validity assessments such as Cronbach's alpha (α), Composite Reliability (CR) and the Average Variance Extracted (AVE) for all constructs were consistent with the recommended threshold values of .70, .70 and .50, respectively (Hair et al., 2019). However, the observed slightly low value of Cronbach's alpha (α) for COVID-19 awareness (.61). However, as Cronbach's alpha is a very conservative test, the CR value should instead be used to assess the internal reliability. The result showed the CR value for the COVID-19 awareness satisfied the recommend value, thus we establish the internal validity for this construct. The lowest value of CR was for COVID-19 awareness (.80) and the highest for satisfaction with e-learning systems (.93). Regarding the AVE values for the constructs, the lowest value was .66 and the highest was .80, see Table 1.

Table 1: Descriptive Statistics

Construct	Items	Loading	Mean	Std.	α	CR	AVE
COVID-19 awareness	CVID2	.82	6.73	.80	.61	.80	.66
	CVID3	.81	6.48	1.01			
Instructors' teaching style	ITS2	.86	4.42	1.58	.76	.86	.68
	ITS3	.77	4.88	1.48			
	ITS4	.84	4.45	1.63			
Course design quality	CDES1	.84	4.91	1.74	.82	.89	.74
	CDES2	.90	4.79	1.47			
	CDES3	.83	4.1	1.66			
Perceived challenges during COVID-19	PCHA2	.79	4.81	1.79	.72	.84	.64
	PCHA3	.88	4.82	1.95			
	PCHA4	.72	5.18	1.85			
Educational institutions preparedness	PEIP2	.88	4.59	1.74	.88	.91	.67
	PEIP3	.86	4.81	1.66			
	PEIP4	.81	4.91	1.67			
	PEIP5	.77	4.77	1.72			
	PEIP6	.75	4.6	1.74			
Satisfaction with e-learning	ESAT2	.91	3.98	1.83	.88	.93	.80
	ESAT3	.92	3.71	1.87			
	ESAT4	.86	3.42	1.84			

4.2   Convergent validity and discriminant validity

We also assessed the convergent validity to make sure that all measures within each construct which are theoretically expected to relate to one another, were in fact related to each other. Regarding the convergent validity, the values of average variance extracted (AVE) were used to establish the convergent validity in this research. As shown in Table 2, all the AVE values were above the recommended threshold of .50. As per discriminant validity assessment to establish that a construct is different from other constructs, we used the Fornell Larcker criterion. As such, we assessed the AVE scores, all values were lower than the shared variance for all model constructs, see Table 2. Therefore, the discriminant validity was established in this research based on Fornell Larcker criterion (Fornell and Larcker 1981).

Table 2: Discriminant Validity

	CVID	CDES	ITS	PCHA	PEIP	ESAT
COVID-19 awareness	.813					
Course design qualiy	.270	.857				
Instructors' teaching style	.230	.570	.823			
Perceived challenges during COVID-19	.083	-.189	-.211	.798		
Educational institutions preparedness	.229	.559	.535	-.129	.817	
Satisfaction with e-learning	.218	.549	.462	-.404	.426	.897

4.3   Structural Results

It should be mentioned that when we report on the structural results, values for the South Korean sample are illustrated in the bracket.

The SEM results showed that the satisfaction with e-learning systems was explained by variance of 28% for the Finnish sample [45% South Korean sample], followed by instructors' teaching style 26% for the Finnish sample [56% South Korean sample] and course design quality 39% for the Finnish sample [34% South Korean sample], see Figure 2. According the SEM results, we found that both (i) the instructors' teaching style ($\beta = .19$; $t = 2.381$; $p = .005$ [$\beta = .24$; $t = 3.101$; $p = .005$]) and (ii) course design quality ($\beta = .40$; $t = 4.693$; $p = .001$ [$\beta = .52$; $t = 6.404$; $p = .001$]) significantly impact the satisfaction with e-learning; thus H1 and H2 were supported by the model and we did not find any differences between two samples.

Moreover, the results showed that while COVID-19 awareness had no significant effect on the instructors' teaching style for both samples; thus, rejecting H3a by the model, it had a significant effect on the course design quality ($\beta = .17$; $t = 1.997$; $p = .005$ [$\beta = .19$; $t = 3.110$; $p = .001$]) for both sample; thus, supporting the H3b. Unlike our expectations, perceived challenges during COVID19 negatively impact both instructors' teaching style ($\beta = -.17$; $t = 2.224$; $p = .005$) and course design quality ($\beta = -.19$; $t = 2.537$; $p = .005$) for the Finnish sample only; therefore, H4a and H4b were both rejected by the model. Finally, perceived educational institutions preparedness directly and positively influenced both the instructors' teaching style

(β = .40; t = 3.884; p = .001 [β = .69; t = 13.191; p = .001]) and course design quality (β = .51; t = 6.647; p = .001 [β = .50; t = 5.847; p = .001]); thus, supporting the H5a and H5b.
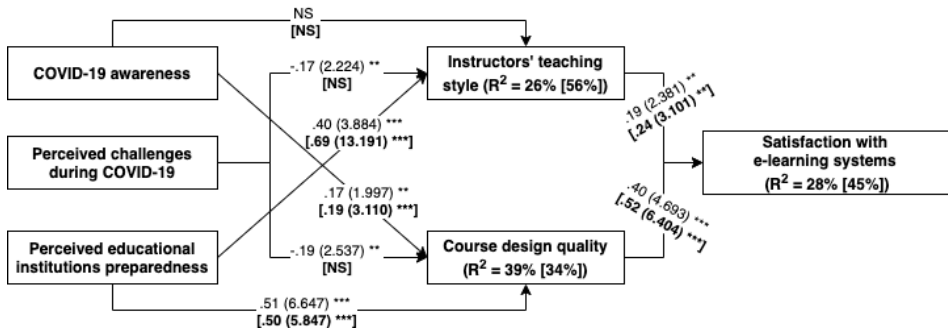
Figure 2: Structural model results (Korean results are presented in the brackets and bold)

We also examined the model to see if the instructors' teaching style and course design quality mediate the paths between COVID-19 related factors to satisfaction with e-learning systems. The mediation test results revealed interesting results. For example, we found that the path between the perceived educational institutions preparedness to satisfaction with e-learning systems was partially mediated by both instructors' teaching style and course design quality for both samples. Also, we found that the path between perceived challenges during the COVID-19 to satisfaction with e-learning was partially mediated by the course design quality for the Finnish sample. For the South Korean sample, the path between COVID-19 awareness to satisfaction with e-learning systems was partially mediated by the course design quality.

5   Discussions

According the SEM results, we could conclude that both intermediate factors (instructors' teaching style and course design quality) that affect how the e-learning process is conducted impact both the Finnish and the South Korean students in a similar manner. This finding indicates, although, we found different results when we examined the impact COVID-19 related factor, for both groups these factors are considered to be important elements of the satisfaction with e-learning systems. Moreover, we also found that, for both groups of students, the effect of COVID-

19 awareness to instructor's teaching style is not significant. In fact, the effect of this factor was only significant to the course design quality. Nonetheless, through the mediation test, we found that the effect of COVID-19 awareness to satisfaction with e-learning is indirect and mediated through course design quality. The SEM results show that the effect of perceived educational institutions preparedness to both instructors' teaching style and course design quality is significant and positive for both groups of students. This is rather important finding, because it shows the importance of the educational institutions and their readiness to provide distance education during the COVID-19. Finally, the effect of perceived challenges during COVID-19 to instructors' teaching style and course design quality was significant only for the Finnish sample. This is rather surprising, as it indicates that for the South Koran sample all COVID-19 related factors such as perceived challenges during the COVID-19 were not important. However, intermediate factors, e.g. educational institutions preparedness deems to be important and significant, as it had the strongest effect on the instructors' teaching style and course design quality effect.

6   Conclusion and Limitations

This paper investigates factors that impact students' satisfaction with e-learning systems during the lockdown of the COVID-19 pandemic. We theoretically contribute to literature in threefold manners. First, we develop and empirically investigate an integrated theoretical model, where not only conventional factors (e.g. instructors teaching style) are incorporated into the model, but also a more contextual related factors in relation to the COVID-19 situation are conceptualised in the model. The results show that contextual factors may directly or indirectly impact students' satisfaction with e-learning systems. Second, by conducting an international comparative research, we contribute to literature by showing different perceptions towards COVID-19 related factors between the Finnish and the South Korean students and how these factors impact their satisfaction with e-learning systems. For instance, for the Finnish students perceived challenges during the COVID-19 is considered to be important but the effect is negative on both instructors' teaching style and course design quality. Also, for the South Korean sample, the school readiness to facilitate the distance learning is considered to be positively associated with both instructors' teaching style and course design quality. Thirdly, regardless of the importance of technology in education, the educational

institutions preparedness to implement and execute e-learning played a central role in boosting students' satisfaction with e-learning systems.

From a more practical standpoint, the results provide useful and new insights for decision-makers at the educational institutions on how advanced learning tools (e.g. LMS) can be used to conduct distance learning and e-learning, while taking contextual factors into account. This is important, because online courses have been available to some extent in both Finnish and South Korean universities. However, due to COVID-19 circumstances it is inevitable that most lectures should be given online. Therefore, it is recommended that educational institutions should have additional efforts and measures to enhance students' satisfaction with e-learning. There are some limitations in this paper too. For example, all participants self-reported that they were students at the time the data was collected, and we were unable to verify this issue as the survey was conducted online. Also, we only collected data from students; however, we believe that teachers' perceptions must also be further studied. Finally, we cannot claim that the results can be generalised and might be applicable only to context of this research.

References

Aavakare, M., & Nikou, S. (2020). Challenging the Concept of Digital Nativeness–Through the Assessment of Information Literacy and Digital Literacy. In International Conference on Well-Being in the Information Society (pp. 211-225). Springer, Cham.

Aboagye, E., Yawson, J. A., & Appiah, K. N. (2021). COVID-19 and E-learning: The challenges of students in tertiary institutions. Social Education Research, 1-8.

Al-Busaidi, K. A., & Al-Shihi, H. (2010). Instructors' acceptance of learning management systems: A theoretical framework. Communications of the IBIMA, 2010, 1-10.

Al-Rahmi, W. M., Othman, M. S., & Yusuf, L. M. (2015). Exploring the factors that affect student satisfaction through using e-learning in Malaysian higher education institutions. Mediterranean Journal of Social Sciences, 6(4), 299.

Alea, L. A., Fabrea, M. F., Roldan, R. D. A., & Farooqi, A. Z. (2020). Teachers' Covid-19 Awareness, Distance Learning Education Experiences and Perceptions towards Institutional Readiness and Challenges. International Journal of Learning, Teaching and Educational Research, 19 (6), 127-144.

Almaiah, M. A., Al-Khasawneh, A., & Althunibat, A. (2020). Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic. Education and Information Technologies, 25, 5261-5280.

Arafat, S., Aljohani, N., Abbasi, R., Hussain, A., & Lytras, M. (2019). Connections between e-learning, web science, cognitive computation and social sensing, and their relevance to learning analytics: A preliminary study. Computers in Human Behaviour, 92, 478-486.

Arbaugh, J.B. (2000). Virtual classroom characteristics and student satisfaction with internet-based MBA courses. Journal of management education, 24(1), 32-54.

Benigno, V., & Trentin, G. (2000). The evaluation of online courses. Journal of computer assisted learning, 16(3), 259-270.

Favale, T., Soro, F., Trevisan, M., Drago, I., & Mellia. M. (2020). Campus traffic and e-Learning during COVID-19 pandemic. Computer Networks, 176, 107290.

Fornell, C. G., & D. F. Larcker. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50.

Garavan, T. N., Carbery, R., O'Malley, G., & O'Donnell, D. (2010). Understanding participation in e-learning in organisations: a large-scale empirical study of employees. International Journal of Training and Development, 14(3), 155-168.

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

Jang, M., Aavakare, M., Nikou, S., & Kim, S. (2021). The impact of literacy on intention to use digital technology for learning: A comparative study of Korea and Finland. Telecommunications Policy, 102154. https://doi.org/10.1016/j.telpol.2021.102154.

Liu, T. Y., & Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. Computers & Education, 55(2), 630-643.

Lu, H. P., & Chiou, M. J. (2010). The impact of individual differences on e-learning system satisfaction: A contingency approach. British Journal of Educational Technology, 41(2), 307-323.

Martín-Rodríguez, Ó., Fernández-Molina, J. C., Montero-Alonso, M. Á., & González-Gómez, F. (2015). The main components of satisfaction with e-learning. Technology, Pedagogy and Education, 24(2), 267-277.

Nikou, S., & Aavakare, M. (2021). An assessment of the interplay between literacy and digital technology in higher education. Education and Information Technologies, 1-23. https://doi.org/10.1007/s10639-021-10451-0.

Nikou, S., & Maslov, I. (2021). An analysis of students' perspectives on e-learning participation–the case of COVID-19 pandemic. The International Journal of Information and Learning Technology. https://doi.org/10.1108/IJILT-12-2020-0220.

Peng, P. J., & Samah, A. (2006). Measuring students' satisfaction for quality education in e-learning university. Unitar E Journal, 2 (1), 11–21.

Radha, R., Mahalakshmi, K., Kumar, V. S., & Saravanakumar, A. R. (2020). E-learning during lockdown of Covid-19 pandemic: a global perspective. International Journal of Control and Automation, 13(4), 1088-1099.

Ramayah, T. and Lee, J.W.C. (2012). System characteristics, satisfaction and e-learning usage: a structural equation model (SEM). Turkish Online Journal of Educational Technology-TOJET, 11(2), pp.196-206.

Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. Computers & Education, 50(4), 1183-1202.

Uppal, M. A., Ali, S., & Gullive, S. R. (2018). Factors determining e-learning service quality. British Journal of Educational Technology, 49(3), 412-426.

Volery, T., & Lord, D. (2000). Critical success factors in online education. International Journal of Educational Management, 14(5), 216-223.

Wright, C. R. (2003). Criteria for evaluating the quality of online courses. Alberta Distance Education and Training Association, 16(2), 185-200.

Linking IT Assets and Competitive Advantage - IT Capabilities of Servitized Business Models

Christoph Brosig,[1] Markus Westner[1] & Susanne Strahringer[2]

[1] OTH Regensburg, Faculty of Computer Science and Mathematics, Regensburg, Germany; email: christoph.brosig@oth-regensburg.de, markus.westner@oth-regensburg.de

[2] TU Dresden, Chair of Business Informatics esp. Information Systems in Trade and Industry, Dresden, Germany; email: susanne.strahringer@tu-dresden.de

Abstract This paper connects research from business model innovation and information systems by exploring critical IT capabilities for servitized business models. The adoption of servitized business models is a major business model innovation strategy. At the same time, digitalization drives the evolution of IT capabilities at these business models. Scholars argue that it remains unclear how IT capabilities enable servitized business models to build a competitive advantage by achieving cost advantages or differentiation. This paper explores IT capabilities that enable building a competitive advantage for servitized business models based on a qualitative analysis of multiple published case studies. The authors identify configurations of IT capabilities among servitized business models. The findings contribute to servitization research by exploring IT capabilities and how they are combined among servitized business models. The insights help practitioners deploy digital technologies and IT assets effectively as building blocks of IT capabilities to advance their servitized business model.

1   Introduction

Since the 1980s, firms move from selling products to offering products as a service (Ulaga and Reinartz, 2011; Vandermerwe and Rada, 1988). Firms pursue this servitization of their business models to improve their competitive advantage (Kindström, 2010; Paschou et al., 2020).

Nowadays, digital technologies and information technology (IT) assets offer new levers to build a competitive advantage for servitized business models (Kohtamäki et al., 2019; Rapaccini and Gaiardelli, 2015). Multiple scholars have explored specific digital technologies and IT assets for digital servitization (Paschou et al., 2020). The resource-based view suggests that firms need to create IT capabilities to build a competitive advantage based on IT assets (Ross et al., 1996).

Despite the increasing number of publications on digital servitization, scholars claim that there is a limited understanding of which IT capabilities enable servitized business models to build a competitive advantage (Coreynen et al., 2017; Grubic and Jennions, 2018). Scholars ask for contributions on how IT capabilities enable different types of competitive advantage (Kohtamäki et al., 2019; Paschou et al., 2020).

We contribute to this discussion by a qualitative analysis of 17 published cases of servitized business models answering two research questions (RQ):

RQ1: Which IT capabilities enable servitized business models to build a competitive advantage?
RQ2: How do IT capabilities enable servitized business models to build a competitive advantage?

Our paper is structured along three main parts to address these two questions. First, we introduce digital servitization and the concept of IT capabilities (section 2). Section 3 describes our case selection and case analysis. Section 4 presents our findings on IT capabilities (RQ1) and configurations of how IT capabilities enable competitive advantages at servitized business models (RQ2). Finally, we discuss our findings and conclude our research.

This paper contributes to business model innovation and information systems (IS) research based on a qualitative analysis of multiple cases. Our paper contributes to the sparse research on IT capabilities of servitized business models and shows their role in building a competitive advantage.

2   Research Background

Servitization describes the transition of a business model from being product-centric to being service-centric (Vandermerwe and Rada, 1988). While product-centric business models focus on the sales of products, service-centric business models employ products to deliver outcomes as a service (Reim et al., 2015). Scholars suggest mapping business models along a continuum of product- to service-centric (Reim et al., 2015).

Digital technologies offer new levers to build a competitive advantage for a business model undergoing servitization (Kohtamäki et al., 2019). Scholars have introduced digital servitization to label the service transition of a business model enabled by digital technologies (Rapaccini and Gaiardelli, 2015).

For such a transition, it is critical to understand how digital technologies and IT assets enable a competitive advantage. The resource-based view offers an explanation based on the notion of assets and capabilities. Firms invest in assets and create capabilities to employ these assets to build a competitive advantage. The concept of capabilities links assets and competitive advantage (Grant, 1991).

An IT capability describes the ability to employ IT assets to support and enhance a firm's strategy or work processes to build a competitive advantage (Lu and Ramamurthy, 2011; Ross et al., 1996). This competitive advantage can be a cost advantage or differentiation (Porter, 1985). Scholars distinguish IT capabilities employing various IT assets. There are three types of IT assets: Tangible IT assets include, e.g., hardware, software, or data assets. Intangible IT assets refer to, e.g., IT management practices. Human IT assets are, e.g., specific IT skills (Ross et al., 1996).

Over the last years, scholars have introduced digital capabilities as types of IT capabilities (Côrte-Real et al., 2020; Krishnamoorthi and Mathew, 2018). Digital capabilities employ stacks of IT assets as digital technologies to support and enhance

a firm's strategy or work processes to build a competitive advantage (Brosig et al., 2020). In this study, we refer to the overarching concept of IT capabilities to cover the range of IT assets.

3   Research Methodology

In this section, we describe the data selection and data analysis of our case-based approach. We decided to analyze published case studies about servitized business models due to the early stage of this research stream (Yin, 2014).

First, we set up a case base. We searched seven literature databases and selected case studies in a two-step approach. Figure 1 summarizes the search parameters and the screening stages of contributions for our case base.

Search parameters	• **Search terms:** *(Servitization* OR *"Integrated Solution"* OR *"Service transition") AND "Case Study"* • **Search area:** Title, Abstract, Keywords • **Publication date:** (01st January) 2015- (20th March) 2020 • **Databases:** AISel, EBSCO Academic Search Complete, EBSCO Business Source Complete, ProQuest, ScienceDirect, Web of Science, Xplore		
Screening stages Amount of results	Initial search results 460	Potential results 135	Results 15
	not discounted for duplicates	discounted for duplicates	

Figure 1: Overview of case study search and screening

We adopted our search terms from three extensive servitization literature reviews and chose the consistently used terms (Baines et al., 2017; Kowalkowski et al., 2017; Rabetino et al., 2018) searching in title, abstract, and keywords. We restricted our search to contributions from 2015 until 2020 (time of data collection), as most servitization literature associated with digitally-enabled service transition was published since then (Paschou et al., 2020). Before we screened the data, we chose three screening criteria, whether the contribution (1) is based on a case study (single- or multiple-case studies), (2) indicates competitive advantage of the case firm, and (3) provides information about the employment of IT assets in the case context linked to the competitive advantage. We obtained 17 cases from 15 contributions. Table 1 shows our case base, including reference, name of the case firm as stated in the original reference, industry, and the respective customer group.

C. Brosig, M. Westner &S. Strahringer:
Linking IT Assets and Competitive Advantage - IT Capabilities of Servitized Business Models

115

Table 1: Overview of the case base

Case ID	Reference	Case name	Industry	Customer Group
[1]	Lim et al. (2015)	*undisclosed*	Car manufacturer	B2B/B2C
[2]	Beltagui (2018)	Eng. Co.	Power systems provider	B2B
[3]	Chen and Møller (2019)	*undisclosed*	Farming equipment provider	B2B
[4]	Niño et al. (2015)	*undisclosed*	Chemical equipment provider	B2B
[5]	Saarikko (2015)	DigitalCo	Telecommunication	B2B
[6]	Bressanelli et al. (2018)	Alpha	Household appliances provider	B2C
[7]	Robinson et al. (2016)	Laing O'Rourke	Construction provider	B2B
[8]	Sklyar et al. (2019)	Navicula	Maritime equipment provider	B2B
[9]	Reim et al. (2016)	Alpha	Construction machinery provider	B2B
[10]	Rapaccini et al. (2019)	Alfa	Building Equipment Provider	B2B
[11]	Dalenogare et al. (2019)	*undisclosed*	Building Equipment Provider	B2B
[12]	Weeks and Benade (2015)	*undisclosed*	Building Equipment Provider	B2B
[13]	Clegg et al. (2017)	Coen	Construction	B2B
[14]	Coreynen et al. (2017)	Beta	Electronic Switchboards Provider	B2B
[15]	Rymaszewska et al. (2017)	Company A	Manufacturing Machinery Provider	B2B
[16]	Rymaszewska et al. (2017)	Company B	Power Generators Provider	B2B
[17]	Rymaszewska et al. (2017)	Company C	Power Transformers Provider	B2B

We followed the resource-based view for our analysis: first, we coded IT assets among the cases with an open coding approach (Corbin and Strauss, 2015). Second, we analyzed how case firms employ these IT assets to build a competitive advantage, individually or as stacks. As a result, we linked IT assets and competitive advantage by IT capabilities (RQ1). Next, we analyzed configurations of IT capabilities to understand how IT capabilities enable servitized business models to build a competitive advantage (RQ2).

4   Results

In this section, we present three IT capabilities of servitized business models and configurations of how these IT capabilities enable a competitive advantage. For each IT capability, we introduce three examples from our cases. An overview of all examples across our cases can be obtained from the authors.

4.1   IT Capability to Connect the Value Chain

The first IT capability employs IT assets to connect the value chain to achieve cost advantages. Table 2 shows three examples based on our cases.

Table 2: Selected cases with IT capability to connect the value chain

Case ID	Observed IT Assets	Employment of IT Assets to Generate Competitive Advantage	Competitive Advantage
2	• Engine usage data *(tangible IT asset)* • Virtual engine testing models *(tangible IT asset)*	• Connect value chain (from maintenance delivery to product development) to reduce the efforts to resolve technical malfunctions by virtual engine simulation with engine usage data	Cost advantage
10	• Cloud-based management accounting system accessible to service network partners *(tangible IT asset)*	• Connect value chain (accounting) among service partners to uncover costs across the service network and eliminate them	
13	• Inventory management system externally accessible to suppliers *(tangible IT asset)*	• Connect value chain (inventory management) with suppliers to ensure availability of materials to avoid project delays at additional costs	
Cases showing same IT capability [Case IDs]: [2, 6, 7, 9, 10, 11, 12, 13]			

Case firms use primarily tangible IT assets, e.g., software systems, to distribute information internally along their value chain, e.g., from maintenance operations to product development or from maintenance operations to spare parts handling. Some case firms offer integration points to external stakeholders, like suppliers or service partners, to connect to their value chains. This connection enables efficient orchestration of processes, e.g., product development or maintenance delivery, and (human) resources, e.g., available maintenance technicians or spare parts. Case firms achieve cost advantages as a competitive advantage from this IT capability.

C. Brosig, M. Westner &S. Strahringer:
Linking IT Assets and Competitive Advantage - IT Capabilities of Servitized Business Models

117

4.2   IT Capability to Connect Products

The second IT capability employs IT assets to connect products to differentiate by value-adding services. Case firms introduce tangible IT assets, e.g., product-integrated sensors, data transmission devices, or software systems to access connected products remotely. Table 3 shows three examples from our cases.

Table 3: Selected cases with IT capability to connect products

Case ID	Observed IT Assets	Employment of IT Assets to Generate Competitive Advantage	Competitive Advantage
3	• Software farm management system to connect to farm components *(tangible IT asset)* • Software developers, user interface experts, and user experience experts to build software system *(human IT asset)*	• Connect products (farm components to farm management software) to enable digital farm monitoring as an additional service	Differentiation by value-adding services
8	• Customer portal to manage condition data of maritime vessels and to access 3rd party maritime software *(tangible IT asset)* • Vessel condition data *(tangible IT asset)* • 3rd party maritime software *(tangible IT asset)*	• Connect products to enable monitoring of condition data for onshore operations of maritime vessels and to offer 3rd party software access as services	
17	• Logging device for power transformer data with internet connection *(tangible IT asset)* • Usage and operational fault data of power transformers *(tangible IT asset)*	• Connect products to enable access to power transformer operations metrics to offer use-based advisory to prolong life-cycle	
Cases showing same IT capability [Case IDs]: [1, 2, 3, 4, 5, 6, 7, 8, 9, 14, 15, 16, 17]			

Some cases explicitly mention the employment of human IT assets for this IT capability, e.g., IT skills to integrate sensors into products or skills to develop and deploy code for respective software systems. These IT assets make products connected to offer value-adding services, e.g., remote monitoring, remote maintenance, or use-based advisory. Case firms offer these value-adding services to differentiate as a competitive advantage.

4.3 IT Capability to Interconnect Value Chain and Products

The third IT capability employs IT assets to interconnect a value chain and products. This interconnection enables differentiation by performance-based contracts. This IT capability is the ability to employ integration points - between value chain and products - as IT assets to build a competitive advantage. Table 4 offers an outline of three examples from our cases.

Table 4: Selected cases with IT capability to interconnect value chain and products

Case ID	Observed IT Assets	Employment of IT Assets to Generate Competitive Advantage	Competitive Advantage
6	• IoT device in washing machines to extract and send data to the provider *(tangible IT asset)* • Washing machine usage data *(tangible IT asset)* • Data analytics tools to detect careless usage of product *(tangible IT asset)*	• Interconnect connected products (usage data) with the value chain (contract monitoring) to operate performance-based contracts	Differentiation by performance-based contracts
9	• IoT device in building equipment to extract and send data to the provider *(tangible IT asset)* • Building equipment condition data *(tangible IT asset)* • Software systems on availability of maintenance services and spare parts *(tangible IT asset)*	• Integration of connected products (condition data indicating maintenance needs) with the value chain (service systems and inventory data) to schedule maintenance delivery for performance-based contracts	
15	• IoT device in machine to extract sensor data and send to the provider *(tangible IT asset)* • Machine usage and performance data *(tangible IT asset)* • Cloud-based platform to access machine data for service organization *(tangible IT asset)*	• Interconnect connected products (usage/performance data) with the value chain (service organization) for remote support in performance-based contracts	
Cases showing same IT capability *[Case IDs]*: *[2, 6, 7, 9, 10, 11, 12, 13]*			

Case firms link product data to value chain information. Several firms connect their products to monitor product condition and usage data, as covered in section 4.2. In distinction, this IT capability focuses on the link of such product data to the providers' value chains, e.g., to anticipate product failure. Firms displaying this IT capability distribute the product data as information along the value chain, e.g., to activate maintenance provision. Interconnecting product data with value chain

C. Brosig, M. Westner &S. Strahringer:
Linking IT Assets and Competitive Advantage - IT Capabilities of Servitized Business Models

119

information is critical to achieving agreed performance levels of products as a service.

4.4   Configurations of IT Capabilities Among Servitized Business Models

We analyzed how the three IT capabilities are distributed in our case firms and found case evidence for four out of eight possible combinations. We refer to each combination as a configuration where each of the IT capabilities is present or absent. Figure 2 shows an overview of these configurations and associated cases.

Configurations of IT Capabilities	Presence of IT Capabilities… (marked as ⊘)			Observation of Configurations	
	to Connect the Value Chain	to Connect Products	to Interconnect Value Chain and Products	Number of Cases	Observed in Cases *[ID]*
A				0	
B	⊘			4	*[10, 11, 12, 13]*
C		⊘		7	*[1, 3, 4, 5, 8, 14, 17]*
D	⊘	⊘		0	
E			⊘	0	
F	⊘		⊘	0	
G		⊘	⊘	2	*[15, 16]*
H	⊘	⊘	⊘	4	*[2, 6, 7, 9]*

Figure 2: Overview of configurations of IT capabilities among servitized business models

We found case evidence for configurations B, C, G, and H, but not for A, D, E, and F. The case evidence supports how configurations of IT capabilities enable servitized business models to build competitive advantage.

The IT capability to connect products could be a sufficient IT capability for firms to build a competitive advantage, as shown in the configurations C, G, and H. Still, due to the lack of evidence for configuration D, this cannot be confirmed.

The IT capability to connect the value chain is present both individually (configuration B) or in combination with other IT capabilities (configuration H). In contrast, the IT capability to interconnect value chain and products is only present in combination with other IT capabilities (configurations G, H), in particular with the IT capability to connect products.

5   Discussion

In this section, we discuss our findings in comparison to existing literature. Our study makes two contributions: we identify IT capabilities for servitized business models (RQ1), and we find configurations of IT capabilities that enable servitized business models to build a competitive advantage (RQ2).

We find three IT capabilities among servitized business models that employ IT assets, (1) the IT capability to connect the value chain to achieve cost advantages, (2) the IT capability to connect products to achieve differentiation, (3) the IT capability to interconnect the value chain and products to achieve differentiation. Our study confirms the importance of IT capabilities in linking IT assets with competitive advantage: case firms employ different IT assets to build a similar competitive advantage. Some case firms employ similar stacks of IT assets to build different competitive advantages. IT capabilities help to understand these equifinal ways how IT assets contribute to building a competitive advantage.

We show configurations of IT capabilities among our case firms. Configuration B includes case firms focusing on the IT capability to connect the value chain to streamline processes and resources. This configuration is similar to the nature of capabilities of product-oriented business models striving for efficient processes (Kohtamäki et al., 2019; Reim et al., 2015; Ulaga and Reinartz, 2011). Configuration C is based on the presence of the IT capability to connect products. Case firms with configuration C offer their services to support the use of the product: services integrate with the product in use. Case firms with configuration C are similar to use-based solution providers with services as an integral part of their offering to maximize product efficiency for the customer (Kohtamäki et al., 2019; Ulaga and Reinartz, 2011). Configurations G and H include configurations of IT capabilities with the IT capability to interconnect value chain and products. Case firms with these configurations differentiate at least by offering product performance as a service. Literature labels similar business models as result-oriented or outcome providers (Kohtamäki et al., 2019; Reim et al., 2015).

C. Brosig, M. Westner &S. Strahringer:
Linking IT Assets and Competitive Advantage - IT Capabilities of Servitized Business Models

121

We do not observe configurations A, D, E, and F in our cases. Configuration A is not to be found due to our initial case selection. It would not have contributed to clarifying the IT capabilities of servitized business models. The lack of case evidence for configurations E and F could indicate that the IT capability to interconnect value chain and products is dependent on the IT capability to connect products. Thus, it could be that configurations E and F are theoretically not possible. In contrast to configurations A, E, and F, configuration D could be available among cases beyond our case base.

Based on our insights on configurations of IT capabilities, we derive the assumption that specific configurations of IT capabilities support specific types of servitized business models along the continuum from product- to service-centric.

6   Conclusion

Our paper helps answer the call for interdisciplinary research at the frontier of business model innovation and IS research (Kohtamäki et al., 2019; Paschou et al., 2020). Section 4 provides an overview of three IT capabilities at servitized business models, (1) the IT capability to connect the value chain, (2) the IT capability to connect products, and (3) the IT capability to interconnect value chain and products (RQ1). We find five configurations of how IT capabilities enable building competitive advantage from rather product- to service-centric servitized business models (RQ2).

Practitioners profit from our synthesis of business model innovation and IS research by obtaining transparency about IT capabilities for servitized business models (Baines et al., 2017). Our configurations of IT capabilities offer starting points to invest in assets that may be used to build a competitive advantage. From our case evidence, practitioners also learn that individual IT assets per se do not build a competitive advantage for servitized business models (cf. (Wiener et al., 2020)).

Our study is not free from limitations: first, to ensure external validity, we collected cases from multiple research fields. Most cases cover the B2B area, consistent with previous servitization research (Paschou et al., 2020). Therefore, our findings may not be generalizable to the B2C area. Second, we sampled cases from fields where IT capabilities are not the primary research contribution. In some cases, the selected

cases may not exhaustively cover all IT capabilities of a servitized business model. We decided to mitigate this risk by sampling a broad set of cases to cover contributions from multiple perspectives.

For future research, we propose further analyses of IT capabilities for servitized business models. Researchers should continue to analyze how IT capabilities differ among different types of servitized business models along the continuum from product- to service-centric. Researchers could use our hypothesis as a starting assumption. For their analyses, they could apply a configurational approach as IT capabilities of servitized business models appear to create equifinal links between IT assets and competitive advantage.

References

Baines, T., Bigdeli, A. Z., Bustinza, O. F., Shi, V. G., Baldwin, J. S., & Ridgway, K. (2017). Servitization: Revisiting the State-of-the-Art and Research Priorities. International Journal of Operations & Production Management, 37 (2), 256-278.

Beltagui, A. (2018). A Design-Thinking Perspective on Capability Development. International Journal of Operations & Production Management, 38 (4), 1041-1060.

Bressanelli, G., Adrodegari, F., Perona, M., & Saccani, N. (2018). Exploring How Usage-Focused Business Models Enable Circular Economy Through Digital Technologies. Sustainability, 10 (3), 639.

Brosig, C., Westner, M., & Strahringer, S. (2020). Revisiting the Concept of IT Capabilities in the Era of Digitalization. Proceedings of the 22nd International Conference on Business Informatics, Antwerp.

Chen, K.-L., & Møller, C. (2019). A Thousand-Mile Journey Begins With the Very First Step: The Case of a Product-Centric Manufacturing Firm's Transformation Towards Servitization. CIRP Journal of Manufacturing Science and Technology, 27, 102-113.

Clegg, B., Little, P., Govette, S., & Logue, J. (2017). Transformation of a Small-to-Medium-Sized Enterprise to a Multi-Organisation Product-Service Solution Provider. International Journal of Production Economics, 192, 81-91.

Corbin, J., & Strauss, A. (2015). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (Vol. 4). Thousand Oaks: Sage Publications.

Coreynen, W., Matthyssens, P., & Van Bockhaven, W. (2017). Boosting Servitization Through Digitization: Pathways and Dynamic Resource Configurations for Manufacturers. Industrial Marketing Management, 60, 42-53.

Côrte-Real, N., Ruivo, P., & Oliveira, T. (2020). Leveraging Internet of Things and Big Data Analytics Initiatives in European and American Firms: Is Data Quality a Way to Extract Business Value? Information & Management, 57 (1), 103141.

Dalenogare, L. S., Baseggio, M. M., Ayala, N. F., Le Dain, M.-A., & Frank, A. G. (2019). The Contribution of Smart Glasses for PSS. Procedia CIRP, 83, 318-323.

Grant, R. M. (1991). The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. California Management Review, 33 (3), 114-135.

Grubic, T., & Jennions, I. (2018). Remote Monitoring Technology and Servitised Strategies – Factors Characterising the Organisational Application. International Journal of Production Research, 56 (6), 2133-2149.

C. Brosig, M. Westner &S. Strahringer:
Linking IT Assets and Competitive Advantage - IT Capabilities of Servitized Business Models

123

Kindström, D. (2010). Towards a Service-Based Business Model – Key Aspects for Future Competitive Advantage. European Management Journal, 28 (6), 479-490.

Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., & Baines, T. (2019). Digital Servitization Business Models in Ecosystems: A Theory of the Firm. Journal of Business Research, 104, 380-392.

Kowalkowski, C., Gebauer, H., & Oliva, R. (2017). Service Growth in Product Firms: Past, Present, and Future. Industrial Marketing Management, 60, 82-88.

Krishnamoorthi, S., & Mathew, S. K. (2018). Business Analytics and Business Value: A Comparative Case Study. Information & Management, 55 (5), 643-666.

Lim, C.-H., Kim, M.-J., Heo, J.-Y., & Kim, K.-J. (2015). A Conceptual Framework for Designing Informatics-based Services in Manufacturing Industries. Procedia CIRP, 30, 72-77.

Lu, Y., & Ramamurthy, K. (2011). Understanding the Link Between Information Technology Capability and Organizational Agility: An Empirical Examination. MIS Quarterly, 35 (4), 931-954.

Niño, M., Blanco, J. M., & Illarramendi, A. (2015). Business Understanding, Challenges and Issues of Big Data Analytics for the Servitization of a Capital Equipment Manufacturer. Proceedings of the International Conference on Big Data, Santa Clara.

Paschou, T., Rapaccini, M., Adrodegari, F., & Saccani, N. (2020). Digital Servitization in Manufacturing: A Systematic Literature Review and Research Agenda. Industrial Marketing Management, 89, 278-292.

Porter, M. E. (1985). The Competitive Advantage: Creating and Sustaining Superior Performance. New York: The Free Press.

Rabetino, R., Harmsen, W., Kohtamäki, M., & Sihvonen, J. (2018). Structuring Servitization-Related Research. International Journal of Operations & Production Management, 38 (2), 350-371.

Rapaccini, M., & Gaiardelli, P. (2015). Smart Services Initiatives in Product-Centric Companies. Proceedings of the 4th Spring Servitization Conference, Aston.

Rapaccini, M., Mauro, S. G., Cinquini, L., & Tenucci, A. (2019). Servitization of SMEs through Strategic Alliances: a Case Study. Procedia CIRP, 83, 176-181.

Reim, W., Parida, V., & Örtqvist, D. (2015). Product–Service Systems (PSS) Business Models and Tactics – A Systematic Literature Review. Journal of Cleaner Production, 97, 61-75.

Reim, W., Parida, V., & Sjödin, D. R. (2016). Risk Management for Product-Service System Operation. International Journal of Operations & Production Management, 36 (6), 665-686.

Robinson, W., Chan, P., & Lau, T. (2016). Finding New Ways of Creating Value A Case Study of Servitization in Construction. Research-Technology Management, 59 (3), 37-49.

Ross, J. W., Beath, C. M., & Goodhue, D. L. (1996). Develop Long-Term Competitiveness through IT Assets. Sloan Management Review, 38 (1), 31-42.

Rymaszewska, A., Helo, P., & Gunasekaran, A. (2017). IoT Powered Servitization of Manufacturing – An Exploratory Case Study. International Journal of Production Economics, 192, 92-105.

Saarikko, T. (2015). Digital Platform Development: A Service-Oriented Perspective. Proceedings of the 23rd European Conference on Information Systems, Münster.

Sklyar, A., Kowalkowski, C., Sörhammar, D., & Tronvoll, B. (2019). Resource Integration Through Digitalisation: A Service Ecosystem Perspective. Journal of Marketing Management, 35 (11-12), 974-991.

Ulaga, W., & Reinartz, W. J. (2011). Hybrid Offerings: How Manufacturing Firms Combine Goods and Services Successfully. Journal of Marketing, 75 (6), 5-23.

Vandermerwe, S., & Rada, J. (1988). Servitization of Business: Adding Value by Adding Services. European Management Journal, 6 (4), 314-324.

Weeks, R., & Benade, S. (2015). The Development of a Generic Servitization Systems Framework. Technology in Society, 43, 97-104.

Wiener, M., Saunders, C., & Marabelli, M. (2020). Big-Data Business Models: A Critical Literature Review and Multiperspective Research Framework. Journal of Information Technology, 35 (1), 66-91.

Yin, R. K. (2014). Case Study Research: Design and Methods (5th ed. Vol. 5). Thousand Oaks: Sage Publications.

Data Management Platforms: An Empirical Taxonomy

Joschka A. Hüllmann, Ajith Sivakumar &
Simone Krebber

University of Münster, Business School, Münster, Germany; e-mail: huellmann@uni-muenster.de, ajith.sivakumar@uni-muenster.de, s_kreb01@uni-muenster.de

Abstract Data management platforms (DMPs) are a widely used means of placing targeted advertising, for example, commercial or political advertisements. However, only a few academic papers shed light on the platforms' mechanisms. These mechanisms' opacity makes it hard for consumers to understand what happens with their data, and regulators struggle to implement effective regulations. Hence, we develop a taxonomy to understand and compare different characteristics of DMPs. Following Nickerson's (2013) method and combining an inductive and deductive approach, eight dimensions emerge that differentiate DMPs. We evaluate the taxonomy's applicability and test it with a set of nine DMPs, which we select by feasibility, relevance, and popularity. The application shows that the eight dimensions cover the significant features that explain most of the variance in characteristics between DMPs. The evaluation revealed opportunities for further development of the taxonomy.

1   Glorious Times for Data Management Platforms

Investments in targeted commercial advertising are at their peak and projected to grow further in the upcoming years (Zenith, 2019), with entire business models being based on the concept (e.g., Facebook, Instagram, YouTube). Likewise, targeted advertisements are a common practice in political election campaigns. A groundbreaking event was the Cambridge Analytica scandal revealed in 2018. Personal information was used to categorise voters into specific segments to positively influence them with targeted political advertisements regarding the election campaign in 2016 (Berghel, 2018; Cadwalladr & Graham-Harrison, 2018). This event escalated the debate about targeted advertising in politics and sparked controversy about the topic. Discussions center around implications such as data privacy and the manipulation of opinions and consumer behaviour. The enabler of targeted advertising, be it in politics or business, are real-time bidding systems. Real-time bidding systems comprise multiple actors, whose interplay permits the dynamic placement of advertisements targeted to specific users. One of these actors is the data management platform (DMP). As the name suggests, DMPs are information systems used to manage consumers' data that is collected through online tracking amongst other means. They gained importance as an enterprise technology in recent years (Poleshova, 2017) as the digitalisation of everyday life co-occurs with collecting personal data at an unprecedented scale through tracking online activities on digital devices (Hüllmann et al., 2021). While data in its raw format adds little value, it becomes a valuable asset for businesses, researchers, and other stakeholders by gathering, managing, processing, and analysing it. DMPs provide such functionalities and support systems for all kinds of data-related operations, which are especially important in the context of real-time bidding systems (Poleshova, 2017).

DMPs are an important actor in the real-time bidding system. Understanding their functions and mechanisms is a prerequisite for discussing and instating effective regulations surrounding privacy and the manipulation of consumers' opinions and behaviors. Nevertheless, a taxonomy that characterises the detailed mechanisms of how DMPs add value has not yet been established in the academic literature[1]. Thus, in this study, we focus on the functions performed by DMPs and put forward the following research question:

---

[1] Our search for DMP literature and especially a taxonomy on DMPs in the three academic search engines Scopus, Web of Science, and Google Scholar did not yield any results.

RQ: How can we distinguish the functionalities of data management platforms? What are the discriminating characteristics?

We address this question by developing a taxonomy that provides an overview and comparison of DMPs. The taxonomy is developed following the Nickerson et al. (2013) approach using an iterative process that combines inductive and deductive phases. It is evaluated by applying it to nine DMPs. Our taxonomy contributes to a better understanding of DMPs, their functionalities, and their mechanisms. Investigating the characteristics of DMPs (e.g., availability of anonymisation functions, real-time processing) contributes knowledge to the debate on targeted advertising and privacy with implications for future regulations.

2 Data Management Platforms in Real-Time Bidding Systems

DMPs are most prominent in the real-time bidding process (see Figure), in which there are two main stakeholders: the advertisers and website operators. To realise real-time targeted advertising, ad-networks take offers of advertising slots from supply-side platforms and match these with bids for advertising slots from demand-side platforms. For the matching process, demand-side platforms use real-time processed data from DMPs (e.g., consumer segments), predefined selection criteria, empirical values, and publishers' predictions about whether the advertising space will be worthwhile (Dawson, 2014; Zhang et al., 2014). If all conditions are met, the deal is closed on the ad-network, and the supply-side platform places the advertiser's advertisement in the advertising space (Wang et al., 2017). A DMP (usually a separate company from the advertising network) collects, stores, manages, and processes the relevant data. Effective coordination between these multiple actors – ad-network, demand- and supply-side platforms, and a DMP – facilitates targeted advertising for defined consumer segments in near real-time. The information from the data or the data itself might be traded on data marketplaces (Lange et al., 2018). While critics contend that arrangements like this can violate consumers' privacy, proponents contend that they serve consumers by presenting them with offers likely to appeal to them.
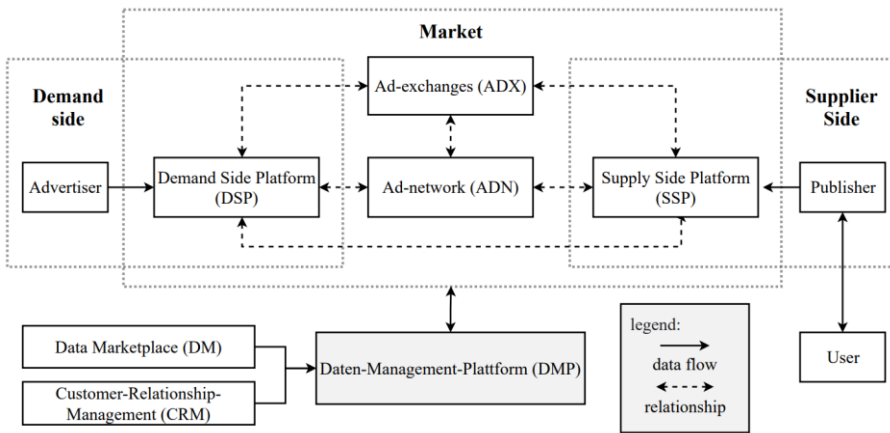
Figure 1: Structure of a Generic Real-Time Bidding System.
Sources: Spiekermann, 2019; Yuan et al., 2012

3   Research Method

To develop a taxonomy of DMPs that allows for their differentiation, we followed the method proposed by Nickerson et al. (2013). The method follows an iterative process through which the taxonomy is gradually built by adding new characteristics. Initially, we curated a longlist of 48 DMPs using internet sources such as articles or reports from auditing companies (e.g., Andrew et al., 2020; Moffett & Chien, 2019). In our iterations to develop the taxonomy, we derived the characteristics deductively from the literature (Kamps & Schetter, 2018; Wang et al., 2017) and inductively from existing DMPs (Adobe, 2020; Google, 2020; Lotame, 2020; Oracle, 2020b). The iterative process terminated after six iterations because all specified subjective and objective termination conditions were met. With regards to objective termination conditions, this means that (1) no dimension or characteristic was merged or split in the last iteration, (2) each dimension was unique and did not repeat, and (3) each characteristic was unique within its dimension. The five subjective conditions used to evaluate the taxonomy's quality after each iteration were conciseness, robustness, extensibility, comprehensiveness, and explainability (Nickerson et al., 2013) (see appendix). In the end, the resulting taxonomy was discussed and jointly evaluated by the three authors. The taxonomy's evaluation was performed using nine DMPs from the longlist (see appendix, Table). We selected these nine DMPs because they met our requirements in terms of 1) **feasibility** (amount of information and trial

version availability), 2) **relevance** (assessing the DMP revenue), and 3) **popularity** (number of clients and awards, which illustrates the influence and reach of the DMPs that can affect people). For each of the nine DMPs, we gathered the necessary information to apply the taxonomy, relying primarily on official information sources, including the DMPs' public websites and trial versions if available. In case of incomplete and missing information on the website, we contacted the support via phone, email, or contact form on their website.

4   Taxonomy

The eight main dimensions are data import, generable data, data sources, web-tracking, data processing functions, external data sources, data export, and data security (see Table). **Data import** distinguishes three different data types, which can be imported in a DMP. **First-party data (FPD)** is obtained by companies through previous consumer contact, website visits (Kreutzer, 2018), or directly from the consumer. The latter includes browsing behaviours or socio-demographic data such as gender or age (Cederholm & Simpson, 2018; Kamps & Schetter, 2018). **Second-party data (SPD)** is first-party data of an external company acquired through direct partnerships (Kamps & Schetter, 2018). **Third-party data (TPD)** is anonymised data provided by data resellers or data marketplaces.

Generable data distinguishes four types of data that DMPs can generate. One is **first-party data** collected by DMPs with the help of tracking mechanisms, for example, collecting it from internal customer management systems. Apart from first-party data, DMPs can identify consumer **segments** based on offline and online data analyses. The granularity, that is, the level of detail in the consumer segment, may vary.

Look-a-like user profiles are generated by DMPs using existing first-party data as a basis. Look-a-like profiles are identified user profiles similar to a company's consumer profiles in terms of indicators such as age, interests, or hobbies. Machine learning is often integrated into the look-a-like modeling process and increases the number of consumers that can be reached with the advertising campaign. DMPs can also determine **unique user profiles** and track them across websites by using internal and external data sources.

The use of cookies is an often **mentioned web-tracking method** of DMPs. While some DMPs track data without using cookies (e.g., fingerprinting), they do not explicitly mention which alternative methods they use. We refer to them as **other methods** in the taxonomy. **Data sources** from which DMPs collect data originate from a company's **website** directly, its **apps**, or **social networks**.

Real-time analysis facilitates the immediate analysis and processing of data, enabling quick decision-making. Data **anonymisation functions, segmentation** of data into consumer segments, and **demand-side platform test functions** are common. DMPs' demand-side test functions are useful when operating with multiple external demand-side platforms to identify the appropriate one for the existing DMP. A wrong choice can harm the advertising campaign to a potential 20% to 40% loss in user profile reach (Joe, 2014). **Machine-learning algorithms** are used to extract meaningful information from unstructured data, such as product popularity. **Waste management** removes records in the DMP not relevant for further processing.

Integrating **external data sources** is crucial for some processes, such as the engendering of look-a-like data or the refinement of existing user profiles. In this regard, DMPs either **provide a platform themselves** with with interfaces to third-party platforms through which data can be acquired. Besides, some DMPs integrate **partner exchange platforms**, which enable companies to establish connections to partners with whom they want to exchange anonymised data. The advantage of this is that the origin of the data is known. Knowing the data source adds transparency and trust on the one hand, and on the other hand, lets people evaluate the quality of the data that originates from the respective source.

Usually, DMPs offer a **data export** option to integrate with interfaces from **customer relationship management software, demand-side platforms, supply-side platforms**, and **ad-networks/ad-exchanges**. Some DMPs provide an export option to **data marketplaces** on which collected data can be sold to external parties. Apart from specific data export options, the export into a **generic local file** (e.g., a consumer data feed) enables companies to use the exported data for various purposes such as business intelligence tools (Chou et al., 2005). Generally, not only raw data but also processed data can be transferred to other platforms (Kamps & Schetter, 2018).

Table 1: Taxonomy on DMPs and Evaluation. A = Adobe, G = Google, L = Lotame, M = MediaMath, Ne = Neustar, Ni = Nielson, O = Oracle, S = Salesforce, T = The Trade Desk. Sources referenced by footnotes were used to establish taxonomy dimensions and characteristics, while the "x" in the evaluation reference DMPs that possess the specific characteristic.

Taxonomy		Evaluation								
Dimension	**Characteristics**	**A**	**G**	**L**	**M**	**Ne**	**Ni**	**O**	**S**	**T**
Data Import [B,C]	First-Party Data [A,B,C,D,E]	x	x	x	x	x	x	x	x	x
	Second-Party Data [A,C,D,E]	x		x		x		x	x	
	Third-Party Data [A,B,C,D,E]	x	x	x	x	x		x	x	
Generable Data [C,D,F,G]	First-Party Data [A,B,C,D,E,G]	x	x	x	x	x	x	x	x	
	Consumer Segments [A,B,C,D,E,G]	x	x	x	x	x	x	x	x	x
	Look-a-like Data [A,B,D,E]	x	x	x	x	x	x	x	x	x
	User Identification [A,B,D,E,G]	x	x	x	x	x	x	x	x	
Data Sources [C,D,F,G]	Apps [A,B,D,E]	x	x	x	x	x	x	x	x	x
	Social Networks [D,E]			x			x	x		
	Websites [A,B,D,E,G]	x	x	x	x	x	x	x	x	x
Web Tracking [C,D,G]	Cookies [A,B,C,D,E]	x	x	x	x	x	x	x	x	x
	Other Methods [A,B,E]	x	x		x	x		x	x	x
Data Processing Functions [A,B,C,F,G]	Real-time Analysis [B,C,D,E,G]	x	x	x	x	x	x	x		
	Data Anonymization [B,D]	x	x	x						
	Data Segmentation [B,C,D,E,G]	x	x	x	x	x	x	x	x	x
	Demand-side Platform Testing [A,B]	x	x							
	Machine Learning Algorithms [A,B,D]	x	x		x	x	x		x	
	Waste Management [D,E]	x		x		x		x		
External Data Sources [D]	Partner Exchange Platform [A,B,D,E]	x	x	x				x		
	Interface for Third Party Data [A,B,D,E]	x	x	x	x	x	x	x		
Data Export [C,F,G]	Ad-network [A,B,C,D,E,G]	x	x	x	x	x	x	x	x	x
	Customer Relationship Management [B,D]		x	x						
	Data Marketplace [A,E]	x				x		x		x
	Demand-side Platform [A,C,D,E,G]	x		x	x	x		x		x
	Export-local-file [A,B]	x	x							
	Supply-side Platform [C,D]			x		x				x
Data Security [F]	Consent Management [D,E,F]			x		x		x	x	x
	Data Security Function [A,B,F]	x	x			x				
	GDPR [A,B,D,E,F]	x	x	x		x		x	x	x

Sources: Adobe[A], 2020; Google[B], 2020; Kamps & Schetter[C], 2018; Lotame[D], 2020; Oracle[E], 2020; Schonschek[F], 2020; Wang et al.[G], 2017

The dimension of **data security** focuses on General Data Protection Regulation (GDPR) compliance, data security, and consent management. The **GDPR** set in place in 2016 by the European Parliament is a set of regulations concerning the processing of personal data in Europe (Art. 1-99 DSGVO). Because the GDPR (Art. 6 Abs. 1 and Art. 7) envisages that companies need users' voluntary consent before collecting their data, **consent management** is often integrated into DMPs.

We evaluated our taxonomy by analysing and categorising nine DMPs, showing that the eight dimensions cover the significant features that explain most of the variance in the taxonomy's characteristics across the DMPs. The evaluation results are presented in Table. Through the evaluation, we identified opportunities for extending the criteria catalogue, such as with a cost and usability dimension.

1   Discussion

The application of the taxonomy showed that DMPs usually involve first-party and third-party data segmentation in real-time, sometimes with the use of machine learning, following the goal of consumer segmentation. Consumer segmentation makes it possible to implement marketing objectives such as serving targeted advertisements, displaying targeted content, political microtargeting, or personalised price discrimination (Badmaeva & Hüllmann, 2019; Klein & Hüllmann, 2018). It can be used as a foundation for placing advertisements or generating look-a-like data (Kamps & Schetter, 2018). The importance of clustering consumer data into segments in the real-time bidding process is reflected in the evaluation of our taxonomy. Data segmentation is the only data processing functionality that all nine evaluated DMPs provide.

Consumer segmentation and the subsequent placement of advertisements have two data-related success factors: data quality and data quantity. First, if the data quality is bad, individuals might have been added to the wrong cluster because the data is inaccurate or inconsistent. Second, the more data is available, the more fine-grained consumer profiles can become, ultimately increasing quality. In light of various web-tracking methods and data export options to other actors in the real-time bidding system, critics argue that DMPs operate in contrast to the individual's data privacy. However, our study shows that while DMPs are not necessarily required to comply

with the GDPR in their origin countries, most of them do comply with the GDPR because of their European customers who are subject to the GDPR.

The DMP's functionalities supplement the functionalities of other actors in the real-time bidding system, such as supply-, demand-side platforms, ad-networks, and data marketplaces. The option of publishers and/or advertisers to export data to demand- and supply-side platforms makes the real-time bidding process more efficient. Real-time analysis is an essential characteristic because it enables advertisers (or advertisers' supply-side platform) to determine in real-time, before the advertisement is placed, whether a publisher's page is related to the advertisement's content. In that way, advertisements can be placed in advertising spaces where consumers of the target group interact. Another advantage is that placing advertisements on unsuitable or reputation-damaging websites can be avoided (Kreutzer, 2018; Zawadzk & Groth, 2014). Without effective data management, the automated and real-time matching of advertisements and advertising spaces on websites and the realisation of targeted advertisement would hardly be possible. These targeted advertisments help to optimise the reach of advertising campaigns (Yuan et al., 2012), which is why DMPs are valued in practice. The performance of targeted advertising in real-time bidding systems depends on the quality and quantity of the used data. A scenario can occur in which not enough data is available for valuable insights. For such cases, DMPs integrate with data marketplaces to provide publishers and/or advertisers the opportunity to buy and integrate external data, enriching subsequent analyses by creating detailed digital consumer profiles or refining existing user data.

Through their characteristic functionalities, DMPs enable effective interaction among different platforms. First-party and third-party data import into a DMP and various export options are exemplary functionalities that facilitate cooperation among actors in the real-time bidding system. For those actors, even though some are direct competitors, the value of cooperating is more significant than not cooperating. Cooperating allows filling gaps in missing functionalities or an expanded portfolio of multiple demand-side platforms to optimize the output of the advertising campaign. Salesforce is, for example, originally known for its customer relationship software but expands its service by collaborating with Google, allowing its customers to use their customer relationship management data with Google Analytics to perform data analyses (Google & Salesforce, 2020). Further, Salesforce

customers can use Google's tracking methods to obtain more data and integrate it into the Salesforce DMP. Cooperating is additionally beneficial for DMPs if they want to offer a more extensive portfolio of platforms (e.g., cooperating with multiple demand-side platforms that offer a DMP themselves). A DMP that integrates different and multiple demand-side platforms has the advantage of reaching a broader range of users, as different DMPs can cooperate with different ad-networks or ad-exchanges, and provide customers with multiple options between different demand-side platforms (Oracle, 2020a; The Trade Desk, 2020).

In the end, DMPs can be seen as a double-edged sword. On the one side, they provide indirect value to consumers as they support the placement of targeted advertisements and thereby ensure that consumers only see the advertisements that are relevant to them. On the other side, DMPs can contribute to the efficient distribution of harmful distorted or fake content to consumer segments because placed advertisements during the real-time bidding process are not evaluated. In that way, misleading content or fake news can be spread. Creating look-a-like data makes this approach scalable by extending existing consumer segments with consumers that have similar profiles.

5   Conclusions, Limitations, and Outlook

With this study we contribute a taxonomy that helps to understand and distinguish DMP functionalities. The taxonomy establishes a common ground for discussions on implications, for example, with regards to data privacy, data security, and the manipulation of opinions and consumer behaviour. It further helps to grasp the DMP's role in real-time bidding systems. Through an inductive and deductive procedure, eight taxonomy dimensions were identified that cover the main functionalities and mechanisms of DMPs. To ensure the reliability and validity of our taxonomy, we evaluated it with nine selected DMPs. Our developed taxonomy and its application clarify which characteristics define a DMP and guide discussions about specific functions and mechanisms of DMPs. The comparison of different DMPs in course of the evaluation is additionally helpful for practitioners when choosing a particular DMP for adoption.

As a limitation, we note that the developed taxonomy categories and characteristics undergo change as data management platforms undergo change. For example, new legal regulations may be set in place over time, or new data sources emerge. In that sense, the taxonomy is only valid until change happens in data management platforms and new characteristics and dimensions emerge. Therefore, future research should investigate the changes that DMPs go through to update the proposed taxonomy. Besides this limitation, future research opportunities are available that extend and refine the taxonomy. We can specifically think of web-tracking, that would benefit from a refinement. In our study, we lacked information on the application of other tracking methods as well as which cookie types are used by DMPs. Altogether our study provides first insights into the functionalities and mechanisms and thereby the role of DMPs in the RTBS. Yet, more research is needed in this field and the meaning of remaining actors needs to be studied to fully understand real-time bidding systems and their implications.

References

Adobe. (2020). Daten-Management-Plattform. Adobe. Retrieved February 28, 2021, from https://www.adobe.com/de/experience-cloud/topics/data-management-platform-dmp.html

Andrew, F., Foo Kune, L., & Schmitt, E. (2020). Gartner Magic Quadrant for Ad Tech. Retrieved February 24, 2021, from http://www.mediamath.com/news/mediamath-named-a-leader-in-gartners-magic-quadrant-for-ad-tech/

Badmaeva, T., & Hüllmann, J. A. (2019). Investigating Personalized Price Discrimination of Textile-, Electronics- and General Stores in German Online Retail. Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI).

Berghel, H. (2018). Malice Domestic: The Cambridge Analytica Dystopia. The IEEE Computer Society, 51, 84–89. https://doi.org/10.1109/MC.2018.2381135

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian. Retrieved February 24, 2021, from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Cederholm, A., & Simpson, I. (2018). Warum First Party Daten eine der wertvollsten Informationsquellen für einen Marketer sind. PIWIK. Retrieved February 28, 2021, from https://piwikpro.de/blog/warum-first-party-daten-eine-der-wertvollsten-informationsquellen-fuer-einen-marketer-sind/

Chou, D. C., Bindu Tripuramallu, H., & Chou, A. Y. (2005). BI and ERP integration. Information Management & Computer Security, 13(5), 340–349. https://doi.org/10.1108/09685220510627241

Dawson, P. (2014). Realtime Advertising in Social Networks. In O. Busch (Ed.), Realtime Advertising: Digitales Marketing in Echtzeit: Strategien, Konzepte und Perspektiven (pp. 177–185). Springer Gabler. https://doi.org/10.1007/978-3-658-05358-1_1

Dun & Bradstreet. (2020). Dun & Bradstreet - Accelerate Growth and Improve Business Performance. Retrieved February 28, 2021, from https://www.dnb.com/

Google. (2020). Google Marketing Platform – Werbung und Analysen in einer Lösung. Retrieved February 28, 2021, from https://marketingplatform.google.com/intl/de/about/

Google, & Salesforce. (2020). Salesforce and Google Cooperation. Google;Salesforce. Retrieved February 28, 2021, from https://cloud.google.com/salesforce?hl=de

Hüllmann, J. A., Krebber, S., & Troglauer, P. (2021). The IT Artifact in People Analytics: Reviewing the Tools to Understand a Nascent Field. Proceedings of the 16th International Conference on Wirtschaftsinformatik (WI).

Joe, R. (2014). The Great DMP Debate. Ad Exchanger. Retrieved February 28, 2021, from https://www.adexchanger.com/data-exchanges/the-great-dmp-debate/

Kamps, I., & Schetter, D. (2018). Performance Marketing - Der Wegweiser zu einem mess- und steuerbaren Marketing – Einführung in Instrumente, Methoden und Technik (1st ed.). Springer Gabler. https://doi.org/10.1007/978-3-658-18453-7

Klein, S., & Hüllmann, J. (2018). Datenkapitalismus akademischer Wissenschaftsverlage. Wirtschaftsdienst, 98(7), 477–480. https://doi.org/https://doi.org/10.1007/s10273-018-2318-3

Kreutzer, R. T. (2018). Praxisorientiertes Online-Marketing: Konzepte - Instrumente - Checklisten (J. Blind & R. Stumpfrock (eds.); 3rd ed.). Springer Gabler. https://doi.org/10.1007/978-3-658-17912-0

Lange, J., Stahl, F., & Vossen, G. (2018). Datenmarktplätze in verschiedenen Forschungsdisziplinen: Eine Übersicht. Informatik Spektrum, 41, 170–180. https://doi.org/https://doi.org/10.1007/s00287-017-1044-3

Lotame. (2020). Data Management Platform. Retrieved February 28, 2021, from https://www.lotame.com/products/data-management-platform/

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. https://doi.org/10.1037/h0043158

Moffett, T., & Chien, A. (2019). The Forrester WaveTM: Data Management Platforms, Q2 2019. Forrester Research, Inc. Retrieved February 28, 2021, fro https://www.adobe.com/offer/forrester-dmp-wave-2019.html#

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

Oracle. (2020a). Media and Data Ecosystem. Retrieved February 28, 2021, from https://www.oracle.com/de/data-cloud/products/data-management-platform/ecosystem.html

Oracle. (2020b). Oracle Data Management Platform (formerly BlueKai). Oracle; Retrieved February 28, 2021, from https://www.oracle.com/de/data-cloud/products/data-management-platform/

Poleshova, A. (2017). Umfrage zur Nutzung und Nutzungsabsicht von Technologien im Marketing weltweit 2017. Retrieved February 28, 2021, from https://de.statista.com/statistik/daten/studie/745360/umfrage/nutzung-und-nutzungsabsicht-von-tech-nologien-im-marketing-weltweit/

Schonschek, O. (2020). So finden Sie die richtige Data-Management-Plattform. High Text Verlag. Retrieved April 15, 2021, from https://www.onetoone.de/artikel/db/138456cr.html

Spiekermann, M. (2019). Data Marketplaces: Trends and Monetisation of Data Goods. Intereconomics, 54(4), 208–216. https://doi.org/10.1007/s10272-019-0826-z

The Trade Desk. (2020). Data Management Platform (DMP). Retrieved February 28, 2021, from https://www.thetradedesk.com/products/data-management-platform-dmp

Wang, J., Zhang, W., & Yuan, S. (2017). Display advertising with real-time bidding (RTB) and behavioural targeting. Foundations and Trends in Information Retrieval, 11(4–5), 297–435. https://doi.org/10.1561/1500000049

Yuan, S., Abidin, A. Z., Sloan, M., & Wang, J. (2012). Internet Advertising: An Interplay among Advertisers, Online Publishers, Ad Exchanges and Web Users. Information Processing & Management, June. http://arxiv.org/abs/1206.1754

Zawadzk, V., & Groth, A. (2014). Granularität schafft Mehrwert für jedes Ziel. In O. Busch (Ed.), Realtime Advertising: Digitales Marketing in Echtzeit: Strategien, Konzepte und Perspektiven (pp. 145–160). Springer Gabler. https://doi.org/10.1007/978-3-658-05358-1_13

Zenith. (2019). Global Programmatic Advertising Spending from 2019 to 2021. Statista. Retrieved April 12, 2021, from https://www.statista.com/statistics/275806/programmatic-spending-worldwide/

Zhang, W., Yuan, S., & Wang, J. (2014). Optimal real-time bidding for display advertising. Proceedings of the 20th ACM SIGKDD, 1077–1086. https://doi.org/10.1145/2623330.2623633

Appendix

Evaluation of the Taxonomy: Table provides an overview of the nine DMPs that we selected from the longlist to be included in our study together with the respective sub-criteria with which we assessed the DMPs' relevance, feasibility, and popularity.

Table 2: Nine Data Management Platforms for Taxonomy Evaluation. In the revenue column, M stands for Million, and B stands for Billion, in the Awards Column, FW stands for Forrester Wave.

DMP	Relevance		Feasibility		Popularity	
	Year	Revenue	Info	Trial	Clients	Awards
Adobe	2019	11,17 B	+	+	148	Gartner
Google Audience	2020	160,74 B	+	-	-	Gartner
Lotame	2020	14,25 M	+	+	>8	FW
Mediamath DMP	2020	104,7 M	+	-	-	Gartner & FW
Nielsen	2018	6,50 B	+	-	-	FW
OneID von Neustar	2017	1,20 B	+	-	>75	FW
Oracle DMP	2019	39,50 B	+	-	6	FW
Salesforce DMP	2020	17,10 B	+	+	>150k	FW
The Trade Desk	2018	447 M	+	Video	10	Gartner

Sources: Andrew et al., 2020; Dun & Bradstreet, 2020; Moffett & Chien, 2019

Subjective Ending Conditions: The subjective ending conditions following Nickerson et al. (2013) are conciseness, robustness, comprehensiveness, extensibility, and explainability. Conciseness is used to limit the number of dimensions, thus focusing on the quality of the taxonomy. Our taxonomy on DMPs has eight dimensions, which from our point of view, is an appropriate number because a possible objective ending condition falls in the range of seven plus or minus two (Miller, 1956). Robustness evaluates whether the chosen dimensions and

criteria are meaningful in the application and allow the differentiation of the investigated objects. The dimensions and characteristics in the taxonomy point to the elementary functions that DMPs can have but that are not mandatory or a must-have and may therefore vary depending on the DMP. Thus, robustness is fulfilled. Comprehensiveness verifies that all dimensions are included and that new objects can be classified. We tested the taxonomy on nine DMPs and the results show that all nine can be classified using the DMP. Even though the sample is small we assume that the taxonomy's comprehensiveness is appropriate. Since we used popular DMPs for our evaluation, future studies should investigate whether the categories are comprehensive for DMPs that are less popular. Extensibility considers the simple extension of the taxonomy. This ending condition is given because the taxonomy's structure is easy, without any subdimensions in place. Researchers who find additional characteristics or dimensions to be necessary can build on our work. Finally, explainability evaluates how exactly the dimensions explain the object. The taxonomy takes into account the fundamental functionalities of DMPs (e.g., data import, data processing functions). In that way, the taxonomy explains what constitutes DMPs and what are their foci (e.g., using first-party data, consumer segmentation). Thus, we argue that also the fifth ending condition is fulfilled and with it our research goal to explore the essence of DMPs.

FAMILIARITY ATTRACTS CONSUMER ATTENTION: TWO METHODS TO OBJECTIVELY MEASURE CONSUMER BRAND FAMILIARITY

URSA BERNARDIC[1] & BENJAMIN SCHEIBEHENNE[2]

[1] University of Geneva, Geneva School of Economics and Management, Switzerland;
e-mail: ursa.bernardic@unige.ch
[2] Karlsruhe Institute of Technology, Institute for Information Systems and Marketing,
Germany; e-mail: scheibehenne@kit.edu

Abstract Brand familiarity is an important and frequently used concept in marketing research and practice. Existing measures of brand familiarity typically rely on subjective self-reports and Likert scales. Here we develop and empirically test two implicit measures to quantify brand familiarity. Based on research in visual attention and computer image processing, observers in a first visual search task are incentivized to quickly find a target brand among varying numbers of competitor brands. In the second approach, we measure the speed at which observers can identify a target brand that is gradually revealed. Both approaches are validated in preregistered experiments. Results show that reaction times predict brand familiarity on an individual level beyond conventional self-reports, even when controlling for "bottom-up" visual features of the brand logo. Our findings offer an innovative way to objectively measure brand familiarity and contribute to the understanding of consumer attention.

1   Introduction and theoretical background

Marketing practitioners and academics have long recognized the importance of brand familiarity – a consumer's prior direct or indirect experiences with the brand (Kent & Allen, 1994). Past research showed that brand familiarity impacts advertisement, evaluation, and consumer choice. Consequently, brand familiarity became a key element in models of brand strength and brand equity (e.g. Aaker, 1997; Erdem, 1998; Keller, 2013). For companies, measures of brand familiarity provide managerial insights on how their brand compares to competitors' brands, which allows for more effective marketing campaigns and hence sustain brand growth and sales (Nielsen Marketing Report, 2020). In marketing research, a substantial body of empirical work has focused on correlations between brand familiarity and other marketing constructs. Findings show that brand familiarity influences brand memory, explicit attitudes toward the brand, and advertising effectiveness in both, traditional advertising (Pieters, Warlop, & Wedel, 2002) and newer formats, such as movies (Brennan & Babin, 2004). Despite this importance of brand familiarity, an overview of recent academic and industry work reveals that brand familiarity has (almost) always been assessed using subjective self-reports, typically based on the Likert Scale where respondents indicate how familiar they are with a given brand with verbal anchors that range from not at all familiar/extremely unfamiliar to very familiar/extremely familiar (Zhou & Nakamoto, 2007).

While such self-reports are cost and time-efficient, they come with a number of disadvantages that limit their validity (De Houwer, 2006). Firstly, they rely on language and are thus not culture-free and can be prone to different interpretations and translations. This makes it hard to compare and aggregate answers in an increasingly globalized market where global brands need to assess and compare their familiarity across different languages, countries, and cultures. Secondly, subjective scales are also hard to incentivize and therefore are more prone to response-order effect (increasing the tendency for respondents to select the first response available to them on the answer scale), donkey vote effect (selecting the same response for all questions), demand effects (the tendency of respondents to respond positively), and dishonesty. However, even when consumers want to answer honestly, there is a mismatch between the way consumers experience and think about the world and the methods marketers use to collect this information (Zaltman, 2003). As such, subjective scales require introspective ability and the psychometric properties of data

from Likert-scales are debated (Li, 2013). For example, there is an ongoing debate about whether a Likert scale is ordinal or interval. A typical ordinal scale can measure the orders of the ratings, but it cannot tell us about the intervals between responses and thus results in information lost during measurement. On the other hand, a typical interval scale implies that the difference between any two consecutive scales reflects equal differences in the variable measured, which leads to information lost during measurement (Wu & Leung, 2017).

Building on recent findings and models of visual attention in marketing (Sample, Hagtvedt, & Brasel, 2020), and in cognitive science (Wolfe & Horowitz, 2017), we empirically test whether and how brand familiarity impacts visual attention, and bridge the gap by proposing an innovative way to objectively measure brand familiarity. Thus the current paper has two objectives. The first is to transfer knowledge from cognitive science and computer visual processing literature to a consumer behavior setting and test whether we can distill top-down effects from bottom-up factors with more ecologically familiar stimuli, brand logos. If this is possible, this will naturally lead to the establishment and proof of concept for an implicit measure of real-world personal brand familiarity. The second is to bring into the spotlight the construct of brand familiarity on visual attention.

2   Experiments

2.1   Experiment 1: Using Visual Search Task to study how brand familiarity impacts visual attention

Up to date, it remains an open question whether consumers would show advantages for familiar or on contrary, novel stimuli. In line with these inconsistent results, a recent review in the area of vision research called for more contribution to the understanding of how familiarity influences visual attention (Wolfe, 2020). Study 1 was designed to provide a first test of the hypothesis that brand familiarity as opposed to brand novelty impacts consumers' attention in a visual search task. Study 1 was preregistered prior to data collection on OSF (OSF link to preregistration).

Design and measurement

A total of 100 participants were recruited on the Mechanical Turk website and preselected based on not having color-blindness, owning a mobile phone, and having no extensive connection with China (MAge=35.9, SDAge=10.5, 37% females). Participants were instructed to find different smartphone app logos ("targets") among either 7 or 23 distractor logos. Half of the targets consisted of familiar app logos from the Google Play store, the other half consisted of unfamiliar logos from apps that were primarily used in China. Each participant completed 200 trials that were split into 4 blocks. At the beginning of each trial, the target logo was displayed in the center of the screen for 1 second. Then, a new screen appeared with a 5x5 grid that contained the target among the distractor logos at random positions. Figure 1 shows an example of this setup. As soon as participants identified the target brand on this grid they had to press a button on the keyboard. Once they pressed the button, all app logos disappeared from the grid and participants had to click on the grid position where the target had been displayed. For this last task, there was no time pressure. Participants received feedback after each trial in terms of points, which were exchanged into the monetary bonus at the end of the experiment. Familiar and unfamiliar logos were visually matched according to color, shape and number of characters and were pretested on a similar sample for familiarity. All logos were resized to 160x160 pixels to ensure the equivalent image size.

Figure 3: Screenshots of the visual search task in the first experiment

Procedure

The experiment started with color-blindness, the Ishihara test (Ishihara, 1994), and a screen resolution check. Next, participants read an instruction of the visual search task and completed a training block with 10 trails. After training, participants finished 200 experimental trials, separated into 50-trial blocks that corresponded to

each of four conditions. Besides the main experimental manipulation of target familiarity, we also manipulated distractor familiarity. These yield four experimental conditions in a 2x2 design: finding a familiar target among unfamiliar distractors (FU), familiar target among familiar distractors (FF), and vice versa (i.e. UF and UU). The order of the four blocks was counterbalanced between participants. All conditions were pseudorandomized among participants and randomized on the trial level. As a manipulation check at the end of the experiment, participants rated their familiarity with the app logos on a 4-point (1= unfamiliar to 4=familiar) scale.

Results and discussion

The familiarity ratings at the end of the experiment were significantly higher for the familiar app icons (mean=3.7, SD = 0.7) than the unfamiliar app icons (mean=1.2, SD=0.6), suggesting our manipulation of familiarity was successful.

In line with our prediction, familiar targets (median = 568ms) were found slightly faster than unfamiliar targets (median = 585ms). The average difference of 17 milliseconds and the corresponding effect size was very small though (Cohen's d = 0.08). To test which variables systematically influenced the reaction times, and to avoid the pitfalls of null-hypothesis testing (Baker, 2016) we estimated a linear mixed effect model using lme4 package in R (version 1.1-26, Bates et al., 2015). The regression model assumes random intercepts for individuals and accounts for possible dependencies due to the repeated measurement design. The baseline model includes intercept as a single fixed effect, location of the target, and set size (denoted as m0). Next, we added predicted variables of target familiarity rating as a fixed effect (denoted m1). A tested variable has a credible influence on prediction accuracy if adding it to the regression equation improves model fit. To select the best-fitting model we used Bayesian information criteria (BIC), which takes model complexity into account by introducing a penalty term for the number of parameters in the model. While absolute values of BIC are difficult for interpretation, BIC differences (ΔBIC) between models can be transformed into Bayes factors, which offer more intuitive explanations.

A regression model that contained familiarity as a categorical predictor of reaction times explained the observed data better than a baseline model that did not include the predictor. Table 1 (Exp 1) shows a comparing of both regression models. As can be seen from the table, the Bayesian Information Criteria (BIC) of the difference between a null model (m0) and model that included target familiarity (m1) was 18 (ΔBIC), which translates into a Bayes factor (BF) of 7104 (Wagenmakers, 2007). A BF of 7104 indicates that based on the observed data, the model including the target familiarity rating is 7104 times more probable.

Results from additional exploratory (i.e. not pre-registered) analyses suggest that distractor familiarity also impacts visual search. In particular, search efficiency was higher when the target and distractor familiarity do not match (FTUD and UTFD), than when they do match (FTFD and UTUD). Using a similar model comparison approach as above yields a Bayes factor (BF) of 602 in favor of the model that includes distractor familiarity as a predictor.

To summarize, Study 1 demonstrated that brand familiarity as opposed to brand novelty enhances visual attention. Thus, the results provide initial evidence that visual search tasks provide an implicit measure of brand familiarity.

Table 1: Regression coefficients of the best fitting mixed-effects models

Exp 1: Online study (US vs PCR brands)			
	m0	m1	m2
Fixed effects			
(Intercept)	18.556	18.741	18.646
Target Location	0.884	0.885	0.887
Set Size	0.266	0.266	0.266

Target Rating		-0.075	-0.075
Distractor Condition			0.190
Model fit			
log(likelihood)	-23182	-23168	-23157
AIC	46374	46349	46329
BIC	46410	46392	**46380**

Note. Best fitting models were chosen using BIC which punishes for model complexity and number of parameters. Corresponding Bayesian Factors (BF) for nested models were calculated as BF = exp(ΔBIC/2) (Wagenmakers, 2007).

2.1.1   Experiment 2: Developing a video recognition task

The findings of the first experiment provide converging evidence for the influence of brand familiarity on visual search efficiency. As such, it indicates that the experimental search paradigm that we used can be implemented as an implicit measure of brand familiarity. However, given the relatively small effect size, the task at hand requires relatively many repetitions within subjects. This limits its practical value, for example in the context of quick marketing surveys. To overcome these limitations, this second study at hand proposes an alternative implicit measure that is based on qualitative differences between the perception of familiar and unfamiliar stimuli (OSF link to preregistration). Previous research on face familiarity explored different manipulations of image stimuli, such as Gaussian blur, and contrast negation (Balas, Cox, & Conwell, 2007). While such image manipulation impaired perception for face familiarity studies, neither Gaussian blurring, linear stretching, nor contrast negation did not challenge the perception of brands (Sandford, Sarker, & Bernier, 2018). Therefore, we developed a perceptual decision task in which participants watch a video that gradually changes from a noisy mask to a given target

brand. Participants are instructed to press a button as soon as participants recognize the brand. The main dependent variables are reaction time and recognition accuracy.
Design and measurement

A total of 73 students from Swiss University, participated in the lab study (MAge=22.2, SDAge=1.99, 57% females). As preregistered, all participants who did have an extensive connection to another country (Switzerland for Slovene participants, and Slovenia for Swiss participants) or preferred that their data would not be included were excluded from further analysis.

Stimuli and Procedure

To test whether a video task could be used to distinguish the different levels of brand familiarities, we manipulated brand familiarity by using national brands from two countries, Switzerland and Slovenia. A pre-selected set of brand logos from both countries from the second study was presented to participants from Switzerland.

Participants in the experiment were instructed to detect a logo in a video, which went from pure noise to the target brand image. The noisy starting point was generated by drawing a random RGB value for each pixel of the video. As the video progressed, random sets of noisy pixels were gradually replaced with pixels from the target image. The number of "flipped" pixes was determined based on an inversely s-shaped function (we used a scaled beta function) such that at the beginning of the video many pixels were flipped while in the middle part the flip rate was decreased. Each participant saw 20 video sequences. The logos in the videos were randomly drawn from the set of 20 videos; representing 10 unfamiliar logos and 10 familiar logos.

To incentivize participants, they received trial-to-trial feedback in terms of points on their speed and accuracy after each round. More precisely, each video lasted for 20 seconds and participants started with 200 points, and for each millisecond participant took, we deducted 1 point from their score. If the participant didn't correctly identify the brand logo, or if the video was played for more than one time, the participant would receive 0 points. At the end of the experiment, the average of all points was exchanged for lottery tickets. Among all lottery tickets from all

participants, we selected one winning number, and participant holding that number received 100 CHF of bonus.
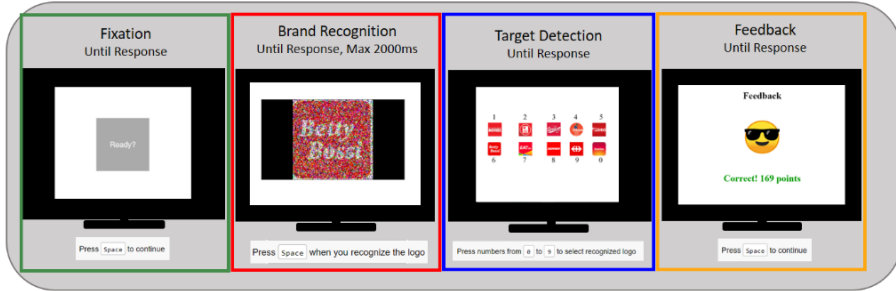
Figure 2: Screenshots of the video task.

Results and Discussion

In line with our prediction, Swiss participants were faster at finding swiss logos than Slovene logos in a video task. Using uninformative priors (Raftery, 1995) the difference between the Bayesian Information Criteria the baseline model (m0) and the extended model (m1) which included familiarity rating was 128, which translates into a Bayes factor (BF) of 4.4e+21 (Wagenmakers, 2007). In other words, familiar brands (swiss) were recognized more than 2 seconds faster (MRT =11.64, CIRT =0.25) than unfamiliar (Slovene) brands (MRT =13.87, CIRT =0.25)

In summary, Study 2 provides evidence that also video task can predict brand familiarity. It demonstrates that visual attention is driven by brand familiarity and that the video task at hand provides an implicit measure of brand familiarity

3   Final Discussion and implications

Our research sheds light on a prevalent, yet understudied research question of how brand familiarity affects visual attention. Across studies, we find converging empirical evidence that familiarity affects visual search efficiency. As a methodological contribution, we introduced two implicit measures of brand familiarity that rely on reaction times and thus avoid the pitfalls of subjective self-reports based on Likert scales. These paradigms can be used as a blueprint for researchers and practitioners alike. Moreover, all studies implemented incentive

alignment in which participants are rewarded for correct responses. Together with the use of real-world stimuli, this enhances the external validity of our findings. To the best of our knowledge, this is the first work that explores the effect of brand familiarity on search efficiency while controlling for distractor familiarity and bottom-up effects.

We hope that our work inspires future research in several notable ways. First, it would be interesting to find out if visual search tasks can be designed to quantify both, bottom-up and top-down effects. We see a variety of interesting and unanswered questions in the marketing and branding domain such as whether and how bottom-up features, such as previous location and other design features interacts with top-down features, such as familiarity and goals. Recent research on brand logo design sheds light on the effectiveness of many bottom-up features; however, it is only studied with unfamiliar brands (Lieven, Grohmann, Herrmann, Landwehr, & van Tilburg, 2015; Luffarelli, Stamatogiannakis, & Yang, 2019). With recent changes in logos of global brands (Google, Messenger from Facebook), empirical research on how these visual elements (bottom-up effects) interact with brand familiarity, is needed.

Second, we see the potential for the refinement and further development of objective measures based on implicit response times. Here, a nearby goal would be to further refine the proposed methods to achieve larger effect sizes and hence fewer repetitions within subjects. The current work further hints at the possibility of future research at the intersection between marketing and cognitive (neuro)science. For example, in visual attention, recent research by Sample, Hagtvedt, and Brasel (2019) could provide worthwhile theoretical foundations. In the long run, this work may also inform the effective design of new logos. In addition, future work may also test the observed top-down effects on visual attention with neuroscientific methods, and thus further strengthen the link between marketing and basic research on (visual) cognition and cognitive neuroscience.

To conclude, observations that familiar brands are detected faster than unfamiliar brands improve our understanding of effective marketing and brand positioning, especially in attention bottleneck settings (i.e. mobile phones, online shops, crowded shelves), where attention is a scarce resource. Current research also presents a new look at fundamental studies on how to objectively measure brand familiarity. The

methods presented here may help brand managers to monitor brand familiarity, give insights about visual attention attributes of their logos, and determine how their brands can stand out to consumers during the very short exposures they have in online settings and even more on mobile screens.

References

Aaker, J. L. (1997). Dimensions of brand personality. Journal of Marketing Research, 34(3), 347–356. https://doi.org/10.2307/3151897

Balas, B., Cox, D., & Conwell, E. (2007). The effect of real-world personal familiarity on the speed of face information processing. PLoS ONE, 2(11), 1–5. https://doi.org/10.1371/journal.pone.0001223

Brennan, I., & Babin, L. A. (2004). Brand placement recognition: The influence of presentation mode and brand familiarity. Handbook of Product Placement in the Mass Media: New Strategies in Marketing Theory, Practice, Trends, and Ethics, 6491, 185–202.

De Houwer, J. (2006). Handbook of implicit cognition and addiction. Sage Publications, Inc.

Erdem, T. (1998). An empirical analysis of umbrella branding. Journal of Marketing Research, 35(3), 339–351. https://doi.org/10.2307/3152032

Keller, K. L. (2013). and Measuring , Brand Managing Customer-Based Equity, 57(1), 1–22.

Kent, R. J., & Allen, C. T. (1994). Competitive Interference Effects in Consumer Memory for Advertising: The Role of Brand Familiarity. Journal of Marketing, 58(3), 97. https://doi.org/10.2307/1252313

Li, Q. (2013). A novel Likert scale based on fuzzy sets theory. Expert Systems with Applications, 40(5), 1609–1618. https://doi.org/10.1016/j.eswa.2012.09.015

Lieven, T., Grohmann, B., Herrmann, A., Landwehr, J. R., & van Tilburg, M. (2015). The effect of brand design on brand gender perceptions and brand preference. European Journal of Marketing, 49(1), 146–169. https://doi.org/10.1108/EJM-08-2012-0456

Luffarelli, J., Stamatogiannakis, A., & Yang, H. (2019). The Visual Asymmetry Effect: An Interplay of Logo Design and Brand Personality on Brand Equity. Journal of Marketing Research, 56(1), 89–103. https://doi.org/10.1177/0022243718820548

Malinowski, P., & Hübner, R. (2001). Malinowski&HÜbner(2001).pdf, 63(3), 458–463.

Pieters, R., Warlop, L., & Wedel, M. (2002). Breaking through the clutter: Benefits of advertisement originality and familiarity for brand attention and memory. Management Science, 48(6), 765–781. https://doi.org/10.1287/mnsc.48.6.765.192

Qin, X. A., Koutstaal, W., & Engel, S. A. (2014). The hard-won benefits of familiarity in visual search: Naturally familiar brand logos are found faster. Attention, Perception, and Psychophysics, 76(4), 914–930.

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. Sociological Methodology, 25(1995), 111–163.

Sample, K. L., Hagtvedt, H., & Brasel, S. A. (2020). Components of visual perception in marketing contexts: a conceptual framework and review. Journal of the Academy of Marketing Science, 48(3), 405–421. https://doi.org/10.1007/s11747-019-00684-4

Sandford, A., Sarker, T., & Bernier, T. (2018). Effects of geometric distortions, Gaussian blur, and contrast negation on recognition of familiar faces. Visual Cognition, 26(3), 207–222. https://doi.org/10.1080/13506285.2017.1407853

Shen, J., & Reingold, E. M. (2001). Visual search asymmetry : The influence of stimulus familiarity and low-level features, 63(3), 464–475.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive. Psychonomic Bulletin and Review, 14(5), 779–804.

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search, 56(5), 495–500.

Wolfe, J. M. (2020). Visual Search: How Do We Find What We Are Looking For? Annual Review of Vision Science, 6, 539–562. https://doi.org/10.1146/annurev-vision-091718-015048

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. Nature Human Behaviour, 1(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Wu, H., & Leung, S. O. (2017). Can Likert Scales be Treated as Interval Scales?—A Simulation Study. Journal of Social Service Research, 43(4), 527–532.

Zaltman, G. (2003). How customers think: Essential insights into the mind of the market. Harvard Business Press., 27(3), 480. https://doi.org/10.2501/s026504870808013x

Zhou, K. Z., & Nakamoto, K. (2007). How do enhanced and unique features affect new product preference? the moderating role of product familiarity. Journal of the Academy of Marketing Science, 35(1), 53–62. https://doi.org/10.1007/s11747-006-0011-3

Physical vs. Digital Interactions: Value Generation Within Customer-Retailer Interaction

Manuel Geiger, Franziska Jago &
Susanne Robra-Bissantz

Technische Universität Braunschweig, Chair of Information Management, Braunschweig, Germany; e-mail: m.geiger@tu-bs.de, f.jago@tu-bs.de, s.robra-bissantz@tu-bs.de

Abstract The traditional retail sector is currently facing major challenges, particularly due to digitalisation and the associated changes in customer behaviour, increasing demands in the service world, new technologies and other factors. The COVID-19 pandemic has accelerated and intensified this process. From a retailer's point of view, it is essential to create value for the customer through digital interactions. In this article, a study based on the *Value in Interaction Model* investigates whether it is possible for physical retailers to make a digitally supported interaction as valuable as the direct contact in the store and what influence this has on the Perceived Relationship Quality. The results show that the difference in perceived value between the physical and digital retailer interaction is relatively small. This proves that when the interaction layers are actively designed with a focus on value, a digital interaction can be almost as valuable as the traditional in-store interaction.

1   Introduction

Not only since the COVID-19 pandemic the stationary retail sector has been facing major challenges. Digitalization (Hagberg et al., 2016) and the accompanying changes in customer behaviour (Spaid & Flint, 2014), new and innovative competitors with disruptive approaches and advantages, increasing demands in the service world (e.g. same day delivery), new technologies (e.g. emotion-based IS support (Meyer et al., 2021)) and sales channels - all this means a major change to stationary retail and the associated traditional mechanisms and approaches. The development shows that retailers must avoid a further loss of customer contact (HDE, 2019) at all costs. We suggest that the interactions with the customers and their design must be placed in the center of attention. To create meaningful and valuable interactions, Geiger et al. (2020) have proposed the *Value in Interaction Model* (consisting of three layers: *Relationship Layer*, *Matching Layer* and *Service Layer* (see Figure 1). Customers access the digital offers of companies via digital interfaces, they use digital mediation platforms or comparison offers, inform themselves in web shops or via apps. Ultimately, a more or less successful and thus, valuable digital interaction then decides which products or services the customer chooses. It is no longer sufficient for a retailer to have only competences to deliver its standard service offering. The interactions should be actively designed on the three layers to generate (positive) value for all participating actors. As described, it becomes apparent that the stationary retail sector has major problems in designing value-added IT-supported interactions. While larger companies usually have both the financial and human resources to drive such developments, smaller ones often lack directly implementable solutions. One such potentially promising and easy-to-deploy service is the use of a messenger channel for customer communication.

In the context of this article, a study based on the *Value in Interaction Model* examines whether it is possible for physical retailers to make a digitally supported interaction just as valuable as the direct contact in the store and what effect this has on the Perceived Relationship Quality (PRC). After briefly explaining the *Value in Interaction Model* in the second section, section 3 deals with the PCR. Section 4 focuses on the methodology of the study and data collection before deriving the hypotheses in Section 5. While the results of the survey are presented in section 6, we finally draw a conclusion and give an outlook in section 7.

M. Geiger, F. Jago &S. Robra-Bissantz:
Physical vs. Digital Interactions: Value Generation Within Customer-Retailer Interaction

155

2   Value in Interaction

In marketing, theories like Service Logic (SL) or Service Dominant Logic (S-D logic) have been developed, which show companies how they can successfully design services in a very strongly customer-centric view (e. g. Grönroos, 2006; Vargo & Lusch, 2004). These service-centric theories have changed the way of thinking about what happens in business. The focus lies on the value for the customer, which is always created by a service. It is then no longer the provider with its product that creates value, but the value arises from the fact that the customer makes use of the provider's competences – called Value in Use (Grönroos, 2006). This value is measured solely from the added value that the customer derives from it. The dedicated consideration of the Value in Use of a service has proven to be a starting point for successful market offers. It is therefore obvious to also measure digital interactions by the value they offer for the customer. Wikström (1996) already pointed out that value is created in dialogue between actors within interactions. Interactions refer to practices in which actors are integrated into each other's processes (Grönroos & Ravald, 2011). They always should serve to realise specific purposes. However, the human being as a social being achieves a value in the interaction itself during communication. As described by Geiger et al. (2020b), the *Value in Interaction* can be created through digital services in the interaction and is based on the providers special competences
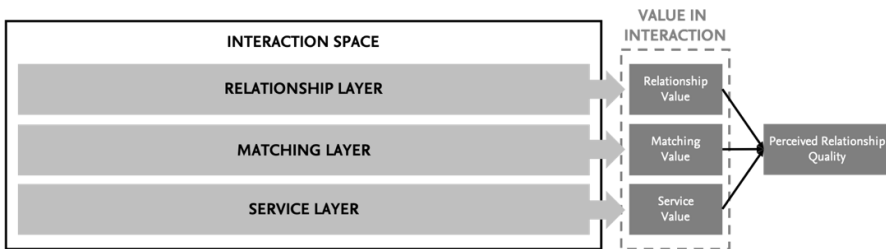
Figure 1: *Value in Interaction Model* and Perceived Relationship Quality.
Source: based on Geiger et al. (2020a)

The basis of any interaction is a connection between the actors in a shared *Interaction Space* (Grönroos, 2006), which can be provisioned by both actors. Such an *Interaction Space* can be the physical store of a retailer, but also a digital space, such as a website, an app or the usage of a messenger. Through interactions in this *Interaction Space*, the

actors have the opportunity to engage with the other actor or to influence their behaviour (Geiger et al., 2020a). However, this is only successful if the interaction is also seen as valuable by both actors (Fyrberg & Jüriado, 2009). From the provider's point of view, the goal is to open up *Interaction Spaces* with customers, to expand them if possible, or to be able to open them up again and again. The *Value in Interaction* arises within such an *Interaction Space*. It develops through and during the interaction, it unfolds its effect in the moment and thus influences the further processes of joint value creation (co-creation). In addition to the learning effects from successful interactions for follow-up interactions, value at the three layers also plays a long-term, direct role in the context of the actors' relationship (Geiger et al., 2021). However, the mere existence of an interaction does not lead to value. Rather, it depends on the quality of the interaction (Fyrberg & Jüriado, 2009). Thus, an interaction characterised by mediocre or even negative aspects (lack of quality) will have a negative impact on the PCR of the customer with the service or product (Geiger et al., 2021). Initial studies have shown that the composition of the *Value in Interaction* Model is basically suitable for significantly influencing the PRQ (Geiger et al., 2021). As a result, a large proportion of the PRQ can be explained by the *Value in Interaction*. This shows that the *Relationship Layer*, *Matching Layer* and *Service Layer* should be taken into account from a company's perspective when designing any interaction. If a company, and here in particular the bricks and mortar retail, manages to satisfy the needs of the customer on the individual layers in the interaction, this positively influences the PRQ. In addition to the actual competences in service provision, this requires further competences in order to be able to actively shape the individual layers of the *Value in Interaction*. Interactions that are adapted to the needs are thus relevant in order to build a high relationship quality between actors.

3   Perceived Relationship Quality

From a business perspective, an interaction with customers should always positively influence the relationship between the actors in order to contribute to shaping a long-term relationship. The Relationship Value described in the *Value in Interaction Model* consists, among other things, of the relationship-relevant advantages and disadvantages that the customer makes use of (Cronin et al., 1997; Dodds et al., 1991; Grewal et al., 1998). A relationship-relevant advantage can be, for example, higher esteem and a more confidential relationship, a disadvantage a resulting dependency. However, the advantages are not derived from the value, but primarily

M. Geiger, F. Jago &S. Robra-Bissantz:
Physical vs. Digital Interactions: Value Generation Within Customer-Retailer Interaction

157

from the PRQ (Hennig- Thurau & Klee, 1997). This thus depends closely on the expectations of both parties as well as their subjective evaluation of the satisfaction - and this concretely in every contact, every interaction between the actors. Thus, the value resulting from the *Relationship Layer* in the *Value in Interaction Model* is relatively more important in the initial stage of a relationship than at a later stage. The longer the relationship lasts, the more important the PRQ becomes. As already proven (Geiger et al., 2021), it can therefore be assumed that a successful interaction, which results in a positive *Value in Interaction*, also positively influences the PRQ and is thus a cornerstone for a long-term customer relationship.

Relationship quality consists of several components, on which research is largely unanimous. Based on the long-term accepted view of Hennig-Thurau & Klee (1997) (based on e.g. Crosby et al. (1990), Dorsch et al. (1998), Garbarino & Johnson (1999) and Smith (1998)) PRQ can be measured by (1) customer satisfaction, (2) the trust of the customers and (3) commitment to the relationship. The PRQ of the two parties involved has a significant impact on the duration and intensity of the underlying relationship (Hennig- Thurau & Klee, 1997). Accordingly, the PRQ is one of the most important determinants in the evaluation of a relationship in terms of permanence and intensity.

4   Research Methodology and Data Collection

This paper aims to find out whether a digital interaction between retailer and customer leads to a comparable value generation in the context of *Value in Interaction* and what effect this has on the PRQ. In the following section, the research methodology used, and the data collection are presented.

4.1   Operationalisation of the Model Scales

In order to obtain robust results, validated scales from existing research were used for data collection. The scales used in this paper for the components under investigation are well established and well founded in the literature. The scale of *Relationship Value* is captured by the "Relationship Value" (RV), which is represented by the four items from (Nguyen & Nguyen, 2011). The *Matching Value* is measured via the "Decision Convenience" (DC), "Access Convenience" (AC) and the "Benefit Convenience" (BC) (Colwell et al., 2008). Based on the work of Ruiz et al. (2008),

the *Service Value* consists of the "Service Value" (SV), the "Service Quality" (SQ) and the "Perceived Sacrifice" (PS). PRQ, as already mentioned, consists of the three scales "Relationship Satisfaction" (RS), "Trust" (TR) and "Commitment to the Relationship" (CR) (Adjei et al., 2010; Morgan & Hunt, 1994; Wulf et al., 2001). All scales were translated from English into German using DeepL[1] and adapted to the scenarios. In addition, items with inappropriate content were excluded due to the subject of the study. All scales were measured on a 5-point Likert scale (5=strongly agree). The survey also asked about gender, age and if they can empathise well with the described situation. Internal consistency for each of the scales was examined using Cronbach's alpha. By eliminating items, a substantial increase in alpha could be achieved (see section 6).

4.2 Data Collection and Sample

The data was collected via an online survey conducted in German and distributed via various mailing lists of a German university and via platforms like SurveyCircle and Pollpool[2]. Before the survey was made available to the public at the end of December 2020, a pretest was held with five participants. The actual survey took three weeks. Participation was voluntary in all cases.

Two different scenarios for (a) interaction in physical retail (scenario 1) and (b) digital interaction of physical retail via messenger (scenario 2) were described in detail. In order to make it easier for the test persons to empathies, the textual description of the situation was underpinned with pictures. The use cases were about a gift search for a third person one day before Christmas. In scenario 1 (S1), the consultation took place in a bookstore, in scenario 2 (S2) the bookstore interacted via WhatsApp Messenger. To ensure comparability, the interaction via messenger was identical to the interaction in physical retail. Where an exact transfer of the physical interaction into the digital interaction was not possible, adequate services were used (e. g. direct takeaway of the gift in scenario 1 vs. same-day delivery in scenario 2). The allocation to the two scenarios was done randomly with a probability of 50 % in each case. The data was analysed using SPSS Statistics 25.

---

[1] DeepL: www.deepl.com
[2] https://www.surveycircle.com/de/ and https://www.poll-pool.com/: Study dissemination platform for generating participants

M. Geiger, F. Jago &S. Robra-Bissantz:
Physical vs. Digital Interactions: Value Generation Within Customer-Retailer Interaction

159

150 participants completed the entire questionnaire. Five data sets had to be eliminated due to uniform response behaviour. In the end 145 valid responses were available. The demographic information on the sample and other characteristics of the subjects are listed in Table 1.

Table 1: Demographic Information

Characteristic	Distribution	
Sample Size	145 (scenario 1: 74, scenario 2: 71)	
Age	Range:	19 – 59 years
	Mean:	28,4
	Median:	26,0
Sex	male:	53 (scenario 1: 28, scenario 2: 25)
	female:	92 (scenario 1: 46, scenario 2: 46)
	other:	0
Empathise well with the situation	scenario 1: 94,6% (70 participants)	
	scenario 2: 78,9% (56 participants)	

5   Derivation of hypotheses

The direct social interaction in 1-to-1 counselling in physical retailing has advantages in shaping the relationship. Creating an equivalent experience in the digital space seems more difficult due to the lack of human interaction and related physically visible expressions (e. g. emotions via voice pitch or body language) (Otto & Chung, 2000). It can therefore be assumed that direct interaction in physical retail has an advantage over digital interaction in shaping the interaction on the *Relationship Layer* and thus in achieving value.

H1: Physical retail interaction can achieve higher *Relationship Value* than a digital interaction performed by physical retail.

Information is needed to match the interaction components with the needs of the actor to design the relationship and service layer. This information about the customer is either already available in physical retail or it is the responsibility of the sales staff to find it out. In the context of this study, a stand-alone interaction was investigated. Thus, there was no existing information about the customer and the concrete needs. Due to the personal interactions and direct responses, it must be

assumed that interaction in physical retail has an advantage over digital interaction in shaping the interaction on the *Matching Layer* and thus in achieving value.

H₂: Physical retail interaction can achieve higher *Matching Value* than a digital interaction performed by physical retail.

If the goal of a customer is to find a gift, as in the context of the study conducted, physical retail can also express its advantages. Inspiration is a core function in retailing (Böttger, 2015). With creating a stimulating shopping environment and due to the service of physical examination and direct availability (Otto & Chung, 2000) physical retailer can inspire their customers in their stores. Therefore, it can be assumed that direct interaction in physical retail has an advantage over digital interaction in shaping a valuable *Service Layer*.

H₃: Physical retail interaction can achieve higher *Service Value* than a digital interaction performed by physical retail.

As already examined by Geiger et al. (2021) in a recent study, the three layers of *Value in Interaction* have an influence on the PRQ. Following the explanations of the preceding hypotheses, it can therefore also be assumed that direct interaction in physical retail has an advantage over digital interaction when it comes to PRQ.

H₄: Physical retail interaction can achieve a higher PRQ than a digital interaction performed by physical retail.

6   Results

For the following comparison of the two scenarios on the different layers of *Value in Interaction* and PRQ, different statistical methods were used. To ensure valid results, the internal consistency of the scales was checked using Cronbach's α. Thus, no item of the *Relationship Value* scale, two of the eight items in the *Matching Value* (MV) scale for S1, two of the seven items of the *Service Value* (SV) scale for S1 and S2 and three (S1) respectively one (S2) of the 16 items of the PRQ scale had to be eliminated. The data was tested for normal distribution using the Shapiro-Wilk Test. The results indicate a non-normal distribution for all scales ($p<0.01$). Since ordinal scaled data was analysed, the Mann-Whitney-U-Test (U) was used to find out whether the central tendencies of the independent samples differ. Since the sample

M. Geiger, F. Jago &S. Robra-Bissantz:
Physical vs. Digital Interactions: Value Generation Within Customer-Retailer Interaction

161

is larger than 30, we report the asymptotic 2-sided significance. The results are shown in Table 2.

Table 2: Statistical results

		α	Mean Rank	*U*	*Z*	Sig.[3]	Mdn	*r*
RV	S1	.797	77.54	2291.000	-1.337	.181	-	-
	S2	.824	68.27				-	
MV	S1	.876	83.89	1821.500	-3.195	.001***	4.67	.265
	S2	.853	61.65				4.13	
SV	S1	.894	83.19	1873.000	-2.997	.003***	4.50	.249
	S2	.864	62.38				4.17	
PRQ	S1	.900	83.98	1814.500	-3.215	.001***	3.85	.267
	S2	.903	61.56				3.53	

Significance level (two-tailed): *** < 1 %

Relationship Value: For *Relationship Value* there was no statistically significant difference in S1 and S2, $U = 2291.00$, $Z = -1.337$, $p = .181$. *H₁ must be rejected for this reason.*

Matching Value: A comparison of the two mean ranks between S1 (83.89) and S2 (62.38) shows that the two groups might have a different central tendency. The *Matching Value* is higher with the physical interaction; exact Mann-Whitney-U-Test: $U = 1821.500$, $p = .001$. *H₂ can thus be confirmed.*

Service Value: Again, a comparison of the two mean ranks between S1 (82.19) and S2 (62.38) shows that the two groups might have a different central tendency. The *Service Value* is higher with the physical interaction; exact Mann-Whitney-U-Test: $U = 1873.000$, $p = .003$. *H₃ can thus be confirmed.*

Perceived Relationship Quality: Finally, when comparing the two mean ranks between S1 (83.98) and S2 (61.56), it can be reported that the two groups might have a different central tendency as well. The PCR is higher with the physical interaction; exact Mann-Whitney-U-Test: $U = 1814.5000$, $p = .001$. *H₄ can therefore also be confirmed.*

---

[3] Asymptotic 2-sided significance

7   Conclusion and Outlook

Due to the high values, the results basically show that the use cases were suitable to represent a valuable interaction in retail. In the end, three of the four hypotheses were confirmed. While no statement can be made about the value creation on the *Relationship Layer*, the physical interaction manages to generate more value on the *Matching Layer* and the *Service Layer*. As was to be expected, this also leads to a higher PRQ. The biggest difference was .45 on the *Matching Layer*. At this layer, it thus seems to be a particular challenge to find out the exact need of the customer in the context of a digital interaction. This is also understandable, as the use case in question here involved an initial contact between the customer and the retailer. It can be assumed that on the basis of several successive interactions, a better knowledge base can be created by data storing and interpreting the different interactions. Surprisingly, however, the difference in perceived value (*Matching Value* and *Service Value*) between the physical and the digital retailer interaction is rather small. This shows that when the layers are actively designed with a focus on value, a digital interaction can be almost as valuable as the traditional in-store interaction. Previous studies have shown that the three layers of *Value in Interaction* are capable of significantly influencing the PRQ (Geiger et al., 2021). PRQ for digital retail interaction is .32 lower than physical interaction. So, when it comes to relationship quality, the additional benefit between the different interaction channels also seems to be low. In order to be able to actively shape the individual layers of the *Value in Interaction*, further competences are required in addition to the actual competences in standard service delivery (Geiger et al., 2020b). In addition, many former customers are no longer (physical) accessible to retailers due to declining customer frequency (HDE, 2019), with the COVID-19 pandemic accelerating this process by five years (IBM, 2021). Customer behaviour itself is changing (Spaid & Flint, 2014) and especially the younger prefer to shop online instead (Sabanoglu, 2017). Accordingly, it is all the more important for retailers to place digital interactions and their valuable design at the heart of their business. With regard to the limitations, it must be taken into account that the scales used were created by different authors and thus may have been perceived differently by the participants. The extent to which it is possible to achieve a higher value with digital interactions or whether digitally supported interactions (digital plus direct interaction) are the best way to generate value should be further researched. Even though there are already initial studies on the impact of the three layers on PRQ (Geiger et al., 2021), a precise analysis of this relationship

M. Geiger, F. Jago &S. Robra-Bissantz:
Physical vs. Digital Interactions: Value Generation Within Customer-Retailer Interaction

163

should be carried out in the context of the use cases described here. In further research, the concrete influencing components of an interaction are also to be identified in order to develop concrete guidelines and design patterns for the active design of interactions on this basis. In addition, the technologies currently discussed in IS and their applications such as emotion recognition (Meyer et al., 2019), personality mining (Ahmad et al., 2021), AI or chatbots are to be examined in relation to the *Value in Interaction Model*. The aim is to find out how these technologies have to be integrated into the interactions between retailers and customers in order to generate value and what contribution they make to the PRQ in comparison to each other.

References

Adjei, M. T., Noble, S. M., & Noble, C. H. (2010). The influence of C2C communications in online brand communities on customer purchase behavior. Journal of the Academy of Marketing Science, 38(5), 634–653. https://doi.org/10.1007/s11747-009-0178-5

Ahmad, R., Siemon, D., & Robra-Bissantz, S. (2021). Communicating with Machines: Conversational Agents with Personality and the Role of Extraversion. 11. https://doi.org/10.24251/HICSS.2021.492

Böttger, T. (2015). Inspiration in Marketing: Foundations, Process, and Application [University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs]. https://www1.unisg.ch/www/edis.nsf/SysLkpByIdentifier/4395/$FILE/dis4395.pdf

Colwell, S. R., Aung, M., Kanetkar, V., & Holden, A. L. (2008). Toward a measure of service convenience: Multiple-item scale development and empirical test. Journal of Services Marketing, 22(2), 160–169. https://doi.org/10.1108/08876040810862895

Cronin, J. J., Brady, M. K., Brand, R. R., Hightower, R., & Shemwell, D. J. (1997). A cross-sectional test of the effect and conceptualization of service value. Journal of Services Marketing, 11(6), 375–391. https://doi.org/10.1108/08876049710187482

Crosby, L. A., Evans, K. Rr., & Cowles, D. (1990). Relationship Quality in Service Selling: An Interpersonal Influence Perspective. Journal of Marketing, 54(3), 68–81. https://doi.org/10.2307/1251817

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. Journal of Marketing Research, 28(3), 307–319. https://doi.org/10.2307/3172866

Dorsch, M. J., Swanson, S. R., & Kelley, S. W. (1998). The Role of Relationship Quality in the Stratification of Vendors as Perceived by Customers. Journal of the Academy of Marketing Science, 26(2), 128–142. https://doi.org/10.1177/0092070398262004

Fyrberg, A., & Jüriado, R. (2009). What about interaction?: Networks and brands as integrators within service-dominant logic. Journal of Service Management, 20(4), 420–432. https://doi.org/10.1108/09564230910978511

Garbarino, E., & Johnson, M. S. (1999). The Different Roles of Satisfaction, Trust, and Commitment in Customer Relationships. Journal of Marketing, 63(2), 70–87. https://doi.org/10.2307/1251946

Geiger, M., Moser, J.-M., & Robra-Bissantz, S. (2021). Value in Interaction meets Relationship Quality: Der Wert von Interaktionen für eine langfristige Kundenbeziehung. HMD Edition 2020, 10.

Geiger, M., Robra-Bissantz, S., & Meyer, M. (2020a). Focus on Interaction: Applying Service-Centric Theories in IS. 33RD BLED EConference Enabling Technology for a Sustainable Society, 8. https://doi.org/10.18690/978-961-286-485-9.46

Geiger, M., Robra-Bissantz, S., & Meyer, M. (2020b). Wie aus digitalen Services Wert entsteht: Interaktionen richtig gestalten. HMD Praxis der Wirtschaftsinformatik, 57(3), 385–398. https://doi.org/10.1365/s40702-020-00611-0

Grewal, D., Monroe, K. B., & Krishnan, R. (1998). The Effects of Price-Comparison Advertising on Buyers' Perceptions of Acquisition Value, Transaction Value, and Behavioral Intentions. Journal of Marketing, 62(2), 46–59. https://doi.org/10.2307/1252160

Grönroos, C. (2006). Adopting a service logic for marketing. Marketing Theory, 6(3), 317–333. https://doi.org/10.1177/1470593106066794

Grönroos, C., & Ravald, A. (2011). Service as business logic: Implications for value creation and marketing. Journal of Service Management, 22(1), 5–22. https://doi.org/10.1108/09564231111106893

Hagberg, J., Sundstrom, M., & Egels-Zandén, N. (2016). The digitalization of retailing: An exploratory framework. International Journal of Retail & Distribution Management, 44(7), 694–712. https://doi.org/10.1108/IJRDM-09-2015-0140

HDE. (2019). Online-Monitor 2019. Handelsverband Deutschland. https://einzelhandel.de/index.php?option=com_attachments&task=download&id=10168

Hennig-Thurau, T., & Klee, A. (1997). The impact of customer satisfaction and relationship quality on customer retention: A critical reassessment and model development. Psychology & Marketing, 14(8), 737–764. https://doi.org/10.1002/(SICI)1520-6793(199712)14:8<737::AID-MAR2>3.0.CO;2-F

IBM. (2021). IBM U.S. Retail Index. https://www.ibm.com/industries/retail

Meyer, M., Helmholz, P., Rupprecht, M., Seemann, J., Tönnishoff, T., & Robra-Bissantz, S. (2019). From the Inside Out: A Literature Review on Possibilities of Mobile Emotion Measurement and Recognition. 27. https://aisel.aisnet.org/bled2019/23

Meyer, M., Siemon, D., & Robra-Bissantz, S. (2021). Emotion-based IS support for Customer-Salesperson Interaction. 10. https://doi.org/10.24251/HICSS.2021.538

Morgan, R. M., & Hunt, S. D. (1994). The Commitment-Trust Theory of Relationship Marketing. Journal of Marketing, 58(3), 20–38. https://doi.org/10.2307/1252308

Nguyen, T. D., & Nguyen, T. T. M. (2011). Enhancing relationship value between manufacturers and distributors through personal interaction: Evidence from Vietnam. Journal of Management Development, 30(4), 316–328. https://doi.org/10.1108/02621711111126800

Otto, J. R., & Chung, Q. B. (2000). A Framework for Cyber-Enhanced Retailing: Integrating E-Commerce Retailing with Brick-and-Mortar Ret. 8.

Ruiz, D. M., Gremler, D. D., Washburn, J. H., & Carrión, G. C. (2008). Service value revisited: Specifying a higher-order, formative measure. Journal of Business Research, 61(12), 1278–1291. https://doi.org/10.1016/j.jbusres.2008.01.015

Sabanoglu, T. (2017). Online shopping preference in the United States as of 2017, by age group. Statista.Com. https://www.statista.com/statistics/242512/online-retail-visitors-in-the-us-by-age-group/

Smith, J. B. (1998). Buyer–Seller relationships: Similarity, relationship management, and quality. Psychology & Marketing, 15(1), 3–21. https://doi.org/10.1002/(SICI)1520-6793(199801)15:1<3::AID-MAR2>3.0.CO;2-I

Spaid, B. I., & Flint, D. J. (2014). The Meaning of Shopping Experiences Augmented By Mobile Internet Devices. Journal of Marketing Theory and Practice, 22(1), 73–90. https://doi.org/10.2753/MTP1069-6679220105

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a New Dominant Logic for Marketing. Journal of Marketing, 68(1), 1–17. https://doi.org/10.1509/jmkg.68.1.1.24036

Vargo, S. L., & Lusch, R. F. (2008). Service-dominant logic: Continuing the evolution. Journal of the Academy of Marketing Science, 36(1), 1–10. https://doi.org/10.1007/s11747-007-0069-6

M. Geiger, F. Jago &S. Robra-Bissantz:
Physical vs. Digital Interactions: Value Generation Within Customer-Retailer Interaction

165

Wikström, S. (1996). Value creation by company-consumer interaction. Journal of Marketing Management, 12(5), 359–374. https://doi.org/10.1080/0267257X.1996.9964422

Wulf, K. D., Odekerken-Schröder, G., & Iacobucci, D. (2001). Investments in Consumer Relationships: A Cross-Country and Cross-Industry Exploration. Journal of Marketing, 65(4), 33–50. https://doi.org/10.1509/jmkg.65.4.33.18386

ALTERNATIVE DATA FOR CREDIT RISK MANAGEMENT: AN ANALYSIS OF THE CURRENT STATE OF RESEARCH

JAN ROEDER

University of Goettingen, Faculty of Business and Economics, Goettingen, Germany;
e-mail: jan.roeder@uni-goettingen.de

Abstract Determining credit risk is important for banks and non-banks alike. For credit risk management, the heterogeneous data generated today can potentially complement the established data such as balance sheet ratios. It has not yet been clearly shown which alternative data sources, such as social media or satellite data, provide added value and how this value can be extracted effectively. This review provides an overview of the intersection between these areas and develops a research agenda. The analysis of the 29 identified papers shows that the use of financial news is analyzed most frequently. Social media has also been used to some extent. The use of other alternative data sets, such as geospatial data, has been analyzed infrequently. The empirical evidence suggests that alternative data can provide both explanatory and predictive benefits in credit risk management. Convergence in terms of analytical approaches and evaluation offers the potential to advance the field.

1   Introduction

Determining credit risk is an essential task for banks and non-banks. For example, non-banks need to monitor their accounts receivables exposure and banks judge the creditworthiness of borrowers (Koulafetis 2017). In addition to established credit risk indicators (i.e., balance sheet ratios or market-based indicators (Altman and Saunders 1997)), the interest in using supplementary alternative data sources (ranging from sensor to social media data) has increased. The promising aspect is a potentially low rivalry of such data for risk management, i.e., the usefulness does not necessarily diminish with increasing dissemination of the data (Monk et al. 2019). Additionally, Mengelkamp et al. (2015) call for a further investigation of how user-generated content, which can be understood as a type of alternative data, can be integrated into corporate credit risk analysis. Although several literature reviews are situated at the intersection of alternative data (specifically text data) and finance research in general (Loughran and McDonald 2016; Nassirtoussi et al. 2014; Xing et al. 2017), there is none for the intersection of *credit risk* and *alternative data* specifically. Therefore, this paper's underlying research questions (RQ) are: 1) *What is the current state of research on using alternative data for supporting credit risk management?* And building on that: 2) *What are research gaps that future research should address?* After outlining the theoretical foundations, I define relevant parameters for the literature review. Afterwards, the results are presented and analyzed. Based on the gained insights, a research agenda is formulated, which can help to guide future research.

2   Theoretical Background

2.1   Alternative and Heterogeneous Data

Alternative data describes potentially decision-relevant but underutilized data sources, which are only available in unstructured form and cannot be used in established forecasting or risk models without prior processing (Monk et al. 2019). The potential originates from the idea that these data sources can contain important signals, for example, to identify changed customer behavior or risk situations. Especially because social media posts or anonymized credit card transactions occur with a higher frequency compared to more traditional information sources like earnings conference calls, these data sources could help to improve our risk understanding (Monk et al. 2019). The spectrum of alternative data ranges from app

usage data, anonymized credit card transactions, point of sales data, or job advertisements to data on the utilization of cruise ships (alternativedata.org 2020). Roughly classified, such data can originate from individual processes, business processes, or even sensors (Monk et al. 2019). However, the universe of alternative data is so broad that no exhaustive enumeration can be provided. It should also be noted that alternative data is ultimately a collective term. There may be differences in the volume, granularity, relationality, or accuracy of the data (Monk et al. 2019; Roeder et al. 2020). This also means that different techniques for processing, storage, and analysis may be necessary.

2.2   Credit Risk Management

Credit risk describes the threat that a borrower does not repay a granted loan or fails to meet contractual obligations (Caouette et al. 2008). Types of financial risk besides credit risk include strategic risk, market risk, or compliance risk (Lam 2014). The credit risk management (CrRM) process includes the identification of risk, credit risk assessment (CRA), treatment of risk, and implementation of actions (Van Gestel and Baesens 2008). In CRA, a distinction can be made between accounting-based models (e.g., Z-score by Altman (1968)) and market-based models (Das et al. 2009). Empirical findings indicate that both approaches can be combined to increase the explanatory power (Das et al. 2009). Since the second Basel accord permits banks to use internal-rating-based approaches, they use more advanced methods (McNeil et al. 2015). Machine learning models have become increasingly relevant for credit risk prediction in recent years, of which Chen et al. (2016) provides an overview. A fundamental distinction is made based on the analyzed entity. Publicly traded companies follow strict requirements regarding corporate disclosure. Hence, accounting ratios, market-based metrics (e.g., credit default swaps (CDS), bond spreads), and credit ratings can be utilized. For private companies, the available data universe is more limited. It includes credit ratings, past transactions, industry specifics or the quality of management (Schumann 2002). However, this paper does not analyze private individuals. In a broader sense, credit risk could include all literature on stock price forecasting, etc. For this *broad* perspective, please refer to existing research (Loughran and McDonald 2016; Nassirtoussi et al. 2014; Xing et al. 2017). In contrast, this literature review focuses on research dealing with the *credit risk* of public and private companies using the automated analysis of *alternative data*. Table summarizes the scope of this paper.

Table 1: Scope of the literature review (bold is examined in more detail in this paper)

Broadness of (credit) risk understanding	Risk management		**Credit risk management**		Credit risk assessment
Entity assessed	**Public comp.**	**Private comp.**		Individual	Sovereign

3   Research Method

This study builds on the work on literature reviews by Webster and Watson (2002) and vom Brocke et al. (2009) to synthesize existing research on alternative data for CRA. First, I define criteria for the relevance of papers. Table shows the inclusion criteria. For research to be relevant, it must examine the intersection between CrRM and alternative data. The focus lies on underexplored data sources such as news, social media, or even sensor data.

Table 2: Relevance criteria applied in the literature search

Required inclusion condition	Deals with CrRM (e.g., identifies, assesses, or monitors credit risk) of companies (public and private).
Label "Relevant"	Uses alternative data (e.g., satellite images, …) that occurs irregularly/with high frequency, uses innovative approaches.
Label "Borderline"	Uses established data sources (e.g., form 10-K data) but utilizes interesting approaches transferable to alternative data.

A comprehensive database search is conducted to identify relevant prior work. Conference and journal papers are included since the research area under investigation is still quite young. The search process is shown in Figure.
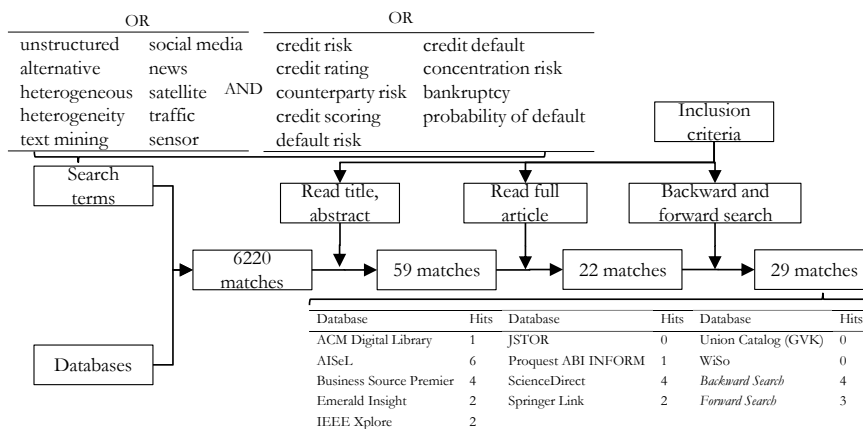
Figure 1: Literature search process

4 Results

4.1 Result Analysis

The papers in **Error! Reference source not found.** show the following **discipline** distribution (not shown in table): Finance (11), information systems (12), and computer science (6). The moderate number of papers can be attributed to the novelty of this research area. The availability of data sources is tightly coupled with the **entity** whose credit risk is **assessed**. As a side note, papers tend to exclude banks because their financial metrics are distributed differently and therefore may not be comparable (e.g., Lu et al. (2015)). The number of papers analyzing public companies is distributed relatively evenly between non-banks (27) and banks (19). Not publicly traded companies were analyzed less frequently (6). The reason could be that more data is available for exchange-listed companies due to the disclosure obligations. Regarding the **granularity** of the analysis, eight studies are carried out at the annual level and five at the quarterly level. In these studies, accounting ratios can be integrated into the analysis straightforwardly. The situation is different for the analyses at a monthly (1), weekly (2), and daily level (4), where traditional regression analyses can be problematic due to autocorrelation issues when using quarterly numbers. Regarding alternative **data sources**, it is apparent that financial news (11) has been analyzed the most, followed by annual reports including 10-K filings (8) and posts on social media (6). Other data sources are transactions on B2B platforms

(1), search engine sentiment (1), or data on staffing decisions (1). No research on using textual financial analyst reports for CrRM could be identified. For the **dependent variables**, a categorical variable is used eight times and the credit rating by rating agencies seven times. The CDS spread is also used seven times as a dependent variable. In addition, the bond spread (2), the LIBOR-OIS spread (1), and equity volatility (1) need to be mentioned. One difference between the two main categorical measures and the CDS spread is that the CDS spread is a continuously updated market-based measure.

Error! Reference source not found. also shows the **statistical models**, which link the endogenous variable to the signals extracted from alternative data sources. Finance-centric literature often uses logistic, linear, and panel regression to ensure sufficient interpretability. Here, the clustering of the data regarding the company and time period tends to be modeled explicitly, as is done using panel regression models in seven cases (e.g., Bao and Datta (2014); Liebmann et al. (2016); Tsai et al. (2016)). Contributions focused on machine learning often use more complex models (e.g., random forest or neural network) and optimize them extensively to optimize the predictive accuracy. Papers also develop custom architectures. For example, Zhao et al. (2019) propose a network architecture to incorporate financial variables and unstructured data. The **data split** (not in **Error! Reference source not found.** due to space limitations) indicates the extent to which a model's forecasting capability has been verified. Fundamentally, both traditional statistical approaches (hypothesis testing and regression), often assessed in-sample, as well as insights stemming from machine learning, which often uses cross-validation, make important contributions. The identified data splits range from classical regression without a split (15) (Bao and Datta 2012; Hu et al. 2018) over the two-way split (10), i.e., train-test split, (Altman et al. 2010) to cross-validation (2), e.g., (Choi et al. 2020) and not available (2). Due to the relatively large number of publications without a data split, statements regarding the out-of-sample performance need to be interpreted cautiously.

Regarding the central **findings** of the identified papers, the importance of the sentiment extracted from financial news for CRA could be shown quite consistently in varying research setups (Janner and Schmidt 2015; Liebmann et al. 2016; Lu et al. 2012; Norden 2017). The same applies to the volume of news (Tsai et al. 2016). The results are more ambiguous for automated topic extraction since the topic models and assigned topic labels vary from study to study. (Bao and Datta 2014; Hajek et al.

2017). For example, the topics "restructuring" and "investment policy" are found to be important in one study (Hajek et al. 2017) and "macroeconomic risk" and "funding risk" in another (Bao and Datta 2014). Fernandes and Artes (2016) and González-Fernández and González-Velasco (2019) represent more unusual approaches. The former uses spatial data to improve the CRA; the latter shows that search engine activity correlates with a credit risk measure. All in all, most papers identify an added value of the variables obtained from alternative data, providing explanatory value or improving the predictive power. Concerning design knowledge, few papers are fully situated in the design science research (DSR) paradigm. Design requirements or principles remain rather implicit. Hristova et al. (2017) propose the *RatingBot* and a process to extract a credit rating from text but the requirements arising from the problem domain could be formulated more explicitly. Zhao et al. (2019) propose a default prediction framework. However, the abstraction towards design principles or even theories is absent. Hence, a lot of untapped potential for research in the DSR paradigm is apparent.

Table 3: Publications identified in the literature review
(If not specified "Entities" includes banks and non-banks; * signals borderline relevance)

Research paper	Entities Granul.	Alt. data source	Risk measure (dep. variable)	Statistical Model	Main results
(Aktug et al. 2015)	PU E	Human resource	Bond spread	Event study, hypothesis test	Credit analyst hiring has an impact on bond but not stock return
(Altman et al. 2010)	PR NB Q	Other	Default category	Log. reg.	Alternative data on legal actions by creditors helps to increase predictive power
(Bao and Datta 2012)*	PU Y	Annual reports	Unsupervised	Not applicable	Proposed Sent-LDA improves identification of risk types
(Bao and Datta 2014)*	PU Y	Annual reports	Equity volatility	Panel reg.	With Sent-LDA, 2/3 risk types not relevant, three show risk increase and five decrease
(Cecchini et al. 2010)*	PU Y	Annual reports	Default category	SVM	Fin. variables with complementary qualitative data achieves the best result
(Choi et al. 2020)*	PU NB Y	Annual reports	Credit rating	SVM, NN, RF	Combination of fin. variables and qualitative data achieves the best prediction

(Fernandes and Artes 2016)	PR NB NA	Spatial data	Default category	Log. reg.	The inclusion of a spatial risk factor improved bankruptcy identification
(González-Fernández and González-Velasco 2019)	PU BA W	Search engine	CDS spread	Linear reg.	Inclusion of search engine-based sentiment index improved credit risk prediction
(Gül et al. 2018)	PU NA	Social media	Credit rating	Multiple criteria decision making	Social media was found to be useful in CRA for half of the companies
(Hajek et al. 2017)*	PU Y	Annual reports	Credit rating	Log. reg., decision tree, NBN, RF, …	Risky firms mention restructuring less and domestic market difficulties more frequently
(Hristova et al. 2017)*	PU Y	Annual reports	Credit rating	Log reg., naïve bayes, SVM, …	Classification model and text representation are important determinants of accuracy
(Hu et al. 2018)*	PU Y	Annual reports	CDS spread	Panel reg.	Evidence for an inverse relationship between readability and CDS spreads
(Janner and Schmidt 2015)	PU E	Financial news	Bond spread	Event study hypothesis test	Explanatory power of corp. news for bonds is comparable to power for stock market
(Liebmann et al. 2016)	PU D	Financial news	CDS spread	Panel reg.	Based on sentiment, CDS traders and equity traders interpret the same news differently
(Lu et al. 2012)	PU NB Q	Financial news	Credit rating	Mod. probit, SVM	News coverage is significantly associated with credit downgrades
(Lu et al. 2015)	PU NB Q	Financial news	Default category	Log. reg.	Distress indicator derived from financial news possesses significant explanatory power
(Mengelkamp et al. 2015)	PU,PR NB M	Social media	Default category	Hypothesis test, k-nearest-neighbor	More social media posts and worse sentiment for financially instable companies; classification accuracy above 50%.
(Mengelkamp et al. 2016)	PU,PR NB NA	Social media	Default category	Frequency counts	Sentiment dictionary achieves 67.9% accuracy compared to 49.97% by domain-independent dictionary

(Mengelkamp et al. 2017)*	PU,PR NB NA	Social media	Default category (proxy)	Log. reg., dec. tree, SVM, …	SVM and bag-of-words show the best performance for Tweet classification
(Norden 2017)	PU D	Financial news	CDS spread	Panel reg.	Financial news show significant influence on the way CDS spreads change
(Onay and Öztürk 2018)	PU NA	Social media	Not applicable	Not applicable	Review shows rising relevance of non-traditional data sources for credit scoring
(Safi and Lin 2014)	PR NB NA	Commerce platform	Solvency proxy	Log. reg.	Measures from commerce platform (membership period, page views) help to explain creditworthiness
(Smales 2016)	PU BA D	Financial news	CDS spread, LIBOR spread	Panel reg.	Significant negative relationship between news sentiment and CDS spread changes
(Tsai et al. 2010)	PU Y	Financial news	Credit rating	Ordered logit/ probit model	Sentiment analysis of corporate news shows explanatory contribution for credit rating
(Tsai et al. 2016)	PU NB Q	Financial news, Ann. reports	CDS spread	Panel reg.	High news volume and negative sentiment are associated with an increase in credit risk
(Yan et al. 2019)	PU NA	Financial news	Entity association	Uni- and bidirect. GRU	Modeling of relation between firms using neural network improves classification
(Yang et al. 2020)	PU D	Financial new	CDS spread	Panel reg.	Inverse relationship of news sentiment and CDS spread; more pronounced in cases of higher analysts' earnings forecasts dispersion
(Yuan et al. 2018)	PU Q	Social media	Credit rating	Log reg., RF, NN, SVM	Topic model that incorporates emotion detection achieves improved accuracy
(Zhao et al. 2019)	PU W	Financial news	Default category	GAM, NN	Combination of financial measures and social media data improves accuracy

Table 4: Definition of abbreviations

Entities	PU (public), PR (private), BA (bank), NB (non-bank)
Granularity	E (Events), D (Daily); W (Weekly), M (Monthly), Q (Quarterly), Y (Yearly), NA (not available)
Stat. model	Logistic regression (log. reg.), panel regression (panel reg.), support vector machine (SVM), neural network (NN), random forest (RF), naïve Bayesian network (NBN), gated recurrent unit (GRU), generalized additive model (GAM)

4.2   Research Agenda

The following **research agenda for future research** on **using alternative data for CrRM** is derived from the analysis. Research gap (*RG*) *1. research on non-public companies.* There is still a lack of research that focuses on small and medium-sized non-public companies since there is less information (e.g., market-based metrics) available. *RG 2. need for research using irregularly occurring/frequent data sources.* There is still untapped potential to analyze signals from alternative data with varying frequency to support decision-making in CrRM. *RG 3. research beyond quarterly frequency.* For the used methods and findings, more research on how alternative data can support CrRM is needed, ranging from the monthly to intraday level. *RG 4. convergence between econometrics and machine learning-based studies regarding methods and evaluation.* There is a distinct divergence between the fields, which calls for more interdisciplinary research to allow rigorous evaluations and more comparable results. *RG 5. Research on design knowledge.* Due to the rather implicit use of DSR in many cases, further research is needed to create expand knowledge related to the design of risk management systems utilizing alternative data (principles and theories).

5   Discussion and Conclusion

In terms of *practical implications*, the studies suggest that alternative data can contain decision-relevant signals. The question arises how the CrRM process and skill profiles (data integration/analysis) for risk managers may evolve. Will the in-house capabilities for CRA need to increase, or will such insights be procured from external entities? This paper *contributes* to *research* by 1) identifying and classifying research addressing the intersection of CrRM and alternative data and 2) deriving a research agenda that encompasses the most prominent research gaps. Like any research, this study has potential *limitations*. Literature reviews are affected by publication bias.

Additionally, the search could fail to identify relevant literature because related terms were not considered. Also, the alternative data that studies use was mostly in English. Moreover, only published research could be considered, not proprietary models used by banks. To answer RQ1, the literature review classified 29 papers in which alternative data provides a basis for enhanced CrRM. The proposed research agenda consists of the most prominent research gaps, thereby addressing RQ2. Overall, the evidence suggests that alternative data can improve CrRM in terms of a better understanding of the risk situation and the predictive performance.

References

Aktug, R. E., Nayar, N., and Salas, J. M. 2015. "Are Credit Rating Agency Analysts Valuable?," The Journal of Risk Finance (16:4), pp. 378-394.

alternativedata.org. 2020. Retrieved 02/25/2021, from www.alternativedata.org.

Altman, E. I. 1968. "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," The Journal of Finance (23:4), pp. 589-609.

Altman, E. I., Sabato, G., and Wilson, N. 2010. "The Value of Non-Financial Information in Small and Medium-Sized Enterprise Risk Management," Journal of Credit Risk (6:2), pp. 1-33.

Altman, E. I., and Saunders, A. 1997. "Credit Risk Measurement: Developments over the Last 20 Years," Journal of Banking & Finance (21:11-12), pp. 1721-1742.

Bao, Y., and Datta, A. 2012. "Summarization of Corporate Risk Factor Disclosure through Topic Modeling," Proceedings of the 33rd International Conference on Information Systems, Orlando, FL.

Bao, Y., and Datta, A. 2014. "Simultaneously Discovering and Quantifying Risk Types from Textual Risk Disclosures," Management Science (60:6), pp. 1371-1391.

Caouette, J. B., Altman, E. I., Narayanan, P., and Nimmo, R. 2008. Managing Credit Risk: The Next Great Financial Challenge. Hoboken, New Jersey: John Wiley & Sons, Inc.

Cecchini, M., Aytug, H., Koehler, G. J., and Pathak, P. 2010. "Making Words Work: Using Financial Text as a Predictor of Financial Events," Decision Support Systems (50:1), pp. 164-175.

Chen, N., Ribeiro, B., and Chen, A. 2016. "Financial Credit Risk Assessment: A Recent Review," Artificial Intelligence Review (45:1), pp. 1-23.

Choi, J., Suh, Y., and Jung, N. 2020. "Predicting Corporate Credit Rating Based on Qualitative Information of Md&a Transformed Using Document Vectorization Techniques," Data Technologies and Applications (54:2), pp. 151-168.

Das, S. R., Hanouna, P., and Sarin, A. 2009. "Accounting-Based Versus Market-Based Cross-Sectional Models of Cds Spreads," Journal of Banking & Finance (33:4), pp. 719-730.

Fernandes, G. B., and Artes, R. 2016. "Spatial Dependence in Credit Risk and Its Improvement in Credit Scoring," European Journal of Operational Research (249:2), pp. 517-524.

González-Fernández, M., and González-Velasco, C. 2019. "An Alternative Approach to Predicting Bank Credit Risk in Europe with Google Data," Finance Research Letters (35).

Gül, S., Kabak, Ö., and Topcu, I. 2018. "A Multiple Criteria Credit Rating Approach Utilizing Social Media Data," Data & Knowledge Engineering (116), pp. 80-99.

Hajek, P., Olej, V., and Prochazka, O. 2017. "Predicting Corporate Credit Ratings Using Content Analysis of Annual Reports – a Naïve Bayesian Network Approach," FinanceCom 2016, Frankfurt, Germany: Springer International Publishing, pp. 47-61.

Hristova, D., Probst, J., and Eckrich, E. 2017. "Ratingbot: A Text Mining Based Rating Approach," Proceedings of the 38th Int. Conference on Information Systems, Seoul, South Korea, pp. 1-19.

Hu, N., Liu, L., and Zhu, L. 2018. "Credit Default Swap Spreads and Annual Report Readability," Review of Quantitative Finance and Accounting (50:2), pp. 591-621.

Janner, S., and Schmidt, D. 2015. "Are Economically Significant Bond Returns Explained by Corporate News? An Examination of the German Corporate Bond Market," Financial Markets and Portfolio Management (29:3), pp. 271-298.

Koulafetis, P. 2017. Modern Credit Risk Management: Theory and Practice. London: Palgrave Macmillan.

Lam, J. 2014. Enterprise Risk Management: From Incentives to Controls. Hoboken, New Jersey: John Wiley & Sons, Inc.

Liebmann, M., Orlov, A. G., and Neumann, D. 2016. "The Tone of Financial News and the Perceptions of Stock and Cds Traders," Int. Review of Financial Analysis (46), pp. 159-175.

Loughran, T., and McDonald, B. 2016. "Textual Analysis in Accounting and Finance: A Survey," Journal of Accounting Research (54:4), pp. 1187-1230.

Lu, H.-M., Tsai, F.-T., Chen, H., Hung, M.-W., and Li, S.-H. 2012. "Credit Rating Change Modeling Using News and Financial Ratios," ACM Trans. on Management Inf. Systems (3:3), pp. 1-30.

Lu, Y.-C., Wei, Y.-C., and Chang, T.-Y. 2015. "The Effects and Applicability of Financial Media Reports on Corporate Default Ratings," Int. Review of Economics & Finance (36), pp. 69-87.

McNeil, A. J., Frey, R., and Embrechts, P. 2015. Quantitative Risk Management: Concepts, Techniques and Tools - Revised Edition, (2 ed.). Princeton, New Jersey: Princeton University Press.

Mengelkamp, A., Hobert, S., and Schumann, M. 2015. "Corporate Credit Risk Analysis Utilizing Textual User Generated Content - a Twitter Based Feasibility Study," Proceedings of the 19th Pacific Asia Conference on Information Systems, Singapore.

Mengelkamp, A., Rojahn, M., and Schumann, M. 2017. "Evaluating Machine Learning Algorithms for Sentiment Classification of Tweets in Credit Assessment," ECSM 2017 4th European Conference on Social Media, Vilnius, Lithuania, pp. 209-216.

Mengelkamp, A., Wolf, S., and Schumann, M. 2016. "Data Driven Creation of Sentiment Dictionaries for Corporate Credit Risk Analysis," Proceedings of the 22nd Americas Conference on Information Systems, San Diego, CA.

Monk, A., Prins, M., and Rook, D. 2019. "Rethinking Alternative Data in Institutional Investment," The Journal of Financial Data Science (Winter 2019), pp. 14-31.

Nassirtoussi, A. K., Aghabozorgi, S., Ying Wah, T., and Ngo, D. C. L. 2014. "Text Mining for Market Prediction: A Systematic Review," Expert Systems with Applications (41:16), pp. 7653-7670.

Norden, L. 2017. "Information in Cds Spreads," Journal of Banking & Finance (75), pp. 118-135.

Onay, C., and Öztürk, E. 2018. "A Review of Credit Scoring Research in the Age of Big Data," Journal of Financial Regulation and Compliance (26:3), pp. 382-405.

Roeder, J., Muntermann, J., and Kneib, T. 2020. "Towards a Taxonomy of Data Heterogeneity," Proceedings of the 15th International Conference on Wirtschaftsinformatik, Potsdam, Germany.

Safi, R., and Lin, Z. 2014. "Using Non-Financial Data to Assess the Creditworthiness of Businesses in Online Trade," Proceedings of the 18th Pacific Asia Conf. on Information Systems, Chengdu, China.

Schumann, M. 2002. "Bonitätsbeurteilungen Und Kreditprüfungen Im Electronic Business," in Electronic Business: Theoretische Aspekte Und Anwendungen in Der Betrieblichen Praxis, R. Gabriel and U. Hoppe (eds.). Heidelberg: Physica-Verlag HD, pp. 413-427.

Smales, L. A. 2016. "News Sentiment and Bank Credit Risk," Journal of Empirical Fin. (38), pp. 37-61.

Tsai, F.-T., Lu, H.-M., and Hung, M.-W. 2010. "The Effects of News Sentiment and Coverage on Credit Rating Analysis," Proceedings of the 14th Pacific Asia Conf. on Inf. Systems, Taipei, Taiwan.

Tsai, F.-T., Lu, H.-M., and Hung, M.-W. 2016. "The Impact of News Articles and Corporate Disclosure on Credit Risk Valuation," Journal of Banking & Finance (68), pp. 100-116.

Van Gestel, T., and Baesens, B. 2008. Credit Risk Management: Basic Concepts: Financial Risk Components, Rating Analysis, Models, Economic and Regulatory Capital. Oxford, United Kingdom: Oxford University Press.

vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R., and Cleven, A. 2009. "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process," Proceedings of the 17th European Conference on Information Systems, Verona, Italy.

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," MIS Quarterly (26:2), pp. xiii-xxiii.

Xing, F. Z., Cambria, E., and Welsch, R. E. 2017. "Natural Language Based Financial Forecasting: A Survey," Artificial Intelligence Review (50:1), pp. 49-73.

Yan, C., Fu, X., Wu, W., Lu, S., and Wu, J. 2019. "Neural Network Based Relation Extraction of Enterprises in Credit Risk Management," 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), Kyoto, Japan, pp. 1-6.

Yang, S., Liu, Z., and Wang, X. 2020. "News Sentiment, Credit Spreads, and Information Asymmetry," North American Journal of Economics & Finance (52).

Yuan, H., Lau, R. Y. K., Wong, M. C. S., and Li, C. 2018. "Mining Emotions of the Public from Social Media for Enhancing Corporate Credit Rating," 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), Xi'an, China, pp. 25-30.

Zhao, Y., Shen, Y., and Huang, Y. 2019. "Dmdp: A Dynamic Multi-Source Default Probability Prediction Framework," Data Science and Engineering (4:1), pp. 3-13.

HEALTH LITERATURE HYBRID AI FOR HEALTH IMPROVEMENT; A DESIGN ANALYSIS FOR DIABETES & HYPERTENSION

LUUK PA SIMONS, MARK A NEERINCX &
CATHOLIJN M JONKER

Delft University of Technology, Faculty of Computer Science, Delft, Netherlands;
e-mail: l.p.a.simons@tudelft.nl, M.A.Neerincx@tudelft.nl, C.M.Jonker@tudelft.nl

Abstract Increasingly, front runner patients and practitioners want to use state-of-the-art science for rapid lifestyle based cure of diseases of affluence. However, the number of new health studies per year (>500.000) is overwhelming. How to quickly assess state-of-the-art and use new opportunities for rapid patient DIY (Do-It-Yourself) health improvement? In order to develop a health literature hybrid AI to aid DIY rapid health improvement, we analyze user side functional requirements. A cross case design analysis is conducted for hypertension and T2D (Type 2 Diabetes), two major cardiometabolic conditions in our society. Our analysis shows that current DIY health support is 'watered down' advise, prone to medicalizing rather than empowering patients. We propose hybrid AI user requirements and discuss how a 2030 hybrid AI health support system can stimulate new ways of working in health and cure.

1   Introduction

When it comes to lifestyle related diseases like cardiovascular disease, Type 2 Diabetes (T2D), dementia and colorectal-, prostate- and other forms of cancer, it turns out that key to our health is our biological self-repair. In virtually all our cells and tissues, damage is being repaired on a continuous basis (Li, 2019). This fact is still largely underutilized by patients and by healthcare professionals. Nor are we using the options to rapidly improve self-repair effectiveness (with biometric progress feedback on a daily basis) from healthy lifestyle choices on foods, exercise, sleep etc (Greger & Stone, 2016).

Already in 2009 Safeway CEO and the corporate Coalition to Advance Healthcare Reform have calculated that 74% of health costs come from only four conditions (cardiovascular disease, type 2 diabetes, obesity and cancer) which are largely preventable or reversible (Burd, 2009). The Lancet EAT committee reiterated this urgency to use options for prevention and reversal of disease more effectively: we cannot afford our current approach, not in health nor in ecology (Willett, 2019).

As discussed elsewhere, health improvement options are welcomed by many (though not all) patients around the moment of diagnosis (Simons, 2020a). There are groups of front runner patients and practitioners who want to use state-of-the-art science for rapid lifestyle based cure of diseases of affluence. Moreover, research increasingly shows that from a biology perspective, health self-repair is more effective than current 'best available' medical treatments (largely because self-repair is biologically more plausible and more advanced, thanks to millions of years of evolution, Greger & Stone, 2016, Li, 2019). The number of well conducted RCT's (Randomized Controlled Trials) showing rapid health improvements within a matter of hours, days or weeks is rapidly growing, largely in the domains of cardio- and metabolic conditions, plus increasingly so in the onco- and neurology domains: depression and even dementia (Greger & Stone, 2016, Bredesen 2017, 2018, Ornish & Ornish 2019, Simons 2020a, 2021a, 2021b).

However, DIY health priorities are difficult to choose, since the number of yearly new studies on health is so large that the field can be overwhelming. For example, even when limiting the search to only the year 2019, Scholar Google finds >500.000 studies on 'health', of which >60.000 are on 'healthy lifestyle'. Furthermore, 2019

L. PA Simons, M. A Neerincx &C. M Jonker:
Health Literature Hybrid AI for Health Improvement; A Design Analysis for Diabetes & Hypertension

183

has >150.000 studies on 'obesity' and >180.000 studies on 'cardiovascular health'. In short, every working day of the year there are >2000 new studies on health: good luck keeping up with that! And whether you are a practitioner or a patient, you likely have tasks which preclude reading many hours of literature every day. Given this enormous amount of literature, it is also quite easy to get lost in sub-branches, while losing sight of the bigger picture.

In order to help practitioners and DIY patients to navigate this massive amount of science and help them capture, assess and use the best and most recent available evidence on lifestyle interventions for disease reversal, we aim to develop a health literature AI. Thus, the main ***research question*** is:

> What are user requirements for a health literature AI in order to support successful DIY healthy lifestyle choices for health self-repair?

2   Literature

From a biology and health engineering perspective, some of the most promising recent health discoveries use our innate mechanisms for rapid bodily self-repair (Li, 2019). We want to help people experience and measure improved health, possibly within days, with rapid feedback of progress from health measurements.

For design purposes, we take a '2030' view from the future, using 'optimism by method': assuming maximum use of the dynamic nature of our biology for self-repair and temporarily ignoring current healthcare barriers. Our aim is to promote cure via rapid health self-repair feedback cycles. This needs an approach with personal iteration cycles, see Figure 1, using (Cross, 1994, Simons, 2020a) goals analysis (problem space), intervention planning (solution space) and measurement portfolio (evaluation space).

We can translate this health iteration cycle into DIY health questions for the hypertension and T2D cases of this paper. DIY health questions for a patient (or a practitioner guiding him/her) may become:

1. What is the underlying biology of the condition (causes, outlook, risk factors)?
   (= Problem Space)

2. What are the most effective lifestyle interventions (& their attractiveness)?
   (= Solution Space)

3. What are suitable health tracking options (behaviors, symptoms, biometrics)?
   (= Evaluation Space)
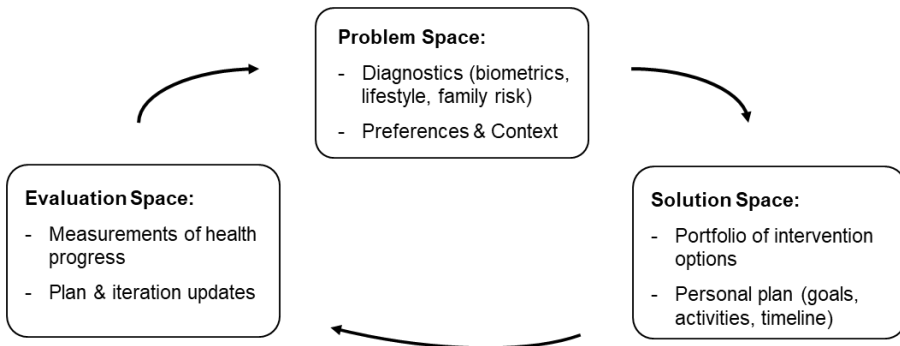
Figure 1: Personal iteration cycles for rapid health self-repair.[1]

Various forms of goal setting based on personal preferences and individual coaching, eTool use like microlearning for health, Quantified Self (QS) progress tracking and peer coaching have all been shown to aid success (Simons, 2015, 2016, 2020b). Generally, it is important that patients can set their own priorities and plans, while also using practitioner support (Simons, 2014). This is not only important from a personal perspective (motivation plus a suitable fit with personal preferences and context), but it is also important from the perspective of science and up to date evidence. Lifestyle advice for patients is often outdated, due to slow adoption in health care. On average, new findings take about two decades before they enter standard clinical practice (Balas & Boren, 2000). The practical implication is that for front runner **patients** with ambition in DIY health improvement, science has a lot more effectiveness and evidence to offer than is visible in regular patient lifestyle guidelines. The same challenge exists for **practitioners** wanting to support DIY

---

[1] Besides biology opportunities of self-repair, overall health iteration success depends on the full picture of choosing personal goals and behaviours that are best suited for one's preferences and context. See Simons (2010, 2013, 2014) for information on intervention planning.

L. PA Simons, M. A Neerincx &C. M Jonker:
Health Literature Hybrid AI for Health Improvement; A Design Analysis for Diabetes & Hypertension

185

health for their patients. Below, we describe how we use a cross-case analysis to find user requirements that must be fulfilled in order for the health literature AI system to aid DIY health intervention choices.

3 Method

Our research question is a design analysis question. The analysis is an example of design research rather than design science (Vaishnavi & Kuechler, 2004), since design research aims at generating (domain specific) knowledge for solving a given problem. Our analysis will follow design cycle phases 1 and 2 of (Verschuren & Hartog, 2005): '1. first hunch' and '2. assumptions and requirements'. Our first hunch is that we need to explicate the gaps in common sources of information for DIY patients (health care lifestyle guidelines and Google Scholar[2]). In other words: which needs or gaps should be filled with the health literature AI to aid DIY health intervention choices? Second, can we formulate 'Voice of the Patient' user requirements? We use the first step from QFD (Quality Function Deployment) for software design. This means we explicate 'the voice of the user/patient', using words that users might use themselves (Simons & Verhagen, 2008, Schockert & Herzwurm, 2018), to indicate their needs when using the AI system. (Next, outside the scope of this paper, come steps to validate this with user testing and to form a QFD matrix translating user requirements to technology attributes.)

Since we want a domain-independent structure of the AI health literature support system, we use two different health domains for our DIY case analyses: hypertension and T2D (Type 2 Diabetes). We see them as suitable cases, since they are relevant (with these conditions impacting respectively 50% and 30% of people in affluent countries), different (managed and researched by different specialists) and obviously lifestyle related. We analyze the Dutch situation: What are some of the main health care lifestyle sites and guidelines that patients encounter? What do we observe if we compare that to leading edge lifestyle interventions?

Our approach is similar to action research in the sense that we have a high level of 'access' to the current practices and patients in these domains,[3] while at the same

---

[2] We take Scholar Google as a reference point for exploring recent studies, since it is so widely used.
[3] By providing 6 months of lifestyle coaching (Simons et al., 2010, 2017) for literally thousands of patients and caregivers in these domains, over the course of the past 10 years.

time trying to help them in navigating the information diversity they encounter. Many 'front runner DIY patients' are not average. Although they are higher educated on average, we see their struggles on a daily basis in trying to digest and use the available health science for their DIY health choices. Simultaneously, we see potential for AI to help them. The user analysis in this paper is meant as a first iteration for 'user requirements' that would support their search and decision needs. A fruitful way to start, is to evaluate the current routes/tools they use and analyze the user needs that become apparent from that process.

In the analysis section below, we will take the following steps for our case and user needs analysis (for T2D & hypertension), in two main paragraphs:

1. (a) Case analysis *Health care advise*: What are some of the main health care lifestyle sites and guidelines that patients encounter for their condition?

1. (b) Evaluation from the design goal perspective: What *omissions* do we see if we compare results from step 1(a) to *leading lifestyle intervention science*?

2. (a) Case analysis *Science*, via Google Scholar: What is the content, diversity, clarity and applicability of the information found?

2. (b) *Translation to 'voice of the patient' user requirements*: How could the AI system support my needs and decisions?

4   Analysis, cases T2D & hypertension

4.1   Health care lifestyle guidelines vs. DIY health decisions

As an exemplary search route for a DIY patient with T2D in the Netherlands, we started with a google search (in Dutch) with: "I have diabetes, what can I do?" This led to a top 3 of respectable online sources: www.thuisarts.nl (most visited NL site for family doctor questions), www.diabetesfonds.nl (NL diabetes research & funding) and www.dvn.nl ('Diabetes Vereniging NL' patient association).

L. PA Simons, M. A Neerincx &C. M Jonker:
Health Literature Hybrid AI for Health Improvement; A Design Analysis for Diabetes & Hypertension

187

Table 1: Case analyses: What is advised vs. omitted on traditional health care sites?

	T2D (Type 2 Diabetes)	**Hypertension**
Advised (a)	-Lower your blood sugar by eating well (fruits, veggies, nuts, yogurt. No sugary drinks) brisk walk 30 min/day or 60 min/day if overweight. -If that doesn't work: pills. -Manage it well: 3-monthly medical checks.	-Stop smoking, eat well (fruits, veggies, wholegrain, fibers, less saturated fat), less salt, brisk walk 2,5 hrs/week, less stress. -Other factors: weight, alcohol, fatty foods (& some meds) -If cardiac risks: pills. -Discuss checkups with doctor.
Omitted (b)[4]	-T2D is >90% avoidable with healthy lifestyle. -Interventions exist that remove >75% of meds in 4 weeks. -Causes: insulin resistance, lipotoxicity, inflammation: 1-wk reset interventions very effective.	-Hypertension >90% avoidable with healthy lifestyle. -Interventions exist that remove >50% of meds in 4 weeks. -Causes: endothelial function & inflammation: food has more & faster effect than medication.

Apart from the similarities, also summarized in Table 1, it is interesting to see that www.thuisarts.nl is more directed towards medication and 3-monthly checks for complications. Whereas the other two sources explain the causal roles of health behaviors and insulin sensitivity better.

Figure 2: Food page of www.dvn.nl directly contradicts www.dvn.nl advise.

---

[4] Sources from longstanding research lines: overall (Roberts & Barnard, 2005), in T2D (Hu, 2001, Fuhrman & Sorensen, 2012, Simons, 2016, 2021a) in hypertension, endothelial health and inflammation (Niskanen, 2004, Franzini, 2012, Rodriguez-Leyva, 2013, Dickinson, 2014, Kapil, 2015, Siervo, 2015, Greger & Stone, 2016).

A similar search for **hypertension** gave as top 3 sources: www.thuisarts.nl again, www.hartstichting.nl (cardiac research funding & patient education) and www.zorgkaartnederland.nl (patient association to compare care providers). Of these, www.hartstichting.nl gives most lifestyle support, but not anywhere close to scientific state-of-the-art.

Three aspects are fascinating about these sources: (1) all the relevant and evidence based health **facts they do not give**, **see 'Omitted' in Table 1** summary (2) tendencies to medicalize instead of empower patients (3) the **contradictions and biases** that persist from Dutch food culture. As two examples of bias, all three T2D sources are clear that saturated fats make things worse. Which they give as one of the reasons that meats should be avoided. Still, Figure 2 shows what the very first picture is on the www.dvn.nl healthy foods page: a meat based dish. And we all know that 1 picture speaks louder than 1000 words… A second example of Dutch food bias is cheese. Despite its high saturated fat content, all three T2D sites say that cheese is perfectly healthy for T2D patients, without providing any justification. The cheese advice is biologically implausible and it contrasts with large empirical studies (Guasch-Ferré, 2017, Drouin-Chartier, 2019) showing clear T2D risk reductions when replacing cheese and butter with less harmful foods[5].

4.2   Scientific studies vs. DIY health decisions

As illustrated in section 4.1, healthy lifestyle advice on main patient support sites is watered down and prone to cultural and historical biases. In other words: outdated and not suited to patients or practitioners that prefer high impact interventions. Hence, the question is: what if we go directly to the scientific state-of-the-art, how easily will we find clear and actionable answers? Though one could argue that scientific studies are not useful since they are not written for DIY health questions, one could also argue the opposite: when looking for the latest findings and evidence, what better place to look than science? The AI for DIY health we aim for, is meant to bridge both sides of this equation.

---

[5] Outside our scope, there are ample discussions (Campbell & Campbell, 2016, Fuhrman & Sorensen, 2012, Greger & Stone, 2016, Greger, 2019) of how our health institutions are living in bubbles of 'not rocking the boat', leading to culturally biased and watered down advice. Which is quite different from the high impact interventions that leading edge DIY patients and practitioners are looking for.
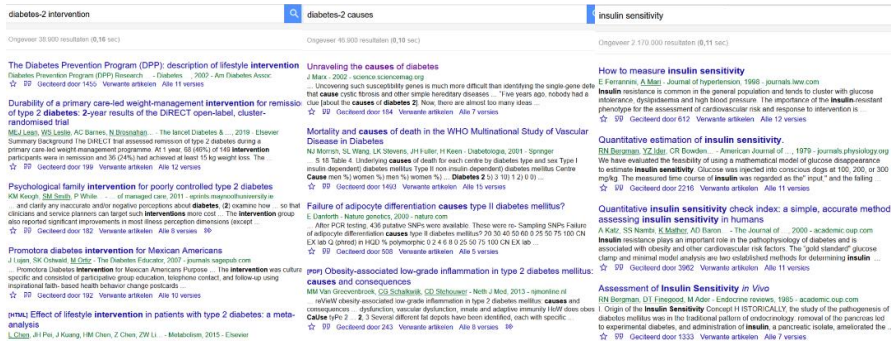
Figure 3: Illustration of study diversity when searching for DIY health answers.

One sees when using Google Scholar, see Figure 3, that the body of scientific studies is not only large, but also highly diverse, with many different subdisciplines in science having their own language and focus. For example, the search results for measuring insulin sensitivity (or -resistance) are way too diverse and technical for helping a patient with his/her daily or weekly progress tracking question. A simple 'ask your doctor to measure it via an OGTT (Oral Glucose Tolerance Test)' would be more helpful. In Table 2 we summarize our main Scholar search findings with regard to the section 2 patient questions: causes, interventions and measurements.

Table 2: Use cases science: study overview & contents evaluation

	T2D (Type 2 Diabetes)	**Hypertension**
Study Search Content	-**Causes**: diverse papers, many with a genetics, cell or pharma focus, or on complications (cardiac, renal, retina etc). Different results per population. Psycho-socio-cultural factors. -**Interventions:** widely varying results & difficult to assess why. Reviews= 'average' results, not highest impact. -**Measuring:** either 'medi-tech' details or quarterly checks & sugar management or 'modest' QS for walking, weight loss etc.	-**Causes**: many forms (resistant, pulmonary, nondipping, secondary) correlates & co-morbidities of hypertension. -**Interventions:** apart from many drugs intervention also a long lifestyle interventions tradition. Hard to find and compare dose-response for components: salt, meat, smoke, sports, stress, alcohol, fruits, veggies, fiber etc. -**Measuring:** Many on 24-h ambulatory monitoring. 'Manage it' is checkups (& often drugs).

Another finding is that Google Scholar search results aid in exploring the field, but they are not qualified overviews, see also Figure 3. Overviews exist in the various academic subdisciplines, like literature reviews or meta-analyses, but they often match poorly to the more action-focused 'voice of the patient' questions we hear on a daily basis. These are questions on e.g. feasibility of interventions, what is most useful to do and to measure and how to deal with dilemma's and tradeoffs?

In **answer to our Research Question** and including the concerns above, we get as draft **'voice of the patient' user requirements** for the AI system:

1. What are the main causal lifestyle factors that I can potentially influence?
1.1. How large are the effects per causal factor?
1.2. What is the quality of evidence to support this?
2. What are the most effective lifestyle interventions?
2.1 Which are relatively easy and/or attractive for me?
2.2 Which offer rapid, noticeable health results?
3. How can I rapidly measure my health progress?
3.1 Which measurements are low cost & practical for DIY?
3.2 Which are reliable health progress indicators (=have good external validity)?
4. Which attributes above need tradeoff decisions?

If we then look at for example questions 2.1 (intervention ease and attractiveness) or 2.2 (rapid results), we find that most academic overviews are not outlined along these lines. The AI system will need to provide functionality to fill that void and help answer these questions for front runner DIY patients and practitioners.

5   Discussion & Conclusion: AI for next level Quantified Self

An important limitation to our study is that our results still need validation via user testing. Preferably via a Wizard of Oz type of study, with questions like: What would you like to know? Which searches would you use? How would result XYZ help you? What display of results would you like? And in terms of design process, the next QFD step has to be taken: translation of the user requirements to technical attributes which fulfill those requirements for the AI system.

L. PA Simons, M. A Neerincx &C. M Jonker:
Health Literature Hybrid AI for Health Improvement; A Design Analysis for Diabetes & Hypertension

191

Still, our analyses illustrate that standard lifestyle guidelines are rather meager for health DIY purposes (section 4.1) whereas the scientific information is huge and hard to assess, with many different 'bubbles' within the scientific community whose discussions are highly specialized and disjunct (section 4.2). When designing AI support, there are three reasons for a **hybrid** AI system (which includes expert mediated interpretations) rather than stand-alone AI. First, human interpretation of research design and study validity are needed to counter 'fabricated pseudo-science' lifestyle studies which are often industry-sponsored (Campbell & Campbell, 2016, Greger & Stone, 2016, Simons, 2020a). Second, to avoid 'newness bias'. For example the PCRM (Physicians Committee for Responsible Medicine) show how 'serious scientists' have abandoned studying cholesterol effects of eggs decades ago, since the results were so clear, leaving the field open to biased egg industry studies under labels like 'recent studies show ..' (Barnard, 2019). Third, due to all kinds of confounding factors, lifestyle intervention successes can be difficult to achieve, thus cluttering the scientific field with mediocre results. If 90% of attempts for a certain intervention were less successful, how do we interpret and present the 10% that were very successful? Although this 10% may not form a majority, they often do lead the way forward for new lifestyle successes.

Conclusion

Front runner patients and practitioners aiming for rapid DIY health improvements have a lot to offer for pioneering the frontiers of a more sustainable and effective '2030' healthcare. This will become even more powerful when they have a shared state-of-the-art health literature view thanks to the hybrid AI system we aim to develop. For diseases of affluence, if "health is what happens between doctors' visits", this is a cheaper, more effective way to deliver health care.

Aknowledgements

References

Balas, E. A., & Boren, S. A. (2000). Yearbook of medical informatics: managing clinical knowledge for health care improvement. Stuttgart, Germany.

Barnard, N. D., Long, M. B., Ferguson, J. M., Flores, R., & Kahleova, H. (2019). Industry Funding and Cholesterol Research: A Systematic Review. American Journal of Lifestyle Medicine, 1559827619892198.

Bredesen, D. (2017). The end of Alzheimer's: The first program to prevent and reverse cognitive decline. Penguin.

Bredesen, D. E., Sharlin, K., Jenkins, D., Okuno, M. et al (2018). Reversal of cognitive decline: 100 patients. J Alzheimers Dis Parkinsonism, 8(450), 2161-0460.

Burd, Steven A. (2009), "How Safeway Is Cutting Health-Care Costs," Wall Street Journal, Eastern Edition, June 12, A15.

Campbell, T. C., & Campbell II, T. M. (2016). The China Study: Revised and Expanded Edition: The Most Comprehensive Study of Nutrition Ever Conducted and the Startling Implications for Diet, Weight Loss, and Long-Term Health. BenBella Books, Inc.

Cross, N. (1994). Engineering Design Methods; Strategies for Product Design (2nd ed. ed.). Chichester: John Wiley & Sons.

Dickinson, K. M., Clifton, P. M., Burrell, L. M., Barrett, P. H. R., & Keogh, J. B. (2014). Postprandial effects of a high salt meal on serum sodium, arterial stiffness, markers of nitric oxide production and markers of endothelial function. Atherosclerosis, 232(1), 211-216.

Drouin-Chartier, J. P., Li, Y., Ardisson Korat, A. V., Ding, M., Lamarche, B., Manson, J. E., ... & Hu, F. B. (2019). Changes in dairy product consumption and risk of type 2 diabetes: results from 3 large prospective cohorts of US men and women. Am J Clin Nutr, 110(5), 1201-1212.

Franzini, L., Ardigo, D., Valtuena, S., Pellegrini, N., Del Rio, D., Bianchi, M. A., ... & Zavaroni, I. (2012). Food selection based on high total antioxidant capacity improves endothelial function in a low cardiovascular risk population. Nutrition, Metabolism and Cardiovascular Diseases, 22(1), 50-57.

Fuhrman, J., & Sorensen, C. (2012). The End of Diabetes. HarperCollins.

Greger, M., & Stone, G. (2016). How not to die: discover the foods scientifically proven to prevent and reverse disease. Pan Macmillan.

Greger, M. (2019) How Not to Diet: The Groundbreaking Science of Healthy, Permanent Weight Loss. Flatiron Books.

Guasch-Ferré, M., Becerra-Tomas, N., Ruiz-Canela, M., Corella, D., Schröder, H., Estruch, R., ... & Salas-Salvadó, J. (2017). Total and subtypes of dietary fat intake and risk of type 2 diabetes mellitus in the Prevención con Dieta Mediterránea (PREDIMED) study. Am J Clin Nutr, 105(3), 723-735.

Schockert, S., & Herzwurm, G. (2018). Agile Software Quality Function Deployment. Software Engineering und Software Management 2018.

Hu, F. B., Manson, J. E., Stampfer, M. J., Colditz, G., Liu, S., Solomon, C. G., & Willett, W. C. (2001). Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. New England journal of medicine, 345(11), 790-797.

Kapil, V., Khambata, R. S., Robertson, A., Caulfield, M. J., & Ahluwalia, A. (2015). Dietary nitrate provides sustained blood pressure lowering in hypertensive patients: a randomized, phase 2, double-blind, placebo-controlled study. Hypertension, 65(2), 320-327.

Li, W. (2019). Eat to Beat Disease: The Body's Five Defence Systems and the Foods that Could Save Your Life. Random House.

Niskanen, L., Laaksonen, D. E., Nyyssönen, K., Punnonen, K., Valkonen, V. P., Fuentes, R., ... & Salonen, J. T. (2004). Inflammation, abdominal obesity, and smoking as predictors of hypertension. Hypertension, 44(6), 859-865.

Ornish, D., & Ornish, A. (2019). Undo it!: How simple lifestyle changes can reverse most chronic diseases. Ballantine Books.

L. PA Simons, M. A Neerincx &C. M Jonker:
Health Literature Hybrid AI for Health Improvement; A Design Analysis for Diabetes & Hypertension

193

Roberts, C. K., & Barnard, R. J. (2005). Effects of exercise and diet on chronic disease. Journal of applied physiology, 98(1), 3-30.

Rodriguez-Leyva, D., Weighell, W., Edel, A. L., LaVallee, R., Dibrov, E., Pinneker, R., ... & Pierce, G. N. (2013). Potent antihypertensive action of dietary flaxseed in hypertensive patients. Hypertension, 62(6), 1081-1089.

Siervo, M., Lara, J., Chowdhury, S., Ashor, A., Oggioni, C., & Mathers, J. C. (2015). Effects of the Dietary Approach to Stop Hypertension (DASH) diet on cardiovascular risk factors: a systematic review and meta-analysis. British Journal of Nutrition, 113(1), 1-15.

Simons, LPA & Verhagen, WP. (2008). Applying value-sensitive design and quality function deployment to healthcare ICT: the case of Dutch primary care unit dossiers. Journal of Design Research 7 (2): 155-176.

Simons, L. P. A., & Hampe, J. F. (2010). Service Experience Design for Healthy Living Support; Comparing an In-House with an eHealth Solution. The 23rd Bled eConference, pp. 423-440. Accessed 2010 from www.bledconference.org

Simons LPA, Hampe JF, Guldemond NA. (2013). Designing Healthy Living Support: Mobile applications added to hybrid (e)Coach Solution. Health and Technology. 3 (1) pp.85-95. DOI 10.1007/s12553-013-0052-9

Simons LPA, Hampe JF, Guldemond NA. (2014). ICT supported healthy lifestyle interventions: Design Lessons. Electronic Markets. 24 pp. 179-192. DOI 10.1007/s12525-014-0157-7.

Simons LPA, Foerster F., Bruck PA, Motiwalla L & Jonker CM. (2015). Microlearning mApp Raises Health Competence: Hybrid Service Design. Health and Technology, 5 pp 35-43. DOI 10.1007/s12553-015-0095-1

Simons, L. P., Pijl, H., Verhoef, J., Lamb, H. J. et al (2016). Intensive Lifestyle (e) Support to Reverse Diabetes-2. In Bled eConference (p. 24), accessed Dec 20, 2016 www.bledconference.org.

Simons LPA, Hafkamp MPJ, Bodegom D, Dumaij A, Jonker CM. (2017). Improving Employee Health; Lessons from an RCT. Int. J. Networking and Virtual Organisations, Vol. 17, No. 4, pp.341–353. DOI https://doi.org/10.1504/IJNVO.2017.088485

Simons, LPA, (2020a). Health 2050: Bioinformatics for Rapid Self-Repair; A Design Analysis for Future Quantified Self, pp. 247-261, 33rd Bled eConference. June 28-29, Bled, Slovenia, Proceedings retrieval from www.bledconference.org. ISBN-13: 978-961-286-485-9, DOI: https://doi.org/10.18690/978-961-286-485-9.17.

Simons, LPA, Heuvel, AC van den, Jonker CM. (2020b). eHealth WhatsApp for social support: design lessons. International Journal of Networking and Virtual Organisations, 23(2), 112-127. DOI https://doi.org/10.1504/IJNVO.2020.108857.

Simons, LPA, Pijl M, Verhoef J, Lamb HJ, van Ommen B, Gerritsen B, Bizino MB, Snel M, Feenstra R, Jonker CM. (2021a). e-Health Diabetes; 50 Weeks Evaluation. Int. J. Biomedical Engineering and Technology, Accepted (DOI will follow)

Simons, LPA, (2021b). Health 2050: Faster Cure via Bioinformatics & Quantified Self; A Design Analysis, Int. J. Biomedical Engineering and Technology, Submitted.

Vaishnavi, V., & Kuechler, W. (2004). Design Research in Information Systems. Accessed Aug 16, 2009 from http://desrist.org/design-research-in-information-systems

Verschuren, P., & Hartog, R. (2005). Evaluation in Design-Oriented Research. Quality and Quantity, 39, pp. 733-762.

Willett, W., Rockström, J., Loken, B., Springmann, et al (2019). Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. The Lancet, 393(10170), 447-492

Smile through the Mask: Emotion Measurement for Stationary Retail

Michael Meyer & Susanne Robra-Bissantz

Technische Universität Braunschweig, Chair of Information Management, Mühlenpfordtstr. 23, 38106 Braunschweig, Germany; e-mail: m.meyer@tu-braunschweig.de, s.robra-bissantz@tu-braunschweig.de

Abstract The global pandemic caused by the coronavirus disease (COVID-19) changes the lives of many people all over the world. In the context of stationary retail, a strong change of customer behavior occurs as mandatory safety measures like wearing facemasks and distance regulations have come into place. The sales personnel's ability to understand and react to customers' emotions is critical for service interactions and the customers' overall satisfaction. Unfortunately, facemasks make it difficult to recognize other's emotions and may lead to misinterpretation and confusion. To address this problem, this paper proposes the design of self-assessment interfaces that offer the customer an easy way to enter their emotions. As part of a Design Science Research (DSR) project, we designed three interfaces and evaluated them over the course of a design cycle. The results indicate that it is possible to use self-assessment technology in stationary retail to measure customer emotions.

1   Introduction

The global pandemic caused by the coronavirus disease (COVID-19) changes the lives of many people all over the world. Besides the physical health issues, COVID-19 evokes negative emotions like fear, sadness and anger (Aslam et al., 2020). In the context of stationary retail, a strong change of customer behavior occurs. Mandatory safety measures like wearing face masks and distance regulations have an influence on the customer (Yang et al., 2021). Stationary retail is a domain that particularly suffers from the grip of the pandemic, because it is already scarred by the digital transformation. The advancing digitalization ensures high sales in e-commerce, but simultaneously poses challenges to stationary retail (Doherty & Ellis-Chadwick, 2010). Since personal interaction with sales personnel is the main advantage of stationary retail, it is crucial to properly support the customer during this challenging time (Otto & Chung, 2000). Customer behavior cannot be explained solely by considering rational aspects (Kahneman, 2003) and is often affected by emotions (van Dolen et al., 2004). The sales personnel's ability to understand and react to the customers' emotions is critical for service interactions and customers' overall satisfaction (Bahadur et al., 2018). To protect oneself and others, the wearing of facemasks is recommended or mandatory, especially in indoor environments. Unfortunately, facemask make it difficult to recognize the emotions of others and may lead to misinterpretation and confusion (Carbon, 2020). To address this problem and support the interaction between customers and sales personnel, this paper proposes the design of self-assessment interfaces that provide a simple way for customers to enter their emotions. Our aim is to determine whether IT-supported self-assessment offers a suitable approach to measure customer emotions in stationary retail. Furthermore, the goal of this contribution is to generate design knowledge in order to provide digital support for stationary retail. Our research follows the Design Science Research (DSR) paradigm (Hevner et al., 2004) and in particular the General Methodology of Design Science Research (Vaishnavi et al., 2015). The research question of this paper is: How can emotion-self-assessment interfaces (ESAI) for stationary retail be designed? The paper is structured as follows: In section 2 and 3, we explain the importance of interaction between customers and sales personnel as well as the relevance of emotions. In section 4, the underlying methodology and the resulting design cycle are explained. Section 5 describes the design of three ESAI based on existing emotion theories, that were

later evaluated in a user test. In section 6, the results of the evaluation are presented. Finally, the papers ends with a conclusion.

2   Customer-Salesperson Interaction

The current difficulties in stationary retail are mainly caused by the digital transformation (Hagberg et al., 2016). Thereby, stationary retail is being threatened by the shift from physical to digital. In particular, e-commerce and the possibility of mobile shopping through smart devices are creating new challenges (Fulgoni, 2015; Reinartz et al., 2019). Although some retailers are able to retain customers through multi-channel strategies, so-called 'internet pure players' account for a large share of the trade (Keyes, 2018). The advantages stationary retail still has are the qualification of the sales personnel and the resulting customer services which can be offered (Otto & Chung, 2000), particularly the option of social interaction with the sales personnel (Gutek et al., 1999). This interaction creates a connection between the sales personnel and the customer within a common interaction space (Fyrberg & Jüriado, 2009), that may contribute to a mutual value creation (Grönroos, 2006). However, the mere existence of an interaction is not enough to ensure value creation (Fyrberg & Jüriado, 2009). An unfitting interaction between the sales personnel and the customer negatively impacts customer satisfaction and salesperson comfort (Groth & Grandey, 2012), whereas a successful interaction can lead to increased trust, stronger loyalty, and improved comfort in future interactions (Geiger et al., 2020b; Grönroos & Voima, 2013). The currently predominant COVID-19 pandemic intensifies the critical situation for stationary retail, not only because stores are oftentimes temporarily closed but because the sales personnel has difficulties to fully recognize the customer's emotions displayed by facial expressions due to facemasks (Adolphs, 2003; Carbon, 2020). In a situation in which the stationary retail is reliant on offering customers good service, this can lead to inappropriate responses. Empathy, care and concern are especially important for ensuring appropriate interactions (Diebner et al., 2020). Therefore, it is vital for the sales personnel to be able to correctly recognize and respond to customer emotions, in order to form the basis for a valuable interaction (Geiger et al., 2020a; Mattila & Enz, 2002; Meyer et al., 2021). ESAI offer the possibility of opening up a common interaction space in which the customer can actively participate in service creation.

3   Customer Emotions

Emotions influence how people think, communicate and interact, and lead to high mental activities that are perceived as positive or negative (Cabanac, 2002). The digital transformation has already produced various technical options for emotion measurement. In contrast to automatic measurement, self-assessment offers a simple way of measuring emotions that does not require cameras, microphones, or biofeedback readings (Betella & Verschure, 2016; Meyer et al., 2019). Furthermore, emotions are subjectively experienced in different ways (Barrett et al., 2006). Thus, self-assessment provides a suitable way to subjectively assess customer emotions (Barrett et al., 2006; Mau, 2009). There are significant correlations between the customers emotion, behavior and satisfaction (Burns & Neisner, 2006; Martin et al., 2008). Positive emotions are caused by a friendly and pleasant behavior of the sales personnel as well as the negotiation of good prices (Menon & Dubé, 2000). Furthermore, a successful interaction between the sales personnel and the customer evokes positive emotions because social needs are met (Lee & Dubinsky, 2003). Customers with positive emotions show higher satisfaction and improved loyalty (Burns & Neisner, 2006). Negative emotions on the other hand reduce customer satisfaction, which can lead to cancellation of purchases, lasting damage to the customer relationship and negative word-of-mouth (Gelbrich, 2010; Ou & Verhoef, 2017). Negative emotions occur when customers are treated rudely or when they are unsure which product to buy or whether they should by it in the first place (Menon & Dubé, 2000). The sales personnel acts as a critical link between the store and the customer and therefore has a strong influence on the customer's emotions (Lee & Dubinsky, 2003). The empathic ability of the sales personnel can support customer satisfaction, whereas a lack of empathy and the resulting inability to understand the customer's emotion can have a negative impact on the interaction and the perception of the service (Agnihotri & Krush, 2015).

4   Methodology

Our research project is based on the systematic and iterative DSR paradigm proposed by Hevner et al. (2004) that combines behavioral science and design-oriented research, and adds rigor as well as theory to generate prescriptive knowledge about the design of artifacts, such as software, methods, models or concepts (Hevner et al., 2004). Hevner et al. state that in order to create design knowledge, the

development, demonstration, justification and evaluation of an artifact is important. This design knowledge covers three fundamental aspects in DSR: (1) the problem space, (2) the solution space and (3) knowledge that describes the effectiveness of the solution through the generated artifact(s), which is called evaluation. The evaluation describes to what extent the constructed novel artifacts (solution space) address the problem space and satisfy the stakeholders of the problem. Our objective is to design ESAI for stationary retail. In this way, we plan to contribute design knowledge (solution space) to address an emotion-based support for the interaction between the customer and the sales personnel (problem space). This covers knowledge on how ESAI can be designed, including expository instantiations as representations of the design knowledge for purposes of testing (Iivari, 2015). We follow a model for comprehensive DSR projects that involve multiple design steps (see Figure 1) (Vaishnavi et al., 2015).
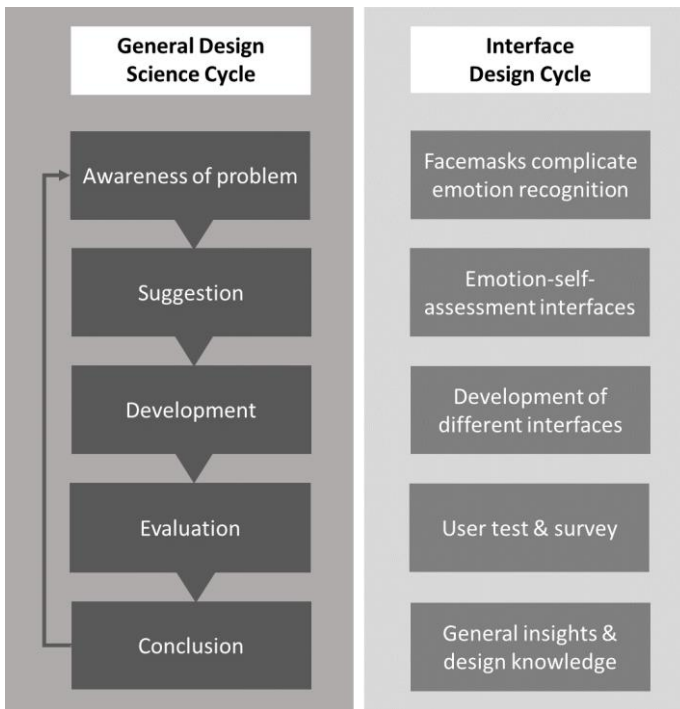
Figure 1: Design Cycle

5   Artifact Design

To approach our research objective, asking whether IT-supported self-assessment offers a suitable approach to measure customer emotions in stationary retail, we designed three different ESAI (see Figure 2). Following the "exploring by building"-principle, the design of the ESAI is very diverse and has an explorative character (Vaishnavi et al., 2015). Due to the high diffusion of smartphones among retail customers, all interfaces were implemented in the form of mobile applications (Fulgoni, 2015). The interface design on the left side of Figure 2 is based on the **Circumplex Model of Affect (CMoA)** (Russell, 1980). The CMoA classifies a variety of emotions using the two dimensions valence and arousal. However, the interface was modified by reducing the complexity of the original. This resulted in four emotional situations. The situations 1 and 2 represent high arousal situations, which are either negative (angry/frustrated) or positive (excited/happy), whereas situations 3 and 4 represent situations of low arousal, which can be negative (sad/tired) or positive (pleased/relaxed). For each emotional situation, an input field was realized in the interface. The emotional situation of the customer can be entered by a tap on one of the input fields. **The Affective Slider (AS)-Interface** in the middle of Figure 2 was designed close to the Affective Slider by Betella & Verschure (2016). This ESAI measures the two dimensions valence and arousal by adjusting emoticons on sliders. In the AS-Interface, the user is therefore able to choose how "happy" and how "excited" he or she is. The emoticons provide feedback about the status of the slider through their facial expressions. For example, if the emoticon for valence is on the left edge, it appears sad; if it is on the right edge, it appears happy. The slider for excitement behaves accordingly: If the emoticon is on the left, it looks relaxed, if it is on the right, it looks excited. After the appropriate emotional situation of the customer is set, it can be entered via the submit button. Finally the **Wheel of Emotions (WoE)-Interface** on the right side of Figure 2 is based on the wheel of emotions by Plutchik (2001). Plutchik distinguishes between eight basic emotions which can be expressed in different intensities. Again, the interface was modified by reduction of the complexity. For the design of the interface only the medium intensity emotions (e. g. anger instead of rage, sadness instead of grief) were used, since these fit better with the context of use. To distinguish the eight input fields more clearly, different colors and symbols were applied. To input an emotion, the customer rotates the wheel until the suitable emotion is found, then selects it by tapping on it.
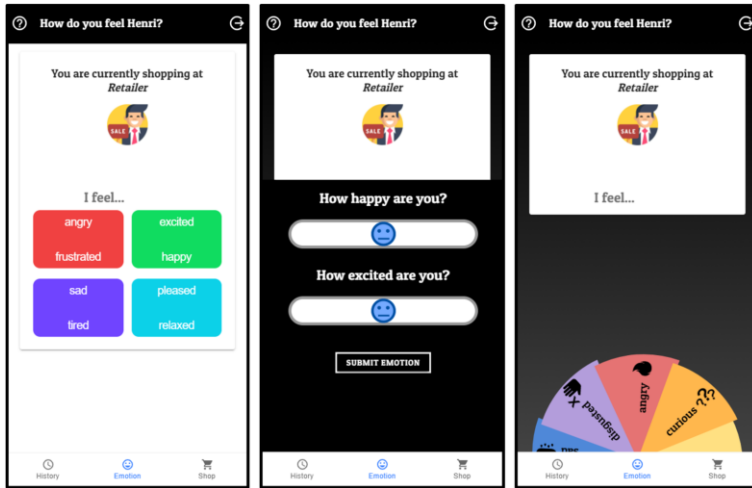
Figure 2: CMoA, AS, WoM

6   Interface Evaluation

Although designed for mobile applications, the user test was implemented through web applications in a browser window (due to the safety reasons during the COVID-19 pandemic). The user test itself consisted of a short introduction to the topic, the interaction with the three ESAI, and subsequent questioning. In order to avoid a preference of one interface, the three interfaces were presented to the participants in a random order. The evaluation of the ESAI was based on the User Experience Questionnaire (UEQ) respectively its modular extension (UEQ+), which allows to freely select and combine individual user experience scales (Schrepp & Thomaschewski, 2019). For the evaluation of the ESAI, the scales *Clarity*, *Visual Aesthetics*, and *Efficiency* were chosen. All UEQ+ scales used a seven-point Likert scale (1 = totally disagree, 7 = totally agree). *Clarity* describes the impression of arrangement, structure and visual complexity of a graphical user interface (Otten et al., 2020). *Visual Aesthetics* measures whether the user perceives the interface as beautiful and appealing (Schrepp & Thomaschewski, 2019). *Efficiency* measures whether the user has the impression that he or she can achieve the goals related to the usage of the interface with minimal effort (Laugwitz et al., 2008). In addition to the UEQ+ scales, various statements and questions concerning emotions and interactions were formulated and rated on a five-point Likert scale (1 = totally disagree, 5 = totally agree). The user test ends with open-ended questions regarding

general criticism as well as missing features. Participants were recruited by distributing the user test via internal mailing lists from a German university. 58 people participated in the survey, with 14 people dropping out before completion, leaving 44 complete datasets. The age of the participants ranges from 16 to 67 years, with an average age of 30.07 years. 59% of the participants are male, 39% female, and 2% other. Table 1 shows the descriptive statistics (mean (M), standard deviation (SD)) of the user test as well as Cronbach´s alphas (α) for the scales *Clarity*, *Visual Aesthetics*, and *Efficiency*.

Table 1: Results of the UEQ+

UEQ+ (7-point Likert-Scale)	CMoA		AS		WoM		α
	M	SD	M	SD	M	SD	
Clarity	5.64	1.12	4.81	1.71	5.56	1.18	0.85
Visual Aesthetics	4.98	1.26	4.33	1.60	5.94	0.85	0.81
Efficiency	5.39	1.29	4.44	1.70	4.90	1.52	0.84

Clarity was rated the highest for CMoA (M = 5.64, SD = 1.12) and the WoE (M = 5.56, SD = 1.18). In terms of *Visual Aesthetics*, the WoE was rated highest by a relatively large margin (M = 5.94, SD = 0.85). *Efficiency* was rated the highest for the CMoA (M = 5.39, SD = 1.29), followed by the WoE (M = 4.90, SD = 1.52). The AS was rated lowest for all three scales (*Clarity*: M = 4.81, SD = 1.71; *Visual Aesthetics*: M = 4.33, SD = 1.60; *Efficiency*: M = 4.44, SD = 1.70). For further analysis, the data was tested for normal distribution using the Shapiro-Wilk test. The results indicate non-normal distribution ($p < 0.01$ for all scales). A non-parametric Friedman test of differences among repeated measures was conducted that retendered a Chi-square value of 55.46, which showed significant differences between the three ESAI ($p < 0.01$). Post-hoc tests (Dunn-Bonferroni tests) pointed out two significantly differences: First, *Visual Aesthetics* significantly differs between the AS and the WoM (Z = -1.03, $p < 0.01$, Cohen's effect size: r = 0.16). Second, *Visual Aesthetics*

significantly differs between the CMoA and the WoE (Z = -0.60, p < 0.05, Cohen's effect size: r = 0.09). The effect sizes correspond to a weak effect.

Table 2: Results of the statements

No.	Statement (5-point Likert-Scale)	Interface					
		CmoA		AS		WoE	
1	I am satisfied with the selection of emotions.	M	SD	M	SD	M	SD
		4.27	0.66	3.30	1.21	4.30	0.63
		M			SD		
2	A good interaction with the sales personal is important to me.	4.68			0.56		
3	Emotions are important in the interaction between customers and sales personnel.	4.02			0.93		
4	IT support for interaction between customers and sales personnel is useful.	3.95			0.91		
5	Giving the customer the possibility to enter his or her emotions is useful.	4.11			0.84		
6	I am willing to share my emotions with the sales personnel.	3.59			1.17		

Tables 2 and 3 show the results of the statements and questions. In terms of satisfaction with the selection of emotions (Statement 1), the WoE received the highest rating (M = 4.30, SD = 0.63), closely followed by the CMoA (M = 4.27, SD = 0.66). This is consistent with the fact that most people say they like WoM the best (Question 1). The participants rate the interaction with the sales personnel as important (M = 4.68, SD = 0.56) and confirm that emotions are an important element in the interaction (M = 4.02, SD = 0.93) (Statements 2 & 3). In addition, the participants agree that the input of emotions (M = 4.11, SD = 0.84) as well as the digital support of the interaction between customers and sales personnel is useful (M = 3.95, SD = 0.91) (Statements 4 & 5). Whether customers are willing to enter their emotions was rated lowest compared to the other statements and shows the largest standard deviation (M = 3.59, SD = 1.17). The evaluation of the open questions revealed further insights about the ESAI. The use of emoticons as well as symbols were repeatedly highlighted positively. The use of colors was also mentioned as being positive. Furthermore, additional functions were desired. One

participant expressed that he/she would like to give the reason for his/her emotion. In addition, two participant would like to indicate whether contact with sales personnel is desired or not. Moreover, one participant would like to indicate his/her location in the store as well as an automatic recognition of whether he/she is entering or leaving the store. Three participants stated that they would not use an ESAI under any circumstances and would rather keep their emotions to themselves, because far too much customer data is already collected in general. Two participants stated that they would only use an ESAI if they received discounts or coupons. In total, 26 participants would use an ESAI in stationary retail, whereas 18 would not use it or were uncertain (Question 2).

Table 3: Results of the questions

No.	Question	Answer		
		CmoA	AS	WoE
1	Overall, which of the three interfaces did you like best?	16	10	18
		Yes	No	Uncertain
2	Would you use the interface that you liked the most in stationary retail?	26	12	6

7   Conclusion & Outlook

The aim of this paper was to answer the question "How can emotion-self-assessment interfaces (ESAI) for stationary retail be designed?". Furthermore, it should be examined whether IT-supported self-assessment offers a suitable approach to measure customer emotions in stationary retail. As part of a DSR project, we designed three ESAI and evaluated them over the course of a design cycle. The user test (n = 44) shows that all ESAI were generally rated as positive (above the midpoint of the used UEQ+ scales) and continued to be perceived as useful by customers (Statement 5). Statistically significant differences could be shown in the evaluation of *Visual Aesthetics* between the WoM and the CMoA as well as the WoM and the AS. The CMoA and the WoE were rated the highest in all UEQ+ scales and were also perceived to be the best overall (Question 1). The reason for this may be due to the fact that the CMoA and the WoM offer the customer discrete emotions or concrete emotional situations. In contrast, the AS offers a more abstract representation of emotions through its two dimensions. In addition, the

CMoA and WoM are characterized by a greater use of colors. The use of colors was highlighted as positive along with the use of emoticons and symbols in the open-ended questions. Therefore, important design knowledge includes the use of discrete emotions as well as the use of colors and emoticons or other symbols to describe the emotions. In the course of this paper, we were able to show that it is possible to use IT-supported self-assessment in stationary retail to capture customer emotions (Statement 5 & Question 2). Furthermore, the evaluation emphasizes the importance of emotions for a successful interaction between customers and sales personnel (Statement 3) (Adolphs, 2003; Geiger et al., 2020a). Derived from the open questions, it became clear that emotions are a very personal topic. Therefore, special attention should be paid to the secure processing and storage of customer data. The resulting design knowledge should be further refined in the next design cycle. A direct comparison of different design features would be a logical step in this process. In addition, a suitable method must be created to display the customer's emotions to the sales personnel. Furthermore, the open-ended questions offer a variety of additional features besides the actual measurement of the customer's emotions, which will be addressed in further research projects. A strong limitation of this work is its highly exploratory nature, which does not allow for direct comparison of individual design features such as the input method, the number and form of emotions displayed, and the use of colors and symbols. Another limitation derives from the fact that the user test was implemented digitally using browser applications instead of the originally planned – and still intended – mobile applications. The designed ESAIs give the customer the opportunity to actively participate in creating a successful interaction between him and the sales personnel by entering his/her own emotions. In this way, stationary retail is supported in utilizing its key competencies of personal interaction properly. However, the success of the ESAI ultimately depends on whether customers want to participate and whether they are honest.

References

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. Nature Reviews. Neuroscience, 4(3), 165–178.

Agnihotri, R., & Krush, M. T. (2015). Salesperson empathy, ethical behaviors, and sales performance: The moderating role of trust in one's manager. Journal of Personal Selling & Sales Management, 35(2), 164–174.

Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., & Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. Humanities and Social Sciences Communications, 7(1), 1–9.

Bahadur, W., Aziz, S., & Zulfiqar, S. (2018). Effect of employee empathy on customer satisfaction and loyalty during employee–customer interactions: The mediating role of customer affective commitment and perceived service quality. Cogent Business & Management, 5(1), 1491780.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2006). The Experience of Emotion. Annual Review of Psychology, 58(1), 373–403.

Betella, A., & Verschure, P. F. M. J. (2016). The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions. PLoS ONE, 11(2).

Burns, D. J., & Neisner, L. (2006). Customer satisfaction in a retail setting: The contribution of emotion. International Journal of Retail & Distribution Management, 34(1), 49–66.

Cabanac, M. (2002). What is emotion? Behavioural Processes, 60(2), 69–83.

Carbon, C.-C. (2020). Wearing Face Masks Strongly Confuses Counterparts in Reading Emotions. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.566886

Diebner, R., Silliman, E., Ungerman, K., & Vancauwenberghe, M. (2020). Adapting customer experience in the time of coronavirus. http://ceros.mckinsey.com/coronavirus-promo

Doherty, N. F., & Ellis-Chadwick, F. (2010). Evaluating the role of electronic commerce in transforming the retail sector. The International Review of Retail, Distribution and Consumer Research, 20(4), 375–378.

Fulgoni, G. M. (2015). The Rise of the Digital Omnivore: What It Means for Advertisers, Publishers, and App Developers. Journal of Advertising Research, 55(2), 115–119.

Fyrberg, A., & Jüriado, R. (2009). What about interaction?: Networks and brands as integrators within service-dominant logic. Journal of Service Management, 20(4), 420–432.

Geiger, M., Robra-Bissantz, S., & Meyer, M. (2020a). Wie aus digitalen Services Wert entsteht: Interaktionen richtig gestalten. HMD Praxis der Wirtschaftsinformatik, 57(3), 385–398.

Geiger, M., Robra-Bissantz, S., & Meyer, M. (2020b). Focus on Interaction: Applying Service-Centric Theories in IS. Proceedings of 33rd Bled EConference Enabling Technology for Sustainable Society, 665–672.

Gelbrich, K. (2010). Anger, frustration, and helplessness after service failure: Coping strategies and effective informational support. Journal of the Academy of Marketing Science, 38(5), 567–585.

Grönroos, C. (2006). Adopting a service logic for marketing. Marketing Theory, 6(3), 317–333.

Grönroos, C., & Voima, P. (2013). Critical service logic: Making sense of value creation and co-creation. Journal of the Academy of Marketing Science, 41(2), 133–150.

Groth, M., & Grandey, A. (2012). From bad to worse: Negative exchange spirals in employee–customer service interactions. Organizational Psychology Review, 2(3), 208–233.

Gutek, B. A., Bhappu, A. D., Liao-Troth, M. A., & Cherry, B. (1999). Distinguishing between service relationships and encounters. Journal of Applied Psychology, 84(2), 218–233.

Hagberg, J., Sundstrom, M., & Egels-Zandén, N. (2016). The digitalization of retailing: An exploratory framework. International Journal of Retail & Distribution Management, 44(7), 694–712.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. 32.

Iivari, J. (2015). Distinguishing and contrasting two strategies for design science research. European Journal of Information Systems, 24(1), 107–115.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. American Economic Review, 93(5), 1449–1475.

Keyes, D. (2018). Amazon captured 4% of US retail sales in 2017. Business Insider. https://www.businessinsider.com/amazon-captured-4-of-us-retail-sales-in-2017-2018-1

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger (Hrsg.), HCI and Usability for Education and Work (S. 63–76). Springer.

Lee, S., & Dubinsky, A. (2003). Influence of salesperson characteristics and customer emotion on retail dyadic relationships. The International Review of Retail, Distribution and Consumer Research, 13(1), 21–36.

Martin, D., O'Neill, M., Hubbard, S., & Palmer, A. (2008). The role of emotion in explaining consumer satisfaction and future behavioural intention. Journal of Services Marketing, 22(3), 224–236.

Mattila, A. S., & Enz, C. A. (2002). The Role of Emotions in Service Encounters. Journal of Service Research, 4(4), 268–277.

Mau, G. (2009). Die Bedeutung der Emotionen beim Besuch von Online-Shops: Messung, Determinanten und Wirkungen (1. Aufl). Gabler Verlag.

Menon, K., & Dubé, L. (2000). Ensuring greater satisfaction by engineering salesperson response to customer emotions. Journal of Retailing, 76(3), 285–307.

Meyer, M., Helmholz, P., Rupprecht, M., Seemann, J., & Tönnishoff, T. (2019). From the Inside Out—A Literature Review on Possibilities of Mobile Emotion Measurement and Recognition. Proceedings of 32th Bled eConference Humanizing Technology for a Sustainable Society, 719–743.

Meyer, M., Siemon, D., & Robra-Bissantz, S. (2021). Emotion-based IS support for Customer-Salesperson Interaction. Proceedings of the 54th Hawaii International Conference on System Sciences, 4426–4435.

Otten, R., Schrepp, M., & Thomaschewski, J. (2020). Visual clarity as mediator between usability and aesthetics.

Otto, J. R., & Chung, Q. B. (2000). A Framework for Cyber-Enhanced Retailing: Integrating E-Commerce Retailing with Brick-and-Mortar Retailing. Electronic Markets, 10(3), 185–191.

Ou, Y.-C., & Verhoef, P. C. (2017). The impact of positive and negative emotions on loyalty intentions and their interactions with customer equity drivers. Journal of Business Research, 80, 106–115.

Plutchik, R. (2001). The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American Scientist, 89(4), 344–350.

Reinartz, W., Wiegand, N., & Imschloss, M. (2019). The impact of digital transformation on the retailing value chain. International Journal of Research in Marketing, 36(3), 350–366.

Russell, J. A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161–1178.

Schrepp, M., & Thomaschewski, J. (2019). Design and Validation of a Framework for the Creation of User Experience Questionnaires. International Journal of Interactive Multimedia and Artificial Intelligence, 5(Regular Issue).

Vaishnavi, V. K., Kuechler, W., & Kuechler, W. (2015). Design Science Research Methods and Patterns: Innovating Information and Communication Technology, 2nd Edition. CRC Press.

van Dolen, W., de Ruyter, K., & Lemmink, J. (2004). An empirical assessment of the influence of customer emotions and contact employee performance on encounter and relationship satisfaction. Journal of Business Research, 57(4), 437–444.

Yang, K., Kim, J., Min, J., & Hernandez-Calderon, A. (2021). Effects of retailers' service quality and legitimacy on behavioral intention: The role of emotions during COVID-19. The Service Industries Journal, 41(1–2), 84–106.

How Digital Market Hosts Control Sellers

SHRADDHA DANANI[1] & JANIS L. GOGAN[2]

[1] Development Institute, Gurgaon, Haryana India; e-mail: shraddhadanani@gmail.com
[2] Bentley University, Waltham Massachusetts, United States of America; e-mail: jgogan@bentley.edu

Abstract How do hosts of digital markets exercise control over sellers? Our three-case study, set in India, reveals that seller control portfolios used by large digital market hosts differ from control portfolios in other contexts (reported in prior research). The platform host neither preselects nor hires most sellers; this limits hosts' control options. The platform supports many short-duration transactions, yet some related processes take place offline – again limiting hosts' control options. In this context of many-sellers, many-buyers, digital market hosts (similar to other controllers) attempt to balance formal and informal controls. By identifying specific control mechanisms that hosts utilize, our study findings provide a useful foundation to support further research on control challenges in digital markets and other digital platforms.

1   Introduction

Control is both necessary and insufficient to digital platform success (Tiwana et al. 2010; Buchwald et al. 2014, Shafiei Gol et al. 2019). A <u>digital market</u> is a digital platform that connects buyers and sellers via their computers or mobile devices. Digital market hosts confront several control challenges. Sellers are not platform employees, and most sellers are independent entities; they voluntarily participate and may exit at any time. Unlike many other digital platforms, on a digital market many sales transactions involve both on-platform and off-platform processes. Because of this, the platform host's control leverage is limited (Felin & Zenger 2014). Their control authority is further constrained by the fact that hosts and sellers are not co-located (difficult to observe off-platform behavior). Reflecting these and other concerns, prior studies indicate that many platform hosts try to coax participants to align with platform priorities (Parker & Van Alstyne 2018), such as by orchestrating participants' interactions (Brown & Grant 2005; Tiwana 2014). Most platform hosts aim to strike a balance between tight and loose control, and between attracting and controlling buyers and sellers (de Reuver et al. 2018, Parker & Van Alstyne 2018).

Some helpful automated controls are embedded in digital platform software (Parker & Van Alstyne 2018), and hosts also have the option of evicting participants who misbehave (Parker & Van Alstyne 2018; Aulkemeier et al. 2019). Since eviction is a last resort, it would be helpful to chronicle in detail how digital market platform hosts actually exercise control over sellers, by closely examining their seller control portfolios, and circumstances that influence which controls are used and when. A recent study (Croitor et al. 2021) investigated sellers' perceptions about digital market hosts' use of two formal and informal control modes (described below). However, to date no prior in-depth study has comprehensively examined how digital market hosts exercise control over sellers. Thus, our three-case study posed the following research question: *How do digital market hosts exercise control over participating sellers?*

1.1   Brief Overview of Prior Control Research

An organization's portfolio of manual and computer-based control mechanisms aims to prevent, detect, and correct adverse events, in ways that align with strategic and operational priorities for organizational control (Cardinal et al. 2017), accounting control (Gelinas & Dull 2008), or IS control (Kirsch et al. 2002;

Choudhury & Sabherwal 2003; Heiskanen et al. 2014; Remus & Wiener 2012, Wiener et al. 2019). Prior studies categorize control mechanisms in two <u>modes</u>: 1) *formal (*process controls and outcome controls) and 2) *informal* (relational controls and mechanisms that support self-control) (Chua et al. 2012; Merchant & Van der Stede 2017). Until recently, prior platform studies articulated control challenges and offered guidance on balancing control portfolios in terms of these higher-level control modes; most prior platform control studies did not closely investigate the specific formal and informal control mechanisms hosts used to achieve balanced control (Yoo et al. 2012; Halckenhaeusser et al. 2020).

A survey of sellers on Amazon and Etsy (Croiter et al. 2021) reveals that control perceived to be strict (e.g., screening mechanisms that block undesired sellers) negatively affect sellers' intrinsic motivation, their perceptions of platform usefulness, and their satisfaction with the platform. Informal relational controls -- what Ouchi (1980) referred to as Clan Control -- positively influenced seller perceptions. Croitor et al. contributed helpful early findings on sellers' attitudes about specific controls, and their behavioral intentions. A recent literature review (Danani, 2021) called for in-depth comprehensive examination of specific control mechanisms that digital market hosts use to exercise control over sellers.

The remainder of this paper is structured as follows. In Section 2 we describe our research method. After presenting findings from our three-case study in Section 3, we briefly discuss those findings which are consistent with prior control studies, and point to other findings which uniquely reflect the digital market context. In Section 4 we discussion contributions, study limitations, and future research opportunities.

2 Research Method

Case research is appropriate for learning 'how' and 'why' managerial phenomena unfold in complex contexts (Yin 2009). Our three-case study sought to learn in detail how digital market hosts exercise control over sellers. We identified three prominent digital markets operating in India (home country of first author). Each digital market serves many consumers and many sellers. MC, GC, and FC (companies anonymized) are each at a mature stage of operations (neither startup nor in decline). Each digital market connects many consumers with 100,000 or more sellers, offering many

products. From left to right, Table 1 summarizes key features of these three digital market cases, in the order in which we gathered data.

For each case, semi-structured interviews were conducted with an operations manager, merchant manager, operations head, and merchant head. Snowball sampling led us to other interviewees. Interview and site observation notes were typed within 24 hours and corroborated/triangulated case findings were confirmed with each firm's operations manager and also with owners of some seller firms. 12 interviews were conducted at MC , 12 at GC; 9 at FC. Interview findings were compared with more than 360 company resources, including training materials, policy documents, manuals, dashboards, digital communications, and observed system interfaces. Here are two triangulation examples: 1) A content manager's interview was corroborated with MC's catalogue creation guideline documents. 2) An operations manager interview was corroborated with training documents, seller portal and operations guideline documents.

Table 1: Three Digital Market Hosts

MultiCart (MC)	GlobalCart (GC)	FastCart (FC)
October 2018--Jan 2019	January 2017-- July 2018	March 2019--May 2019
Launched 2007	Launched 2012	Launched 2010
80M product SKUs	100M + product SKUs	60M + product SKUs
100,000 sellers	400,000 sellers 120,000 active sellers	300,000 sellers 60% to 75% active sellers
100M + consumers	150M + consumers	10M + consumers
In 2007 MC targeted a niche market. Later it expanded into electronics, apparel, appliances, books, toys, other consumer products, and groceries. Today it targets consumers all over India.	GC operates in many countries; this study focused on its operations in India. Its systems and infrastructure connect small to medium size mostly-independent sellers with consumers all over India.	FC does not produce or trade any products under its brand. Its logistics infrastructure services 3000 Indian cities. FC targets consumers in smaller towns. It offers low-price high-volume products.

Our analysis utilized both a positivist lens (we coded case data for known control mechanisms, classified in informal or formal control modes), and a grounded theory lens (we identified control mechanisms not discussed in prior studies and, iterating between data collection and analysis, we identified new control themes). Thus, both open and axial coding described each organization's control portfolio. For example, three open codes – *specify delivery milestones*, s*pecify target timeframe, c*learly defined interaction success criteria -- were grouped into an axial code: *Clearly defined performance criteria.*

When necessary, we re-contacted interviewees to clarify details and obtain supporting documents (e.g., after analysing a merchant manager interview, we asked this interviewee to clarify details about performance metrics and evaluation criteria). Interview findings were corroborated via primary-source or secondary-source documents and other interviews. This helpful triangulation led us to modify some initial concepts. For example: we saw that MC training documents and guidelines transferred <u>process</u> knowledge to sellers. Later, we obtained evidence indicating that training did help sellers perform effectively. Thus, we mapped training to both formal process control and informal self-control.

3   Study Findings

The study findings revealed that hosts' seller control portfolios are comprised of control mechanisms implemented at three levels: system (automated control mechanisms), participants (control exerted by host employees, consumers, peer sellers and seller themselves) and host firm (policies, initiatives, values and culture). Figure 1, a Digital Market Seller Control Framework, summarizes three broad levels of control mechanisms (automated, participant-level, host firm-level), mapped to formal and informal control modes, and influencing consumer-seller interactions.
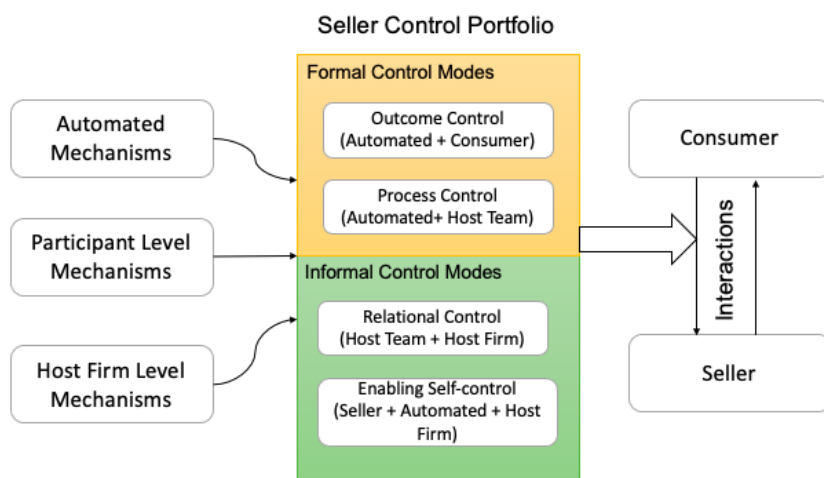
Figure 1: Digital Market Platform Seller Control Framework

Table 2 (shaded in grey) summarizes conventional (single-mode) formal and informal seller controls identified in the three cases. Table 3 (not shaded) summarizes hybrid (multi-mode or multi-mechanism) controls in the three cases.

Table 2 Three-Case Comparison: Single-Mode Controls in Digital Market Platforms

Mechanism	Mode	MC	GC	FC

Formal Controls
    OC: Outcome control PC: Process Control
    a: automated f: firm p: participant				

Verify adherence to catalogue guidelines, participation terms	PC a, p	√	√	√
Measure rate of order acceptance by seller	OC a	√	√	NO
Measure consumer returns (indirectly gauge product quality)	OC a	√	√	NO
Measure product quality through customer returns	OC a, p	√	√	NO
Measure consumer satisfaction on order cycle	OC a	√	√	√

Measure seller order (value + volume) in a given period	OC a	√	√	NO
Measure consumer satisfaction on query/issue resolution	OC a, p	√	√	√
Measure number of completed returns request	OC a	√	√	NO
Informal Controls RC: Relational Control; SC: Support of Self-Control				
Support sellers through community platform	RC p, f	√	√	NO
Assist sellers with registration, catalogue creation, other setup	RC p	√	√	√
Assist sellers with issue resolution	RC p	√	√	√
Connect with seller through calls and meetings	RC p	√	√	√
Encourage sellers to recruit new sellers to the platform	RC p	NO	√	NO
Organize seller group events	RC f	√	√ +	NO
Training: platform norms, values and objectives	RC f	√	√	√
Sellers decide re pricing, promotion, QC, packaging, shipping	SC p	√	√	√
Link financial benefits with order performance	SC a, p	√	√	NO

Most formal controls are enacted via automated systems. An MC operations manager stated that automated controls monitor consumers' product return requests, and that "we do not monitor if the seller packed the right product, as ordered by the consumer." GC's operations manager said "For every order, performance against checkpoint parameters is recorded. The system calculates average value [for] a 30-day [period]."

Table 3 Three-Case Comparison: Hybrid Controls (multiple mechanisms/modes per control)

Mechanism	Mode	MC	GC	FC
Controls that Combine 2 Formal Modes PC: process control; OC: outcome control				
Measure time to pack and ship	PC a, OC a	√	√	√
Measure time to deliver to end consumer	PC a, OC a	√	√	NO
Measure pickup reattempt rate	PC a, OC a	NO	√	NO
Measure time taken to resolve consumer query/issue	PC a, OC	√	√	√
Measure time taken to process refunds on returns request	PC, OC	√	√	NO
Controls that Combine 2 Informal Modes RC: relational control; SC: support for self-control				
Create promotional events	RC, SC	√	√+	NO
Organize seller appreciation events	RC, SC	√	√+	√
Promote seller success stories	RC, SC	√	√+	NO
Provide access to comprehensive training material	RC, SC	√	√	√
Best practices training: QC, packaging, shipping, etc.	RC, SC	√	√	NO
Controls that Combine Formal and Informal Modes				
Training: order delivery, queries, returns), performance criteria	SC, PC	√	√	√

Automatically cancel order if not shipped on time	SC, PC	√	√	√
Display system-generated seller service rating	OC, SC	√	√	NO
Prominently display high performing sellers' products	OC, SC	√	√	NO
Prominently display products highly rated by consumers	OC, SC	√	√	NO

Hosts rely on employee teams to manually measure sellers' content quality, and rely on consumers to judge sellers' product and service quality. All three hosts encourage consumers to evaluate sellers via quantitative and qualitative ratings of product quality, service experience, and seller query resolution. These ratings are displayed on or linked to sellers' product listing pages). Consumer evaluations weigh heavily in hosts' overall seller ratings (along with sellers' sales per evaluation period.

All sellers receive training that explains terminology, processes and instructions. MC and GC community portals target all sellers with these resources. Other informal controls aim to build relationships with sellers. Host teams attempt to keep sellers engaged with their platform (participant-level relational controls). For example, MC and GC invite high-performing sellers to local city chapter events. "Sellers who perform well are very important for us," said an MC Operations Manger. "We need to … support them if there is an issue." Awards and recognition events (firm-level controls) also aim to strengthen high performing sellers' association with the platform. A host merchant coordinator organizes meetings, calls, awards events, advanced training seminars and other events. MC merchant coordinator: "We meet up with them, one to one or in a group setting, region-wise. "

Other control mechanisms encourage seller self-control, and these intertwine with formal controls, such as performance-triggered rewards and penalties. MC's Operations Manager stated that sellers "control their performance. We openly display their performance report card … [Sellers try to] keep their consumers happy and get good ratings." FC manager: "The weighted average of customer ratings for a seller is displayed next to seller name on every product listing. Future customers can view the rating, identify the reason ..." Hosts respond to poor performance with

warnings or penalties. GC Operations Manager: "We observe [a problematic seller] for a fixed number of days. If performance does not improve, we completely deactivate the seller account and remove all listings."

4   Contributions, Limitations and Directions for Further Researc

The three cases reveal that hosts use many formal and informal controls, including providing resources that enable seller self-control. Hosts also deputize consumers to exercise control over sellers, through quantitative ratings and qualitative feedback. Consistent with the 'Goldilocks' challenge (Ghazawneh & Henfridsson 2013), hosts aim for balance; that is, overall control that is neither too-tight nor too-loose (Tiwana 2014; Benlian et al. 2015). In digital markets, the Goldilocks challenge appears to be partly influenced by interdependence among hosts, sellers and buyers, and partly influenced by the fact that sellers and consumers are only loosely tied to the market platform (they can buy or sell elsewhere). In this interdependent yet loosely-coupled context, hosts apply tight system-based controls, and authorize consumers to exercise tight control by evaluating sellers' product and service quality. Hosts offset tight controls with looser informal relational controls and by mechanisms that support seller self-control. We believe a similar balancing of formal/informal and preventive/detective controls likely applies in other contexts characterized by both interdependence and loose coupling -- such as platforms that support ride-sharing, short-term home rentals and other 'sharing economy' services. Future in-depth and holistic case studies set in these other digital platform contexts are still needed.

Our study was based in India, which limits the generalizability of our findings. Future case studies can usefully focus on culturally-different contexts like Europe, North or South America, East Asia, and Africa. An embedded-cases study of a huge multinational like Amazon or AliBaba could inevestigate why and to what extent controls are chosen and exercised differently by headquarters versus managers in different regions. Our study provides a helpful foundation for future case studies as well as large-sample surveys investigating hosts' reliance on specific seller control mechanisms (in differently-configured control portfolios).

Our study did not directly examine how specific controls affect seller employees' attitudes or behaviour (an important early contribution of Croitor et al. 2021). In future studies, it would be helpful to take a 360^0 view of stakeholder responses to a broader set of formal and informal control mechanisms (important, since hosts need to fairly balance sellers' and consumers' interests). Studies informed by service-dominant logic (Lusch & Nambisan 2015) could helpfully explore whether and how value cocreation (or inadvertent value destruction) is associated with differently-configured digital platform control portfolios.

In our three cases, hosts focused on building relationships with high-performing sellers. As for high-potential (but as yet under-performing) sellers (e.g., those serving small but profitable market niches or offering innovative products which consumers do not yet understand): our findings suggest that digital market hosts adopt a 'sink or swim' approach. Perhaps this is because a seller's success potential is hard to spot. Stronger data analytics might help hosts identify high-potential sellers by attending to faint signals that point to consumer acceptance and likely profitability in small market niches. Future design-science studies could contribute, by testing alternative analytic techniques that may strengthen those faint signals.

Our three-case study revealed that digital market hosts allow sellers to decide how to carry out many processes (on-platform and off-platform). Advanced information systems and supporting infra-structures might in future enable hosts to exercise tighter automated control. Our case study provides a basis for comparison with future studies that could chronicle whether and how host control changes as smarter systems (supported by artificial intelligence, blockchains, etc.) take on additional control functions, and also chronicle how host employees, sellers and consumers react to such changes. Given the rapid evolution of ICT, many future studies utilizing multiple research methods, are needed, to continue to shed helpful light on mechanisms of control in digital markets and on other digital platforms.

References

Aulkemeier F., Iacob M., van Hillegersberg J. 2019. Platform-based collaboration in digital ecosystems. Electronic Markets (29), 597–608.
Benlian A., Hilkert D., Hess, T. 2015. How open is this platform? The meaning and measurement of platform openness from the complementors' perspective, Journal of Information Technology, 30:3), 209–228.

Brown A.E. & Grant, G.G. 2005. Framing the frameworks: A review of IT governance research. Communications of the Association for Information Systems (15), 696–712.

Buchwald A., Urbach N., Ahlemann F. 2014. Business value through controlled IT: Toward an integrated model of IT governance success and its impact. Journal of Information Technology (29:2), 128–147.

Choudhury V. & Sabherwal R. 2003. Portfolios of control in outsourced software development projects. Information Systems Research (14:3), 291–314.

Chua C.E.H., Lim W.-K., Soh C., Sia S.K. 2012. Enacting clan control in complex IT projects: A social capital perspective, MIS Quarterly (36:2), 577-600.

Cram W.A., Brohman/ M.K., Gallupe, R. B. 2016a. Hitting a moving target: a process model of information systems control change. Information Systems Journal (26:3), 195–226.

Cram W.A., Brohman K., Gallupe R.B. 2016b. Information systems control: A review and framework for emerging information systems processes. Journal of the Association for Information Systems (17:4), 216–266.

Cram W.A & Gallupe B.R. 2016. A method to evaluate information systems control alignment. Journal of Information Systems (30:1), 117–135.

Croitor, E., Werner, D., Adam, M., Benlian, A. 2021. Opposing effects of input control and clan control for sellers on e-marketplace platforms. Electronic Markets, 1-16.

Danani, S. 2021. How and Why Do Digital Market Platform Hosts Exercise Control Over Selers? PhD Thesis, Management Development Institute, Gurgaon. India.

de Reuver M., Sørensen C., Basole R.C. 2018. The digital platform: A research agenda. Journal of Information Technology (33:2), 124–135.

Eisenhardt K.M. 1989. Building theories from case study research. Academy of Management Review (14:4), 532–550.

Felin T. & Zenger T.R. 2014. Closed or open innovation? Problem solving and the governance choice. Research Policy. (43:5), 914–925.

Fichman R., Dos Santos B., Zheng Z. 2014. Digital innovation as a fundamental and powerful concept in the information systems curriculum. MIS Quarterly (38:2), 329-A15.

Gelinas U.J. & Dull R.B. 2008. Accounting Information Systems, 7th Edition. Thomson Southwestern.

Ghazawneh A. & Henfridsson O. 2013. Balancing platform control and external contribution in third-party development: The boundary resources model. Information Systems Journal, 23(2), 173–192.

Gregory R.W. & Keil M. 2014. Blending bureaucratic and collaborative management styles to achieve control ambidexterity in IS projects. European Journal of Information Systems (23:3), 343-356.

Halckenhaeusser A, Foerderer J., Heinzl A. 2020. Platform governance mechanisms: An integrated literature review and research directions. Proceedings of ECIS 2020.

Heiskanen A., Newman M., Eklin M. 2008. Control, trust, power, and the dynamics of information system outsourcing relationships: a process study of contractual software development. Journal of Strategic Information Systems (17:4), 268–286.

Kirsch L.J. 1997. Portfolios of control modes and IS project management. Information Systems Research (8:3), 215-239

Kirsch L.J., Sambamurthy V., Ko, D-G, Purvis R.L. 2002. Controlling information systems development projects: The view from the client. Management Science (48:4), 484–498.

Lusch R.F. & Nambisan S. 2015. Service innovation: A service-dominant logic perspective. MIS Quarterly (39:1), 155-176.

Merchant K.A. & Van der Stede W.A. 2017. Management Control Systems: Performance Measurement, Evaluation and Incentives, Pearson Education.

Ouchi W.G. 1980. Markets, bureaucracies, and clans. Administrative Science Quarterly (25:1), 129–141.

Parker G., & Van Alstyne M. 2018. Innovation, openness, and platform control. Management Science (64:7), 3015–3032.

Remus U. & Wiener, M. 2012. The amount of control in offshore software development projects, Journal of Global Information Management (20:4), 1-26.

Rustagi S., King W.R., & Kirsch L.J. 2008. Predictors of formal control usage in IT outsourcing partnerships. Information Systems Research (19:2), 126–143.

Shafiei Gol E., Stein M.-K. & M. Avital. 2019. Crowdwork platform governance toward organizational value creation. Journal of Strategic Information Systems.

Tiwana A. 2014. Platform Ecosystems : Aligning Architecture, Governance, and Strategy. Morgan Kaufmann.

Tiwana A., Konsynski B., Bush A.A. 2010. Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. Information Systems Research (21:4), 675–687.

Wiener M., Mähring M., Remus U., Saunders C., Cram W.A. 2019. Moving IS project control research into the digital era: The "why" of control and the concept of control purpose. InformationSystems Research (30:4), 1387–1401.

Yin R.K. 2009, Case Study Research: Design and Methods, (4th Ed). Sage Publications.

Yoo Y., Boland Jr. R.J., Lyytinen K., Majchrzak A. 2012. Organizing for Innovation in the Digitized World. Organization Science, 23(5), 1398–1408.

A Novel Combined Business Recommender System model Using Customer Investment Service Feedback

Asefeh Asemi & Andrea Ko

Corvinus University of Budapest, Doctoral School of Business Informatics, Budapest, Hungary; e-mail: asemi.asefeh@uni-corvinus.hu, andrea.ko@uni-corvinus.hu

Abstract The aim of the study was to present a new business model of an investment recommender system using customer investment service feedback based on fuzzy neural inference solutions and customized investment services. The model designed to support the system's process in investment companies. The type of research was qualitative and used of exploratory study and extensive library research. The model divided into two main parts using customer investment service feedback: data analysis and decision making. In this model, seven group factors proposed to implement the model of the proposed system of investment jobs through the potential investors. Machine learning use in this process and next ANFIS, which is an implementation of the neural art community uses the establishment of fuzzy logic judgment directly forward. The system act like a system consultant, studies the investor's past behavior and recommends relevant and accurate recommendations to the user for most appropriate investment.

1 Introduction

The study of customer behavior in the management of affairs is in many respects almost a new issue, in this regard, there are very few resources to study and learn about the experiences of others in the world. The importance and dimensions of paying attention to the behavior of online customers have not yet received enough attention. Recommendation systems are the tools used to supply pointers to customers based on their requirements (Kanaujia et al., 2017). This research will provide a new method for investment service customization using a recommender system based on ANFIS. We designed a novel combined recommender system framework based on a neuro-fuzzy inference system. In this way, the effective factors from customer experiences will categorize by machine learning techniques and factor analysis. Since a recommender system framework will design to provide suitable and novel customer investment service. These systems act usefully if they implement based on customer experience. The main issue in this research is what business model can be designed for a recommender system ANFIS-based, to present investment recommendations based on investor types and investment indicators?

2 Theoretical Framework

In this research, the customer is the investor. The investor, as a customer, buys services or investment products from the investment company or investment consulting companies. Investor is "a person who puts money into something in order to make a profit or get an advantage" (Cambridge Dictionary, n.d.). The theoretical foundations of this research are summarized in key concepts as the following.

2.1 Investor Behavior

Understanding the investor's behavior as a customer is complex in the decision-making process when buying a product or service. Investor behavior is the result of various cognitive processes, social interactions and social institutions, and the ability of investment firms to predict investor behavior is very important. A deep understanding of investors behavior creates more opportunities to predict and guide their behavior. The use of intelligent recommender systems is also an effective tool in predicting investor behavior. Various factors influence the analysis of investor behavior. One of these factors refers to the experience that the customer or investor

gains in using the services or product of an investment company. This experience has a significant impact on both his loyalty and attracting new investors. One of the important dimensions of customer behavior is its social nature. Although we collect data from investors about their behavior, but the influence of other investors, social institutions and social regulations governing society are also very important in these behaviors. Therefore, the investors can only be understood and examined based on their relationships with other investors and in the framework of a larger social environment. "Customer engagement behavior can serve as a useful framework for classifying and segmenting customers, based on their propensity to engage and the types of engagement behaviors they display" (van Doorn et al., 2010). Of course, investors can be either individuals or organizations. Due to the differences between these two types of investors, there is a lot in common between them.

2.2 Investor Behavior in Investment Decision Making

Most investors do not act individually in decision making and consider the opinions of different people in the investment process. In families, different people may be involved in different stages of the investment decision process. The lower the investment, the smaller the number of people involved in the decision-making process. Of course, depending on the cognitive aspects of individuals and their individual characteristics, how they consult with different people in decision making is different. People who are involved in decision-making may even come from a variety of backgrounds. In a family decision, the number of people present in a decision and the type of people are usually constant. Investor behavior varies in different investment situations and in the decision-making process. This behavior includes how to decide on the type of investment, how to invest, places to invest, review of different portfolios, evaluation of services and products offered by investment companies. The decision-making method varies depending on whether the investor is involved in a new investment or needs an extension of a previous investment. In a simple investment situation, the investor needs to take a series of simple steps, but in a non-simple investment situation, he needs more information and time to ensure the investment decision. Of course, it is necessary to point out that in some investments, such as investing in cryptocurrency, it is a kind of game with money. In this regard, a large amount of information and technical and fundamental analysis is required. Slovic (1972) says the basic tenet of those in charge of helping the investor to make market decisions seems to be "the more information,

the better." Various key factors play a role in investor behavior when making investment decisions. For example, the opinion of specialists, the opinion of people who have experience in the field of investment and are experts in the field of investment. Even the opinion of people who have invested in a field for the first time can be effective in the decision-making process of other investors. Direct or indirect marketing of media and social networks in the field of investment news can have a great impact on investor behavior when making investment decisions. Executors or agents of investment in various fields and their performance are also effective in this process. Finally, it can be said that the most important and effective factor is the opinion of investors who directly use the products and services of an investment company. Awareness of the needs of investors and knowledge of the investment decision process is the basis of the success of an investment company. These companies must be able to pass the investor through various decision-making stages step by step. Including in recognizing the need, gathering information about that investment field of interest to the investor or suitable for the investor, evaluating different options, investing decisions and significant issues in investor behavior after the investment. Adequate knowledge and understanding of investment companies helps them to design effective and successful portfolios for investment. Christensen and Bower (1996) stated that "technological advances can exceed the required performance in a market, technologies that can initially only be used later in emerging markets can attack major markets and move incoming companies to victory over established companies". It can be said that the design of investment proposing systems is one of these technical and effective advances in the investment market.

2.3   Investor experience & feedback

An investor's experience as a customer is the result of the investor's interaction with the company that assists an individual or organization in investing and uses the company's products and services in the investment. This investment can be made directly by the investor or by an intermediary. The experience gained can be during and after the investment. "This interaction is made up of three parts: the customer journey, the brand touchpoints the customer interacts with, and the environments the customer experiences (including digital environment) during their experience. Good customer experience means that the individual's experience during all points of contact matches the individual's expectations. Gartner asserts the importance of

managing the customer's experience" (Verhoef, et al, 2009). Customer experience implies customer involvement at different levels – such as rational, emotional, sensorial, physical, and spiritual (Janakiraman, Meyer, & Morales, 2006). The experience of investors may be gained directly or indirectly. In direct experience, the process of interaction starts from the investor, while in indirect experience, the investor gains the experience from news media in different contexts. Even this experience can be achieved through verbal interaction with other investors. "Customer experience is created by the contribution of not only the customers' values but also by the contribution of the company providing the experience" (Gentile et al, 2007). All of the events experienced by customers before and after a purchase are part of the customer experience. What customer experience is personal and may involve sensory, emotional, rational, and physical aspects to create a memorable experience. In the retail industry, both companies and customers play a big role in creating customer experience (Andajani, 2015). The investor's experiences can be in the form of "investor feedback". Customer feedback exposes their degree of satisfaction and assists product, customer success, and advertising groups to recognize the place there is room for improvement. Companies can gather customer feedback proactively via polling and surveying customers, interviewing them, or asking for reviews (Customer Feedback Definition | Pendo.io Glossary). Gartner believes that "the company's customer outstanding experience greatly influences their long-term exchange behavior and reflects the true drivers of loyalty" (www.gartner.com). The investor feedback helps to measure the satisfaction of the investment company's products and services. Without investor feedback, no company can be assured of the value of the product or service it offers. The more importance is given to investor feedback, the easier it is to retain the investor and the higher the investor loyalty. It is possible to receive feedback in different ways. Depending on the different investor groups, types of services, and products, the appropriate application can be used. The method of receiving investor feedback should be commensurate with their needs and conditions and it should be at any time and in the simplest possible way with proper access. Another important issue is the proper and timely use of investor feedback in the use of products and services. The use of intelligent systems is very effective in skillful and timely analysis of investor feedback.

2.4   Investment Recommender System

The information system here is any kind of system where a lot of information is stored, and the information system can be equipped with Recommender Systems. "Recommender Systems (RSs) are software tools and techniques that provide suggestions for items that are most likely of interest to a particular user" (Burke, 2007), (Resnick et al., 1994), (Resnick &Varian, 1997). Liang (2008) believes that RSs are a type of DSSs that analyzes user behavior and proposes based on its results. Recommender systems are a digital solution supporting financial investments. This digital support is usually implemented by recommender systems, which gives customized offer for customers according to their needs. Figure 1 shows the relation between the basic sections of the ontology with Recommender System. As mentioned before, the information system here is any kind of system with a large amount of stored information. This information system can be equipped with recommender systems. In the recommender investment systems, this program uses the techniques and methods of the recommender system to meet the information needs of the customer in investing. In fact, a recommender system is designed for the user. Customer or investor behavior plays a key role in evaluating the recommender system. We cannot evaluate the recommender system without considering the user as a separate class. For this reason, we consider the behavior of the user (investor) of this system as the main feature of the investor in using the recommender system.
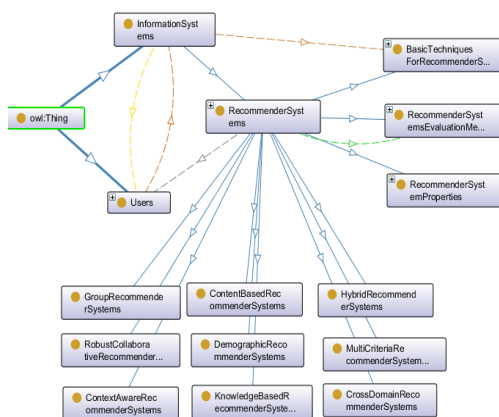
Figure 1: Relation between the basic sections of the ontology with recommender system

Usually, the client or investor is looking for an answer to their information needs in investing. Seek means search, and search is what the user says when the search is done. Of course, instead of a human user, we can also consider a machine user or, in a more advanced form, an investor robot. To implement the core function of the investment recommender system, identifying useful items for the client or investor must anticipate that an item is worth the investment recommendation. To do this, the system must be able to predict the profitability of some items for investment. Even the system can compare the usefulness of some items for investment with others. Based on this comparison, the system then decides which items to recommend for investing based on the customers' group. Various recommendation techniques are used to predict items based on the needs or preferences of the investor.

2.5   Business Model Innovation

Business model innovation is the development of new, unique concepts supporting an organization's financial viability, including its mission, and the processes for bringing those concepts to fruition (Cole, 2015). In this research, we presented a novel model of investment recommender system that supports the processes of achieving the goals of investment companies in the business model. Various technologies in the application of recommender systems are important applications in presenting current business models applied in investment portfolio to investors.

3   Literature Review

Paranjape-Voditel and Umesh (2013) proposed a stock market portfolio recommender machine based totally on association rule mining that analyses inventory records and suggests a ranked basket of stocks. In 2017 proposed a collaborative filtering-based recommender device for monetary analysis based on Saving, Expenditure, and Investment the usage of Apache Hadoop and Apache Mahout (Kanaujia et al.,2017). Hernández et al. (2019) evaluated the state of the art on Financial Technology for the layout of a novel recommender system. They presented a social computing platform that is proposed, based on Virtual Organizations, that allows enhancing person experience in moves that are related to the method of funding recommendation. Tejeda-Lorente et al (2019) proposed a novel recommender system, which is conscious of the risks related to unique hedge funds, considering multiple factors, such as modern-day yields, historic

performance, diversification by way of industry, etc. Tarnowska et al., (2020) presented a Recommender System for Improving Customer Loyalty. This recommender system addressed various important problems. (1) it presents a favored framework to assist managers to decide which moves are possible to have the largest **influence** the internet promoter rating. (2) the consequences are based totally completely definitely on multiple clients. (3) its dietary supplements ordinary textual content mining alternatives. The recommender gadget allows users to view specific, anonymous feedback related to the right clients. (4) ultimately, the computer offers a sensitivity assessment feature.

4   Research Methodology

The study is exploratory research. The qualitative data collection methods applied to collect data from previous research and library studies. In this research, a novel business model for an investment recommender system proposed based on ANFIS that analyses customer data and suggested several recommendations based on customer needs. This model is different compared to existing systems because it found the correlation between potential customers' demographic/personality traits, potential customer's investment indicators, and investment's products and services and on this base, recommends a portfolio based on the customers' needs. An intelligent fuzzy framework uses for generating association rules. The novel methods implement using machine learning and fuzzy logic. Thorough experimentation performs on the Portfolio dataset based on a web-based investment questionnaire. Our approach demonstrates the application of soft computing techniques like data mining, machine learning, factor analysis, and fuzzy classification in the design of recommender systems.

5   A Novel Combined Recommender System Business Model

Based on the extensive study on the previous research, we propose four main steps in our investment recommender system business model. In the first step, the customer's types clustered by an unsupervised machine learning technique based on multiple variables that are extracted from gathered data by the questionnaire. In the second step, the same variables analyze by the factor analysis method to identify customer's investment indicators. This nomination can be based on strong features in each category. Also, the expert viewpoints can use to finalize the indicators. In

the third step, an ANFIS solution develops based on the output of the first step for predicting the customer investment type. An artificial neural community implementation, ANFIS is based totally on Takagi–Sugeno FIS at first presented by Jang (1993). It makes fuzzy logic judgment deployment extra straightforward in contrast to the normal neural network simulations as defined in (Asemi & Asemi, 2014). The input of ANFIS is a summarization of factors in the previous step and the output is scoring categories for the customer. For example, if a category gets the highest score means that the considered customer's investment belongs to this category. In ANFIS two types of rules are contributed for prediction, i) rules which are designed based on data training and ii) rules which we design based on our analysis from customer categorization and expert viewpoints. In this model, we designed membership functions on ANFIS based on the nature of input factors and measurement scales. In the fourth step, a recommender system provides proper recommendations for the customer with a predicted type. The customers can invest based on these recommendations. Figure 2 shows the research framework. It shows how we answered the research question of the study. According to different functions, the research design divided into three phases:

First Phase (Data gathering): The data-gathering phase includes the data acquisition layer and the data storage layer. Second Phase (Data analysis): The second phase includes two functions: (a) machine learning techniques (clustering and factor analysis) and (b) ANFIS. Third Phase (Decision): This phase includes the recommendation layer, and it presents information to the customer and receives their feedback. This feedback shows the probability errors, then the errors refer to the data analysis phase for correction. According to different functions, the business recommender system model structure includes these layers: 1. data acquisition layer, 2. data storage layer, 3. machine learning layer, 4. ANFIS layers (fuzzification, implication rules, normalization, defuzzification, integration, or aggregated output membership function), 5. investment recommendation and feedback layer, or application layer. All the parts of the research framework are specified as follows:

5.1 Data Acquisition Layer

The purpose of this layer is to collect data from the users and find out how conscious their readers are about their finances. In the web-based investment questionnaire, they asked about the savings, spending habits, use of digital financial solutions,

communication ways, satisfaction form different financial companies, banks, & organizations, or their view of the state of the economy in the years to come, and demographic data. The data transfer to the next layer cloud data storage from this layer. This type of system recommends items based on user demographics. The basic premise of these recommender systems is to provide different recommendations for different groups of users. Many websites today offer simple and effective recommendations based on user demographic and personalized information. For example, users are referred to specific websites based on their language or country. The offers may also be customized according to the age of the user. While these approaches have been quite common in the marketing, relatively little recommender systems research has been done on demographic based recommender systems.

5.2   Data Storage Layer

The data storage layer stores all the data of users in the company's private server. The data storage layer adopts a set of different data processing formats so that it can focus on data storage. The investment data reformat in this layer and transfer to the machine learning layer. The customer/user's data analyse based on the attributes. The data classified based on the demographic and personality traits in this section.

5.3   Machine Learning Layer

Clustering and Factor Analysis is the most important part of this layer, which performs data analysis through data mining and machine learning algorithms. This layer takes data from the data storage layer and transfers classified factors to the ANFIS layer. The data entered the machine learning layer after data collection and integration into the storage space for data analysis. In the proposed layer, the attributes classification used to automatically classify the customer types and to identify customer investment indicators.
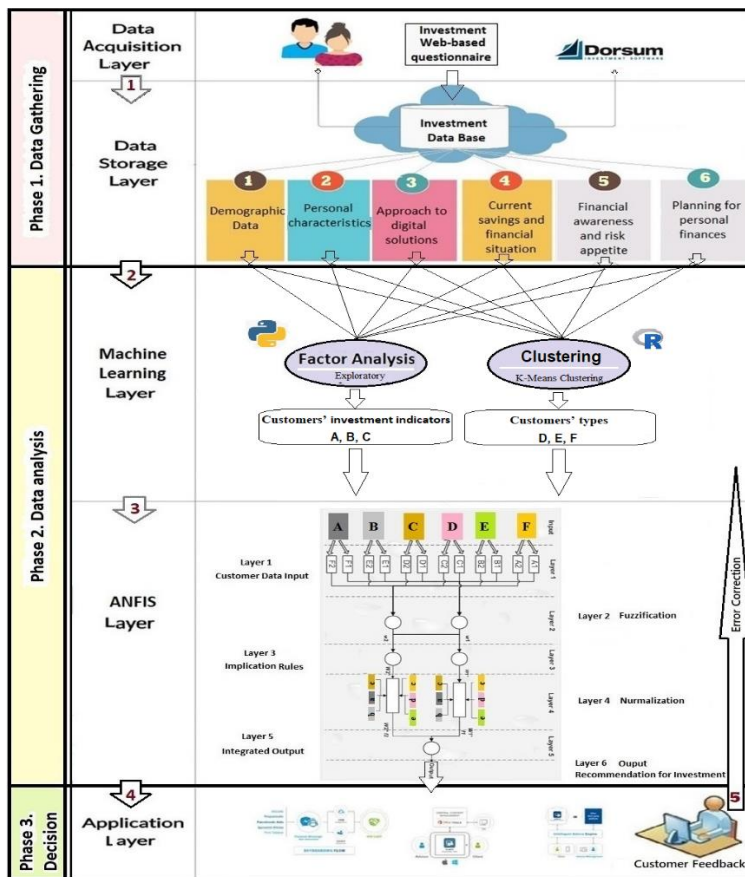
Figure 2: A Novel Business Recommender System Model Using Customer Investment Service Feedback

There are two basic functions in this layer included clustering and the factor analysis for the variables from the previous layer. We clustered due7 to its wide range of features and the ability to compare output in different ways, good guide, efficient graphical interface, compatibility with Windows, and having a comprehensive reference. For the factor analysis the system uses Python. The next function of this layer uses the exploratory factor analysis to indicate customer investment indicators. In this method, the researcher tries to discover the infrastructural structure of a relatively large set of variables. At this stage, there is no initial theory. The researcher must identify and discover the factors involved in the customer's investment that may be hidden. This method is used to summarize a set of variables and new factors

identify and introduce based on the correlation between the variables. These factors are prepared to be transferred to the next layer ANFIS for the next function. This exploratory analysis can be structuring, modelling, or hypothesizing.

5.4   ANFIS Layer

In this layer, the Sugeno fuzzy model will use for the ANFIS system. This system includes five layers as the following (Asemi, et al. 2019).

Layer 1: Neurons in this layer only direct the external input signals to the next layer. This layer is the first hidden layer and the fuzzy layer of the ANFIS model. Fuzzy neurons receive an input signal. Then they decide on the degree of dependence of this signal on the neural fuzzy set.

Layer 2: This layer is the fuzzy rule layer and the second hidden layer. Each neuron in this layer is associated with only one fuzzy Sugeno law.

Layer 3: This layer is the normalization layer and the third hidden layer. Each neuron in this layer receives and calculates signals from all neurons in the third layer. This is called normalized Firing Strength. This value determines to what extent the relevant law is valid for the inputs in the result.

Layer 4: This layer is the fourth hidden layer and the diffusion layer. Each neuron in this layer is related to the corresponding normalized neuron in the fourth layer. It also receives the first input signals (x1, x2...). The defuzzied neuron in this layer calculates the weight of the result of a rule.

Layer 5: This layer is called the output layer. In this output layer, the neurons of the previous stage are added together. Finally, by defuzzification, fuzzy outputs are converted to numeric outputs. There is only one neuron in this layer. The defuzzification method is the same as the Centre of Gravity.

5.5   Application Layer

In this layer, the application mainly customized to customers' needs and displays recommendations based on investment company's products and services. The application layer connected to the data analysis phase, so the end-user (customer) can access the source of investment recommendations in their investment platform using companies' applications. Applications required by the customer to use recommendations in the investment platform are in this part of the model. The system can receive the customer's feedback in this layer and the probability errors refer to the data analysis section for detection.

6   Conclusion

The research objective was to provide a new, novel business model for investment recommender systems using customer investment service feedback based on neuro-fuzzy inference solutions and customized for investment service. The research question was what business model can be designed for a recommender system ANFIS-based, to present investment recommendations based on investor types and investment indicators? To answer the question, a business model of an investment recommender system designed to support the investment process for the customers. Tarnowska et al., (2020) presented a Recommender System for Improving Customer Loyalty. The recommender system designed by Tejeda-Lorente et al (2019) relates to unique hedge funds that consider multiple factors, such as modern-day yields, historic performance, diversification by way of industry, etc. In the proposal of Hernandez et al. (2019), they present the functions of agents and an algorithm that improves the accuracy of the Recommender agent which oversees the Case-based reasoning system. The data corresponds to the user's characteristics, asset classes, profitability, interest rate, history stock market information, and financial news published in the media. Paranjape-Voditel and Umesh (2013) proposal was a recommender machine-based totally on association rule mining. The model presented in this research is based on the ANFIS system. This model is divided into three main parts: data collection from the investor, analysis of investor data and decision making. In the designed model, seven group factors are identified to implement the proposed investment system model through the customer or potential investor data set. These seven groups include: demographic data, personality traits, investor attitudes toward digital solutions, investor current financial status and savings, investor awareness of potential risks, and investor

financial plan information. In the proposed model, the initial data is collected through a web-based platform and transferred to the machine learning section, which is the first part of the data analysis section. In this section, customer investment criteria and types of customers are extracted. Then the types of investors are clustered and investment indicators are factor analyzed. The output obtained from this layer is transferred to the second part of the data analysis section. In the ANFIS layer, data is analyzed in six steps and investment proposals are extracted for each investor cluster. These suggestions are presented to the customer in the application layer using designed applications. Investor feedback is also received to improve and develop the system at this layer. The objective of this business recommender system model is to support the investment companies, individual investors, and fund managers in their decisions by suggesting the investment products and services based on the customers' needs, experiences, and traits.

References

Andajani, E. (2015). "Understanding Customer Experience Management in Retailing". Procedia - Social and Behavioral Sciences. 2nd Global Conference on Business and Social Sciences (GCBSS-2015) on "Multidisciplinary Perspectives on Management and Society", 17–18 September 2015, Bali, Indonesia. 211: 629–633. https://doi.10.1016/j.sbspro.2015.11.082

Asemi, A. & Asemi, A. (2014) Intelligent MCDM method for supplier selection under fuzzy environment. Int J Inf Sci Manag (IJISM) 12(2):33–40.

Asemi, A., Salim, S.S.B., Shahamiri, S.R., Asemi, A., & Houshangi, N. (2019). Adaptive neuro-fuzzy inference system for evaluating dysarthric automatic speech recognition (ASR) systems: a case study on MVML-based ASR. Soft Computing. 23, 3529–3544. https://doi.org/10.1007/s00500-018-3013-4

Burke, R. (2007). Hybrid web recommender systems. In: The Adaptive Web, pp. 377–408. Springer Berlin / Heidelberg

Christensen, C. M., & Bower, J. L. (1996). Customer Power, Strategic Investment, and the Failure of Leading Firms. Strategic Management Journal, 17(3), 197–218. https://doi.org/10.1002/(SICI)1097-0266(199603)17:3<197::AID-SMJ804>3.0.CO;2-U

Cole, B. (2015). What is business model innovation?. Definition from WhatIs.com. SearchCIO. Retrieved March 06, 2021, from https://searchcio.techtarget.com/definition/business-model-innovation

Customer Feedback Definition | Pendo.io Glossary. (n.d.). Pendo.Io. Retrieved April 20, 2020, from https://www.pendo.io/glossary/customer-feedback/

Gartner Says Marketers Must Focus on Helping Customers in Order to Remain Competitive Today. (n.d.). Gartner. Retrieved April 20, 2020, from https://www.gartner.com/en/newsroom/press-releases/2019-04-29-gartner-says-marketers-must-focus-on-helping-customer

Gentile, Ch, Spiller, N., & Noci, G. (2007). "How to Sustain the Customer Experience: An Overview of Experience Components that Co-create Value with the Customer". European Management Journal. 25 (5): 395–410. https://doi.10.1016/j.emj.2007.08.005

Guarino N. (1995). Formal Ontology, Conceptual Analysis and Knowledge Representation: The Role of Formal Ontology in the Information Technology, International Journal of Human-Computer Studies, Vol. 43, Nos. 5/6, pp. 625-640.

Hernández, E., Sittón, I., Rodríguez, S., Gil, A. B., & García, R. J. (2019). An Investment Recommender Multi-Agent System in Financial Technology. In M. Graña, J. M. López-Guede, O. Etxaniz, Á. Herrero, J. A. Sáez, H. Quintián, & E. Corchado (Eds.), International Joint Conference SOCO'18-CISIS'18-ICEUTE'18 (Vol. 771, pp. 3–10). Springer International Publishing. https://doi.org/10.1007/978-3-319-94120-2_1

Holzman, Ph.S. (2020). Personality. Encyclopædia Britannica, Encyclopædia Britannica, Inc. February 24, Available at https://www.britannica.com/topic/personality, Access Date: April 12, 2020

Investment service—FCA Handbook. (n.d.). Retrieved April 20, 2020, from https://www.handbook.fca.org.uk/handbook/glossary/G603.html

Investment. (2020). In Wikipedia. https://en.wikipedia.org/w/index.php?title=Investment&oldid=951351513

Investor. (n.d.). Retrieved February 25, 2021, from https://dictionary.cambridge.org/dictionary/english/investor

Jang, J.R. (1993). "ANFIS: adaptive-network-based fuzzy inference system". IEEE Transactions on Systems, Man and Cybernetics. 23 (3): 665–685. https://doi.10.1109/21.256541

Janakiraman, N., Meyer, R.J., Morales, A.C. (2006). "Spillover Effects: How Consumers Respond to Unexpected Changes in Price and Quality". Journal of Consumer Research. 33 (3): 361–369. https://doi.10.1086/508440. ISSN 0093-5301

Kanaujia, P. K. M., Manjusha, P., & Siddharth, S. R. (2017). A Framework for Development of Recommender System for Financial Data Analysis. International Journal of Information Engineering & Electronic Business, 9 (5),18-27.

Liang, T.P. (2008). Recommendation systems for decision support: An editorial introduction. Decision Support Systems. 45 (3). 385-386. https://doi.org/10.1016/j.dss.2007.05.003.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: Grouplens (1994). An open architecture for collaborative filtering of netnews. In: Proceedings ACM Conference on Computer-Supported Cooperative Work, pp. 175–186

Resnick, P., Varian, H.R. (1997). Recommender systems. Communications of the ACM 40(3), 56–58

Sarkar, D., Bali, R., Sharma, T. (2018) Machine Learning Basics. In: Practical Machine Learning with Python. Apress, Berkeley, CA. Available at: https://doi.org/10.1007/978-1-4842-3207-14

Slovic, P. (1972). Psychological Study of Human Judgment: Implications for Investment Decision Making. The Journal of Finance, 27(4), 779–799. https://doi.org/10.2307/2978668

Tejeda-Lorente, Á., Bernabé-Moreno, J., Herce-Zelaya, J., Porcel, C., & Herrera-Viedma, E. (2019). A risk-aware fuzzy linguistic knowledge-based recommender system for hedge funds. Procedia Computer Science, 162, 916.

van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P. C. (2010). Customer Engagement Behavior: Theoretical Foundations and Research Directions. Journal of Service Research, 13(3), 253–266. https://doi.org/10.1177/1094670510375599

Verhoef, P.C.; Lemon, K.N., Parasuraman, A.; Roggeveen, A., Tsiros, M., Schlesinger, L.A. (2009). "Customer Experience Creation: Determinants, Dynamics and Management Strategies". Journal of Retailing. Enhancing the Retail Customer Experience. 85 (1): 31–41. https://doi.10.1016/j.jretai.2008.11.001

SOLOPRENEUR DIGITAL ECOSYSTEMS: GENESIS, LINEAGE AND PRELIMINARY CATEGORIZATION

GABRIELE PICCOLI,[1] BIAGIO PALESE[2] & JOAQUIN RODRIGUEZ[3]

[1] Louisiana State University and University of Pavia, Business Education Complex, Baton Rouge, Los Angeles 70803, United States of America; e-mail: gpiccoli@cct.lsu.edu
[2] Northern Illinois University, College of Business, DeKalb, Illinois 60115, United States of America; e-mail: bpalese@niu.edu
[3] Grenoble Ecole de Management, 12 rue Pierre Sémard, 38000 Grenoble, France; e-mail: joaquin.rodriguez@grenoble-em.com

Abstract This paper traces the genesis and lineage of solopreneur digital ecosystems. These ecosystems, fostered by a digital environment that is infrastructural, combinatorial and servitized, are enabling the rise to prominence of the solopreneur. We theorize solopreneur digital ecosystems as the latest incarnation of systems beyond firm control, with digital platforms and digital marketplaces as their principal enablers. In an effort to compare them from the perspective of the solopreneur, we categorize solopreneur digital ecosystems on three dimensions: algorithmic control, commoditization, and lock-in. Our work contributes a framework that solopreneurs can use to identify ecosystems in which they can optimally invest their talents and scarce resources. We discuss the findings of this mapping and draw implications for research and practice.

1   Introduction

The prediction that "by reducing the costs of coordination, information technology will lead to an overall shift toward proportionately more use of markets-rather than hierarchies-to coordinate economic activity" (Malone et al 1987 p. 484) has proven true. What the prediction did not foreshadow, however, was that such markets would also be digital. As software continues to "eat the world" (Andreessen 2011), digital ecosystems have dramatically impacted business strategies and society more generally. One such impact is on individual's revenue generating activities and the growing opportunity to unbundle work from employment. As automation and machine learning advancements threaten a larger swath of traditional jobs (Manyika et al 2017), there is an unprecedented opportunity for human talent to be unleashed in digital ecosystems that enable demand and offer matching at a never-before-seen scale (Jin 2020).

According to the *Oxford Dictionary*, a solopreneur is »a person who sets up and runs a business on their own«. Solopreneurs, of course, are not a novelty of the digital age. But interest in solopreneurship has grown noticeably in the last two decades[1] due to the rise of what we call solopreneur digital ecosystems (SDE). SDEs promise to simplify access to (self-)employment, particularly in disadvantaged or marginalized communities, by providing digital tools enabling solopreneurs to organize their work and overcome some of the barriers to employment they usually face (Dillahunt and Malone 2015). However, selection of a SDE is a critical early decision by solopreneurs seeking to maximize the return on their invested time and talents. This article takes the solopreneur perspective, investigating how digital enablers of different SDEs affect solopreneurs' strategic options. It contributes to a cohesive research agenda centered on SDE by providing an early categorization of their digital enablers and a framework to evaluate competing SDE. It also draws implications for research and practice based on the categorization of the ecosystems.

---

[1] The term did not appear until 1996, according to the Google Books ngram viewer, and grew more than 63-fold between its advent and 2019.

G. Piccoli, B. Palese &J. Rodriguez:
Solopreneur Digital Ecosystems: Genesis, Lineage and Preliminary Categorization

241

2   Theoretical foundations

Before the advent of the Internet, organizational information systems were fully controlled by the firm that built, purchased or commissioned them to an outsourcer to operate on the firm's behalf. But the internet and the services built on top of it ushered in an era of systems beyond firm control. Systems beyond firm control are information systems neither designed nor commissioned by a firm that the organization must use to compete (Palese and Piccoli 2018). They are socio-technical arrangements that enable transactions and value exchanges. Online review systems, a prototypical example, have transformed how travelers search and share information. Their features and functionalities are not designed or controlled by travel companies, yet hoteliers, restaurateurs, and other travel operators cannot ignore dominant ones (e.g., TripAdvisor, Booking).

Systems beyond firm control are ecosystems: groups of interacting and interdependent entities and their environments. SDEs are a special kind of ecosystem in which the bulk of participants are solopreneurs serving end-consumers in an infrastructural, combinatorial and servitized environment that provides digital enablers (Piccoli et al 2020). Those digital enablers are digital platforms and digital marketplaces that allow the solopreneurs to organize and commercialize their work without formally joining a company in a traditional employment working arrangement. Digital enablers are the novel instruments solopreneurs use to create their products/services and/or the organizational processes to manage and run a business operation independently.

The definition of SDE advanced here is broad enough to encompass gig economy workers (e.g., Uber), social media influencers (e.g., Instagram) and digital creators (e.g., Twitch). It is in line with recent research that identifies "platform ecosystems" as "semi-regulated marketplaces that foster entrepreneurial action under the coordination and direction of the platform sponsor, or as multisided markets enabling transactions among distinct groups of users" (Jacobides et al 2018 p. 2258).

To categorize the variety of SDE, it is important to recognize the characteristics of the digital enablers that make them possible. To the best of our knowledge, the literature lacks such categorization, which makes it difficult to evaluate the inevitable trade-offs between competing SDEs and offer reliable guidance for maximizing

cospecialized investments in competing ecosystems. We introduce and define the major categories of digital enablers that shape modern SDE.

2.1 Digital Marketplaces

Digital marketplaces are the digital spaces in which buyers and sellers "exchange product information, coordinate, and transact" (Pavlou and Gefen 2004 p. 40). They exercise control over the products and services listed by sellers (Eaton et al 2015). Thus, digital marketplaces differentiate themselves by exercising control over the type, characteristics, number, and quality of products and services offered by sellers. They enable trust between buyers and sellers by guaranteeing levels of customer protection (e.g., refunds), reliable payment transfers (e.g., escrow services), and mechanisms to ensure that transactions are based on accurate and reliable information (e.g., ratings) (Pavlou and Gefen 2004. Finally, they facilitate discovery of products and services by customers (Li et al 2018). Digital marketplaces with a significant number of suppliers incur high search costs and seek to reduce costs by implementing tools that enable customers to easily find products and services of interest.

2.2 Digital Platforms

In line with recent literature, we define digital platforms as evolving sociotechnical systems with modular design architecture that expose digital resources module designers use to produce innovations (Constantinides et al 2018). We define digital resources as a specific class of digital objects (Faulkner and Runde 2019) that a) are modular; b) encapsulate objects of value, namely specific assets and/or capabilities; c) and are accessible by way of a programmatic bitstring interface (Piccoli et al 2020). By specifying a modular architecture and exposing digital resources, digital platforms enable the creation of new modules (i.e., complementors) that extend their functionality. More importantly, they offer combinatorial and servitized resources solopreneurs can leverage to build innovative products and services (i.e., vertical platforms) and devise new business models (i.e., horizontal platforms). Depending on their architecture, characteristics, and variety of digital resources exposed, digital platforms engender different levels of generativity (Zittrain 2006). For example, a highly generative platform like Roblox enables solopreneurs to create an infinite variety of games and applications for the Roblox "metaverse." On the other end of

G. Piccoli, B. Palese &J. Rodriguez:
Solopreneur Digital Ecosystems: Genesis, Lineage and Preliminary Categorization

243

the spectrum (i.e., the extensibility of the platform is zero), the platform becomes a tool. For example, Substack provides a number of resources for solopreneurs in newsletter and podcasting spaces. These individuals can create products (e.g., a newsletter) and run businesses (e.g., manage mailing lists, collect payment) by configuring/using instruments made available by Substack. However, Substack does not expose interfaces that enable complementors to contribute new modules or enable solopreneurs to extend the functionality of existing modules.

2.3 Integrated Platforms and Marketplaces

While early research treated platforms and marketplaces as interchangeable constructs (Rochet and Tirole, 2003), more recent work has articulated the difference between the two (Benlian et al 2015). An increasing number of organizations purposefully integrate and simultaneously manage a digital platform and a digital marketplace (Ghazawneh and Henfridsson 2015). By doing so they concurrently control marketplace and platform functionalities. The integration of the two yields unprecedented power through the simultaneous control of the products/services (via platform ownership) and distribution and monetization channel (via marketplace ownership). In this type of SDE, solopreneurs are required to abide by both a prespecified product or business architecture and marketplace governance rules enforced by the owner of digital enablers at the center of the ecosystem. Examples include ecosystems anchored by such firms as Amazon in retail, Spotify in podcasting, or Deliveroo in food delivery. They represent the latest examples in the evolution of systems beyond firm control.

3 Solopreneur digital ecosystems as algorithmic economies

Solopreneur digital ecosystems are characterized by resources made available by digital platform and digital marketplace owners. Those resources enable transactions between solopreneurs and consumers as well as creation of solopreneurs' products or services. Thus, SDEs become "algorithmic economies" in which decision-making coordination and control functions are embedded in the ecosystem's digital enablers' algorithms (Möhlmann et al 2020). Consider, for example, product visibility in physical retail. It is characterized by limited shelf space, managed through ad hoc contractual agreements, and has stable underlying performance drivers. Conversely, digital shelves in a digital marketplace are theoretically unlimited, and product

visibility is determined by evolving algorithms that operate in real time and are often proprietary and inscrutable. This distinctive characteristic of algorithmic economies represents an added layer of complexity solopreneurs need to manage when joining a digital ecosystem.

4   Preliminary classification dimensions

Keeping with our focus on investigating how the digital enablers of different SDEs affect solopreneurs' strategic options, we advance a preliminary categorization of SDEs based on analysis of the primary digital enablers of each ecosystem. These enablers expose digital, IT and complementary resources (Piccoli and Ives 2005, Piccoli et al 2020) supporting solopreneurs in the development and/or commercialization of their innovations (i.e., products and services). Solopreneurs in turn orchestrate a purposefully arranged set of resources in pursuit of their goals. Such goals are typically commercial, measured in revenue and profits, but can also be personal (e.g., self-actualization, validation). The focus on solopreneurs as the primary beneficiaries of our work requires a categorization that, while concentrating on digital enablers as the unit of analysis, is designed to be practical for solopreneurs deciding in which ecosystems to optimally invest their talents and scarce resources. Specifically, we adopt the following three dimensions:

- Algorithmic control. This dimension captures the automatic enforcement of control mechanisms through algorithms implemented in software programs (Möhlmann et al 2020). It determines the degrees of freedom solopreneurs can exert as they operate within the ecosystem. It includes control over the product or service specifications (e.g., Uber's eligibility for Uber Black as a rating above 4.85), control over the manner in which work is organized (e.g., Uber's expectation that riders accept a ride withing 15-30 seconds), and control over the solopreneur's relationships with customers and the visibility of their offerings (e.g., Uber's algorithmic matching of riders to drivers).
- Commoditization. Commoditization stems from the design features adopted by digital enablers. While the fungibility of what the solopreneurs produce is an important consideration, with highly fungible solopreneur offerings being more substitutable, commoditization is a function of the resources that the enablers expose to solopreneurs and the functionalities

available to users. Consider the example of Instacart, the grocery delivery marketplace, and Dumpling, a competing digital platform designed to offer "everything you need to start, run, and grow your own personal shopping business." Instacart personal shoppers are entirely fungible, since there are no features in Instacart to request a specific shopper and all interactions between the shopper and the customer are managed within the app. Conversely, Dumpling's design is geared toward enabling the shopper to develop and maintain a base of recurring customers who trust her. Over time, personal shoppers on Dumpling become non-fungible to their loyal customers.

- Creator lock-in. Lock-in is a function of switching costs, defined as the current value of all the tangible and intangible co-specialized investments the solopreneur has made in the ecosystem (Piccoli and Ives 2005). The higher the switching costs, the more difficult it is for solopreneurs to continue operating when migrating to a competing ecosystem. Uber drivers lose their driving history and reputation score if they migrate to a competitor (e.g., Lyft). Since history and reputation are critical input to the matching algorithm, or the ability to offer premium services (e.g., Uber Black), lock-in is substantial. Conversely, while a writer migrating from Substack needs to learn how to operate her newsletter in the competing ecosystem (e.g., Revue), Substack writers own their mailing list and payment relationship with subscribers (i.e., switching costs are relatively low). However, "even when switching costs appear low, they can be critical for strategy" (Shapiro and Varian 1999, p. 108), with the critical element being "not the absolute magnitude of the cost of switching, but its size relative to the value received from the [platform resources]" (Piccoli and Ives, 2005 p. 762).

5   Data and Results

While the three dimensions are clearly related, they capture different aspects of the decision-making space solopreneurs must investigate when deciding which ecosystem to select. We screened 200 digital enablers, evaluating the Gross Merchandise Value (GMV), number of active solopreneurs and users in their ecosystems.[2] We selected the top 10 by GMV, by number of users and number of

---

[2] We used the list on Sidehustlestack (https://sidehustlestack.co/) as the starting point of our selection.

solopreneurs. We classified the resulting 30 distinct SDEs[3] by type of primary digital enabler (i.e., digital platforms, marketplaces and integrated platforms and marketplaces) and rated them on a scale from 1 (lowest) to 5 (highest) for algorithmic control, commoditization and lock-in.[4] The top 30 solopreneur digital ecosystems are anchored by 4 digital platforms, 14 digital marketplaces and 12 integrated platforms and marketplaces. We detected significant variability across the three dimensions of algorithmic control, commoditization and lock-in (Table).

Table 1: Descriptive Statistics by Digital Enabler category

	Digital Platforms	**Marketplaces**	**Integrated Platforms and Marketplaces**
Number	4	14	12
Algorithmic control (mean)	2.25	2.64	2.83
Algorithmic control (sd)	0.50	1.08	0.84
Commoditization (mean)	2.25	3.29	3.17
Commoditization (sd)	0.50	0.83	1.19
Lock-in(mean)	2.75	3.14	3.33
Lock-in(sd)	0.50	0.95	1.07

These results indicate that the characteristics of solopreneur digital ecosystems heavily depend on the design choices of their primary digital enablers rather than on uncontrollable or inherent characteristics of the ecosystems. In other words, the solopreneurs compete within an algorithmic economy they can perhaps influence, but certainly cannot control (i.e., a system beyond firm control). Instead control of the system rests with the firms that design, manage and own the primary digital enablers. Our results also show significant differences across the three types of digital enablers, with marketplaces and integrated platforms and marketplaces showing similar patterns, while platforms diverge. This preliminary observation may indicate some mimetic tendencies by competing digital enablers. Alternatively, it may

---

[3] If a platform was among the top 10 in multiple lists, we included it only once. On each list we continued selection until we classified 10 digital enablers of solopreneur digital ecosystems (e.g., Uber was top 10 in all 3 dimensions, but we included it only in the top 10 by GMV).

[4] Following Krippendorff (2018) we provided a classification procedure to two independent coders (available upon request to the authors) and computed inter-rater reliability. We recorded a kappa value of 0.93. In the second stage, a third coder reviewed only the SDEs without full agreement. Those SDEs were discussed among the coders in a consolidation meeting that lead to full agreement.

be that the type of digital enabler constrains, at least in part, its owner's design choices (see discussion).

Table 2: GMV, users and solopreneurs statistics by ecosystems type

Ecosystem type	Digital Platforms	Digital Marketplaces	Integrated Platforms and Marketplaces
GMV (mean)	456,750,000	507,092,857	1,562,829,667
Users (mean)	175,750,000	98,838,182	24,740,000
Solopreneurs (mean)	366,667	1,780,000	1,533,917

SDEs anchored by firms that integrate a platform and a marketplace in our sample have the highest average GMV (Table), about three times higher than an SDE anchored by either a digital marketplace or a digital platform. This result shows the power of integrating the two enablers, likely stemming from their ability to control resources underpinning solopreneurs' products/services and their transactions with consumers. It appears that the most successful firms that integrate both a platform and a marketplace enable superior value propositions and successfully aggregate customers demand, resulting in higher GMV.

Marketplaces, be it as a standalone enabler or when integrated with a platform, attract, on average, a larger number of solopreneurs, with digital platforms only reaching a fifth of the other two types of digital enablers. This result may depend on the draw and incentives SDEs anchored by a digital marketplace create for solopreneurs. As marketplaces provide direct access to customers, it is a simple strategic decision to join – but simple may not imply advantageous (see discussion).

The above argument leads to the expectation that marketplaces also dominate in number of users. The opposite is true in our sample, which suggests marketplaces only draw a subset of consumers in most markets.[5] Conversely, lacking a marketplace, digital platforms focus on providing tools to reach and serve all consumers in an addressable market. When successful, digital platforms empower solopreneurs to serve customers both directly and across marketplaces – resulting in successful digital platforms attracting twice as many customers as the average marketplace and seven times as much as the average integrated platform and

---

[5] Amazon, widely seen as a monopolist in the US, only controls about 35% of ecommerce transactions by value and only 6% of all retail.

marketplace. Charting distributions of SDEs across the three dimensions shows how most combinations are completely absent (87.2%), with 21 of the 30 observations concentrated in only 7 combinations. While the low number of ecosystems codified in this preliminary study is likely responsible for these results, the observation points to some converge toward the dominant designs. Moreover, stricter configurations (e.g., 4, 5, 5) appear viable only when the digital enabler encompasses a marketplace.

6   Discussion and Conclusions

Taken together, our results suggest that while the draw of marketplaces may be inescapable, solopreneurs must be weary of their power. The limited number of digital platforms we categorized, compared to the other two types of digital enablers, suggests that control of a marketplace contributes to growth in terms of GMV and the ability to attract solopreneurs. However, digital platforms dominate in the number of users in the ecosystem. In other words, controlling a marketplace helps firms that own the primary digital enabler of a solopreneur digital ecosystem to also serve as the catalyst for supply. We speculate this feature attracts solopreneurs because it simplifies their "route to market." However, such simplification comes at a cost, making the solopreneurs more dependent on the firm that owns critical marketplace resources since they mediate the solopreneur's ability to develop a digital relationship with customers.

The above result is corroborated by the average score of each digital enabler type on algorithmic control, commoditization and lock-in. Digital enablers that include a marketplace are fairly consistent, with scores that exceed digital platforms by about 30%. We ascribe this result to the control marketplaces exert over solopreneurs' commercialization practices. Moreover, there appears to be a positive correlation between algorithmic control and commoditization, which hints at the need to standardize the signals and variables used for representing solopreneurs' creations. In other words, there may be an implicit commoditization pressure of solopreneurs' creations, even when not purposedly designed by digital enabler owners who, in search of efficiency through algorithmic control, seek to enforce standards in categorization and evaluation of offers. Solopreneurs in these digital ecosystems compete for visibility with only a limited understanding of the algorithm's inner workings and the casual paths that govern the relationship between actions and results. The algorithms can rapidly, continuously, and comprehensively evaluate

solopreneurs' products and services, resulting in an economy that is hypercompetitive and commoditizing toward suppliers (Möhlmann et al 2020).

The above pressures are not present in digital platforms, which lack marketplace control. The search for algorithmic efficiency is focused on work processes that enable the solopreneur and result in a low commoditization score. Digital platforms may inherently spur differentiation and innovation in product/service and business model/operations. We are convinced that design choices by the digital enabler owners are critical; however, preliminary results point to structural differences between the types of enablers.

We summarize our analysis in the following 2x2 matrix, mapping the depth and breadth of commercialization services and of product and/or operations support digital enablers provide (Figure).
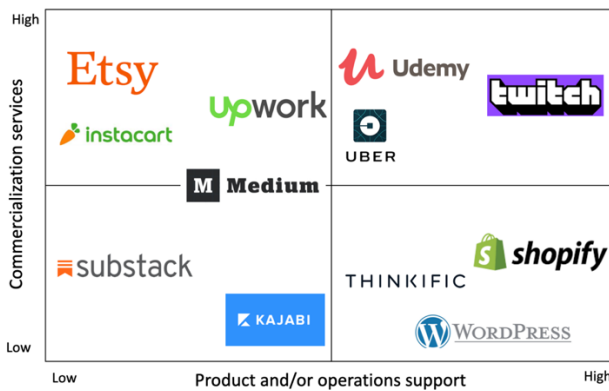
Figure 2: Dimensions of Digital Enablement in Solopreneur Digital Ecosystems
Source: authors

The matrix captures the type of analysis our work offers to solopreneurs seeking to strategically invest their talents and scarce resources. On the one hand, deeper support generally speeds up product/service creation, market access and transaction completion. But when leveraging a wider array of digital resources offered by the enablers, the solopreneur must carefully analyze the design to evaluate the resulting degree of algorithmic control, commoditization and lock-in. As an illustration, we mapped the 10 most representative SDE in our sample. We hope that, despite its

limitations, our preliminary work will inspire future research focused on helping the emerging middle class of solopreneurs (Jin 2020) to take advantage of the increasing opportunities to unbundle work from employment in solopreneur digital ecosystems.

References

Andreessen, M. (2011). "Why Software Is Eating the World," Wall Street Journal. (https://www.wsj.com/articles/SB10001424053111903480904576512250915629460, accessed July 15, 2018).

Benlian, A., Hilkert, D., & Hess, T. (2015). How open is this Platform? The Meaning and Measurement of Platform Openness from the Complementers' Perspective. Journal of Information Technology, 30(3), 209–228. https://doi.org/10.1057/jit.2015.6

Constantinides, P., Henfridsson, O., and Parker, G. G. (2018). Platforms and Infrastructures in the Digital Age, Information Systems Research (29:2), pp. 381–400. (https://doi.org/10.1287/isre.2018.0794).

Dillahunt, T. R., and Malone, A. R. 2015. "The Promise of the Sharing Economy among Disadvantaged Communities," in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, New York, NY, USA: Association for Computing Machinery, April 18, pp. 2285– 2294.

Eaton, B., Elaluf-Calderwood, S., and Sorensen, C. (2015). "Distributed Tuning of Boundary Resources: The Case of Apple's IOS Service System," MIS Quarterly (39:1), 217+

Faulkner, P., and Runde, J. (2019). Theorizing the Digital Object, MIS Quarterly (43:4).

Jacobides, M. G., Cennamo, C., and Gawer, A. (2018). "Towards a Theory of Ecosystems," Strategic Management Journal (39:8), pp. 2255–2276. (https://doi.org/10.1002/smj.2904).

Jin, L. (2020) The Creator Economy Needs a Middle Class. Harvard Business Review.

Jin, L. (2020) Unbundling Work from Employment. Substack. Available at: https://li.substack.com/p/unbundling-work-from-employment

Li, L., Chen, J., & Raghunathan, S. (2018). Recommender system rethink: Implications for an electronic marketplace with competing manufacturers. Information Systems Research, 29(4), 1003-1023.

Malone, T., Yates, J., and Benjamin, R. (1987). Electronic markets and electronic hierarchies. Communications of the ACM 30 (6), 484-497

Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., Dewhurst, M. (2017). A future that works: Automation, employment, and productivity. Mckinsey. Available at: https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Digital%20Disrupti on/Harnessing%20automation%20for%20a%20future%20that%20works/MGI-A-future-that-works_Full-report.pdf

Möhlmann, M., Zalmanson, L., Henfridsson, O., & Gregory, R. W. (2020). Algorithmic management of work on online labor platforms: when matching meets control. MIS Quarterly: Management Information Systems, 1-54.

Palese, B., and Piccoli, G. (2018). Effective Use of Systems Beyond the Firm's Control: The Case of Online Review Systems, In Proceedings of the 39th International Conference of Information Systems, San Francisco, USA

Pavlou, P. and Gefen, D. (2004.) Building effective online marketplaces with institution-based trust Information Systems Research, 15 (1) (2004), pp. 35-62

Piccoli, G., and Ives, B. (2005). II-Dependent Strategic Initiatives and Sustained Competitive Advantage: A Review and Synthesis of the Literature, MIS Quarterly (29:4), pp. 747–776.

G. Piccoli, B. Palese &J. Rodriguez:
Solopreneur Digital Ecosystems: Genesis, Lineage and Preliminary Categorization

251

Piccoli, G., Rodriguez, J, and Grover, V. (2020) Strategic Initiatives and Digital Resources: Construct Definition and Future Research Directions. , In Proceedings of the 41st International Conference of Information Systems, India

Rodriguez, Joaquin, and Gabriele Piccoli. "Competing within Aggregators: Competitive Moves in the Deliveroo Online Delivery Platform." (2020).

Rochet JC, Tirole J (2003) Platform competition in two-sided markets. Journal of the European Economic Association 1(4):990–1029

Shapiro, C., and Varian, H. R. (1999). The art of standards wars. California management review, 41(2), 8-32.

Tapscott, D. (2001). Rethinking strategy in a networked world [or why Michael Porter is wrong about the internet]. strategy and business, 34-41.

Zittrain, J. (2006). The Generative Internet. Harvard Law Review, Vol. 119, p. 1974, Available at SSRN: https://ssrn.com/abstract=847124

DEMOGRAPHIC DIFFERENCES IN THE EFFECTIVENESS OF A PHYSICAL ACTIVITY APPLICATION TO PROMOTE PHYSICAL ACTIVITY: STUDY AMONG AGED PEOPLE

TUOMAS KARI,[1] MARKUS MAKKONEN[1] & LAURI FRANK[2]

[1] Institute for Advanced Management Systems Research, Turku, Finland, [2] University of Jyvaskyla, Jyvaskyla, Finland; e-mail: tuomas.t.kari@jyu.fi, markus.v.makkonen@jyu.fi
[2] University of Jyvaskyla, Faculty of Information Technology, Jyvaskyla, Finland; e-mail: lauri.frank@jyu.fi

Abstract The global population is ageing and simultaneously the life expectancy at older ages is improving. To support healthy and active aging, it is imperative to find solutions to support physical activity (PA) in older age. Digital wellness technologies are a potential solution, but in order for such technologies to be successful, research is needed to gain a better understanding on their use and effectiveness among aged people. To address this need, this study investigated the effectiveness of a physical activity application to promote PA behavior among aged people of different demographics (gender, age, education, marital status). PA levels were measured before taking the application into use and after 12 months of use. The results suggest that a physical activity application can be effective in promoting PA behavior among aged people as there was a notable and a statistically significant increase in walking and total PA levels between baseline and 12-month follow-up. Regarding the demographic differences, there were very few differences in the changes in PA levels between different demographics, suggesting the effectiveness is not subject to the demographic background of the user.

1   Introduction

The global population is ageing and practically all countries are experiencing growth
in the proportion of their older population. Globally, the number of people aged
over 65 years is projected to double by 2050, and at the same time, the life expectancy
at older ages is improving (United Nations 2019). This makes healthcare and policy
makers concerned, and it is imperative to find more solutions to support PA in older
age. PA has significant health benefits and contributes to the prevention of non-
communicable diseases across all age groups, but also aids in upkeeping the ability
to function when a person gets older and is vital to ward off frailty and age-related
illness (Hoogendijk et al. 2019). The World Health Organization (WHO) as well as
several national health institutes provide research-based PA guidelines. Yet,
insufficient PA is a major global problem across all age groups (WHO 2020). For
example, the WHO (2020) recommends that people aged 65 years and older should
do at least 150–300 minutes of moderate PA or at least 75–150 minutes of vigorous
PA per week, or an equivalent combination. Balance and muscle-strength exercises
should also be conducted regularly. Additional health benefits can be reached with
more PA. In Finland, where our study was conducted, only around one fourth of
the people aged 60 years and older meet these guidelines (THL 2019). Thus,
innovative solutions to promote PA among aged people are urgently needed.

Digital wellness technologies, that is, "digital technologies that can be used to support
different aspects of wellness" (Kari et al. 2021), represent a potential solution. There
are various PA devices, applications, and services, the use of which has become
increasingly common for diverse types of users with different PA levels (Kettunen
et al. 2017; Moilanen et al. 2014). Their potential to promote PA behaviors among
aged people has been suggested, but more research on their effectiveness is called
for (e.g., Carlsson & Walden 2017; Larsen et al. 2019; Seifert et al. 2018). To address
this, our study investigates the following research question: *How effectively can a physical
activity application promote PA behavior among aged people of different demographics*? For PA,
we follow the definitions by WHO (2020). The investigated demographics are
gender, age, education, and marital status. By addressing this, we contribute to the
IS stream of research on the ability of digital wellness technologies to influence
behavior change and also provide insights on how they could support changes in PA
behavior among aged people. The study is part of an ongoing research program in
which aged people participants take into use a mobile physical activity application.

2   Digital Wellness Technologies Among Aged People

A key reason for using digital wellness technologies and physical activity applications is often the expected positive effects on PA behavior and health. When using digital wellness technologies, the users are receptive to potential changes in their PA behavior. As Oinas-Kukkonen (2013, p. 1225) notes, "information technology always influences people's attitudes and behaviors in one way or another", either intendedly or unintendedly. There is a plethora of information systems that can be used to persuade (either positive or negative) behavior changes (Oinas-Kukkonen 2013). Such systems are often discussed in the literature using the terms persuasive systems (e.g., Oduor & Oinas-Kukkonen 2021) or persuasive technology (e.g., Fogg 2003). On this line, Oinas-Kukkonen (2013) presented the concept of behavior change support system (BCSS), defined as "a socio-technical information system with psychological and behavioral outcomes designed to form, alter or reinforce attitudes, behaviors or an act of complying without using coercion or deception".

Digital wellness technologies (e.g., physical activity applications) can act as BCSSs in several ways, as by using such technologies, the users are exposed to different user experiences that can act as drivers for future PA behaviors (e.g., Kari et al. 2016a; Karppinen et al. 2016). Hence, using such technologies can potentially change and support PA behaviors in various ways. For example, digital wellness technologies can be used to increase PA levels (e.g., Larsen et al. 2019; Romeo et al. 2019), and to reduce sedentary behaviors (e.g., Stephenson et al. 2017). They can support goal-setting (e.g., Gordon et al. 2019; Kirwan et al. 2013), and some advanced products also include digital coaching features (e.g., Kari & Rinne 2018; Kettunen et al. 2020; Schmidt et al. 2015). Features for social support are another benefit (e.g., Sullivan & Lachman 2017). They can also make PA more compelling via gamification (Kari et al. 2016b; Koivisto & Hamari 2019) or exergaming (e.g., Kari 2014; Kari et al. 2019; Loos & Zonneweld 2016). Moreover, digital wellness technologies can provide different feedback, which can increase the user's awareness of PA and motivate to improve related behavior (e.g., Wang et al. 2016). However, the "learning-effect" resulting from the increased awareness may also lead to use discontinuance (Kari et al. 2017, p. 285). One should also note that sometimes users face negative and harmful experiences with these technologies (Rockmann 2019). The perceived compatibility is also important (Makkonen et al. 2012).

Digital wellness technologies have been found promising in terms of promoting wellness and PA behavior among aged people, however, questions have been posed regarding their effectiveness. In studies focusing on aged people, Changizi and Kaveh (2017) found that mobile health solutions can improve PA, self-management, and self-efficacy among other things. Muellmann et al. (2018) found interventions utilizing eHealth solutions to be effective in promoting PA at least on short-term. Similarly, Yerrakalva et al. (2019) found that mobile health application-based interventions are potential in promoting PA and reducing sedentary time on short term. Larsen et al. (2019) found low-quality evidence for a moderate effect of physical activity monitor-based interventions on PA. Stockwell et al. (2019) found that digital behavior change interventions may increase PA and ability to function. A common conclusion in these studies is that more high-quality studies are required.

3   Methodology

To investigate how effectively can a physical activity application promote PA behavior among aged people of different demographics (gender, age, education, marital status), we examined the changes in PA levels between baseline (before taking the application into use) and follow-up (after 12 months of use). Thus, the participants of this study consist of those partaking in the research program and using the application for a whole 12 months or longer.

3.1   The Physical Activity Application Used in the Study

The application was developed for the target group in the ongoing research program. The application operates on the Wellmo application platform (Wellmo 2019), where its features constitute their own entity. Wellmo supports iOS and Android operating systems. The central features are related to tracking PA, including for example, logging conducted PA and following it through weekly, monthly, and yearly reports. User can also import data from external PA services supported by Wellmo platform.

3.2   Research Setting, Data Collection, and Analysis

The first field groups in the research program started in June 2019, after which new groups have started continuously. The study was conducted in Finland, and the field groups (i.e., participants) were recruited via the Finnish pensioners' associations. No limits beyond age were set for partaking, meaning that the baseline PA level could be anything from very low to very high. Each field group was assigned a field researcher who guided the participants in taking the application into use and using it. The participants used the application in daily life and conducted PA according to their own preferences, meaning they were not provided with any specific PA programs to follow or goals to reach out for, but instead could freely conduct PA how and when they preferred. The application and its use were free of charge for the participants, but they were required to have their own smartphone.

The data on PA levels were collected with surveys and measured as a self-report by using the IPAQ-E (Hurtig-Wennlöf et al. 2010). The IPAQ-E is a modified version of the short-format IPAQ (Craig et al. 2003; Ekelund et al. 2006), and it has been culturally adapted and validated for the elderly (Hurtig-Wennlöf et al. 2010). The IPAQ and IPAQ-E are designed to provide a set of well-developed instruments that can be used to obtain comparable estimates of PA (IPAQ group 2005a). The IPAQ is the most widely used and validated PA questionnaire (Lee et al. 2011; van Poppel et al. 2010). The IPAQ-E focuses on collecting self-reported PA data on sitting time, walking, moderate PA, and vigorous PA from the period of last seven days. For this study, the IPAQ-E questionnaire was translated from Swedish to Finnish following the wording of the Finnish short-format IPAQ. As there were Finnish and Swedish (both are official languages in Finland) speaking participants, both language versions were used. The baseline data was collected with printed questionnaires in organized meetings. The participants were given oral and written instructions on answering. The follow-up data was collected with an online questionnaire, each participant receiving a survey invitation link via email. The local ethical committee was contacted before the start of the research program deeming that no separate approval was needed for this study. All participants gave a written informed consent.

Before the analysis, the *Guidelines for the data processing and analysis of the International Physical Activity Questionnaire* (IPAQ group, 2005b) and the guidelines presented by Hurtig-Wennlöf et al. (2010) were followed. First, before the analysis, the standard methods for the cleaning and treatment of IPAQ datasets were conducted (IPAQ group 2005b, p. 10-11). Second, the results are presented in *median minutes per week* rather than means, together with the interquartile ranges (25th–75th percentile) (IPAQ group 2005b); and as advised by Hurtig-Wennlöf et al. (2010, p. 1853), the results are presented in *time in minutes spent in different intensities* instead of converting into metabolic equivalent of task values (MET) and MET-minutes, except for the *total PA* which is presented in MET-min per week. Total PA was calculated using the formula: 3.3*walking min + 4.0*moderate min + 8.0*vigorous min (IPAQ group 2005b). However, as Hurtig-Wennlöf et al. (2010) note, the used weighting factors correspond to activity-specific MET values in adults and might not be appropriate for older people (e.g., moderate intensity level in the elderly is likely lower than the same intensity level in younger adults). Yet, they assumedly "can still reflect the proportions of PA intensities and are therefore useful for ranking participants with regard to PA" (Hurtig-Wennlöf et al. 2010, p. 1853).

The analysis was conducted with the IBM SPSS Statistics 26 software. The changes in sitting and different types of PA behavior (walking, moderate, vigorous) as well as total PA were analyzed by comparing the PA levels between the baseline and the follow-up. The statistical significance of the overall changes in PA levels were analyzed with the Wilcoxon signed-rank test (Wilcoxon, 1945) as the focus was on medians. For our particular interest in the differences in changes between different demographics, we conducted two different analyses: 1) we compared the differences in changes between the sub-groups of each demographic variable separately by using the Mood's median tests with statistical significance level set to 0.05, and 2) we analyzed the changes in PA levels by using quantile (median) regression with the aforementioned demographic variables as explanatory variables and the baseline PA levels as a control variable. Missing values were handled by excluding the responses of a certain participant to a certain item if s/he had not responded it in both data collection rounds (a participant could have missed the follow-up for some reason or dropped out from the program entirely). Thus, the exact number of respondents slightly varies between the items. For the analysis, we formed broader sub-groups. For example, the different sub-groups of marital status were combined into broader "In a relationship" and "Not in a relationship" sub-groups.

Results

442 research program participants had been partaking for 12 months or more[1]. Of them, 294 responded to the IPAQ-E at both baseline and 12-month follow-up and formed the sample of this study (descriptive statistics reported in Table 1).

Table 1: Descriptive statistics of the sample of this study (N=294)

	n	%
Gender		
Male	109	37.1
Female	185	62.9
Other	0	0.0
Age (mean 69.7 years – standard deviation 4.2 years)		
–64 years	26	8.8
65–69 years	110	37.4
70–74 years	127	43.2
75 years or over	31	10.5
Marital Status		
Married	202	69.2
Common-law marriage	25	8.6
Single	9	3.1
Divorced	36	12.3
Widow(er)	20	6.8
N/A	2	–
Highest level of education		
Primary education	16	5.6
Vocational education	200	69.9
University of applied sciences	20	7.0
University	50	17.5
N/A	8	–
Language		
Finnish	227	77.2
Swedish	67	22.8

---

[1] At the time of conducting the study.

3.3   Overall Changes in the PA Levels

The statistical significance of the overall changes (all participants as one group) in the PA levels were analyzed with the Wilcoxon signed-rank test (Wilcoxon, 1945), the results of which are presented in Table 2. Regarding the overall changes in the *median minutes per week spent in different intensities*, there was a notable and statistically significant increase in walking. The median minutes per week for sitting and vigorous PA remained unchanged, whereas moderate PA had a very small but statistically not significant decrease. As for total PA MET-min, there was a notable and statistically significant increase.

Table 2: Overall changes in the PA levels (N=294)

Overall		Baseline (min/week)		Follow-up (min/week)		Change	p
	n	Median	25th–75th	Median	25th–75th	Median	
Sitting	283	2100	1680–2940	2100	1680–2730	0	0.661
Walking	276	540	273–840	840	405–1260	75	**< 0.001**
Moderate	284	240	120–450	233	90–420	0	0.529
Vigorous	286	60	0–240	60	0–240	0	0.603
Total PA	264	4098	2251–6435	4518	2326–6753	213	**0.037**

3.4   Differences in the Changes in PA Levels

When comparing the differences in changes in PA levels between the sub-groups of each demographic variable separately by using the Mood's median tests (Table 3), there was a statistically significant difference between marital statuses in sitting and walking (those *not in a relationship* increased walking time more than those *in a relationship*, but *they* also increased sitting time, whereas those *in a relationship* did not) and between age groups in walking (those aged *70 years or older* increased walking time, whereas those aged *under 70 years* did not). When explaining the changes in PA levels by using quantile (median) regression with the aforementioned demographic variables as explanatory variables and the baseline PA levels as a control variable (Table 4), there was a statistically significant difference between genders in walking (*females* having a more positive (or less negative) change than *males*) and between marital statuses in sitting (those *not in a relationship* having a more positive (or less negative) change than those *in a relationship*, which in the case of sitting is of course an undesirable change). These results also show that, in general, the higher the

baseline level the less positive (or more negative) change there has been in sitting, walking, moderate PA, vigorous PA, and total PA.

Table 3: Changes in median PA levels between the sub-groups (N=294)

	n	Change (min/week)		n	Change (min/week)		p
		Median	25th–75th		Median	25th–75th	
Gender		**Male**			**Female**		
Sitting	107	-105	-840–420	176	0	-420–420	0.305
Walking	105	20	-180–313	171	150	0–600	0.173
Moderate	104	0	-180–180	180	0	-180–131	0.520
Vigorous	109	0	-140–65	177	0	-60–90	0.987
Total PA	100	-86	-1544–2018	164	501	-1308–2667	0.205
Marital status		**In a relationship**			**Not in a relationship**		
Sitting	218	0	-630–420	63	420	-420–840	**0.001**
Walking	212	25	-140–420	62	240	0–630	**0.021**
Moderate	219	0	-180–180	63	20	-140–180	0.111
Vigorous	221	0	-120–83	63	0	-60–100	0.785
Total PA	203	107	-1581–2376	59	948	-740–2838	0.183
Highest education		**Vocational/Primary**			**Academic**		
Sitting	210	0	-534–420	65	0	-420–420	0.879
Walking	200	65	-55–420	68	93	-136–360	0.779
Moderate	206	0	-180–158	70	0	-120–158	0.903
Vigorous	209	0	-145–90	70	0	-53–101	0.910
Total PA	189	129	-1682–2419	68	806	-977–2541	0.083
Age		**–69**			**70–**		
Sitting	130	0	-630–420	153	0	-420–525	0.565
Walking	131	0	-125–420	145	150	0–520	**0.011**
Moderate	134	0	-180–180	150	0	-180–150	0.403
Vigorous	131	0	-60–90	155	0	-135–90	0.763
Total PA	125	44	-1467–2527	139	530	-1480–2462	0.109

Table 4: Results from the quantile regression (N=294)

	Sitting	Walking	Moderate	Vigorous	Total PA
Gender: Female Ref: Male	38.182	191.45**	-47.14	-19.34	523.44
Age: –69 Ref: 70–	-190.91	-80.13	-12.86	5.13	-529.79
Marital status: Not in a Ref: In a relationship	343.64*	55.66	4.29	-2.76	43.521
Highest education: V/P Ref: Academic	0	-37.11	-8.57	-1.97	-191.20
Intercept	1145.46*	259.74**	158.57**	31.97	1996.48**
Baseline PA level	-0.545***	-0.309***	-0.643***	-0.566***	-0.389***
Pseudo R²	0.178	0.098	0.148	0.148	0.102
The cells present the sizes and the statistical significances of the effects of the demographic variables on changes in PA levels after controlling the baseline PA levels; *p<0.05, **p<0.01, ***p<0.001; Ref = Reference group; Highest education: V/P = Vocational/Primary.					

4 Discussion and Conclusions

The main purpose of this study was to investigate the following research question: How effectively can a physical activity application promote PA behavior among aged people of different demographics? The study participants took into use a physical activity application to track their everyday PA. The self-reported PA was measured by using the IPAQ-E at two time points: before taking the application into use and after 12 months of use. The changes in four types of PA behavior (sitting, walking, moderate, vigorous) and total PA as well as demographic differences were analyzed.

As a response to the research question, in general, the results suggest that a physical activity application could be effective in promoting PA behavior among aged people as there was a notable and statistically significant increase in walking and total PA between baseline and 12-month follow-up. There was no statistically significant change in sitting, moderate PA, nor vigorous PA. These results are mostly in line with a recent systematic review on the ability of physical activity monitors to promote PA behaviors among older adults (Larsen et al. 2019). Moreover, for the vast majority of the participants in the present study the baseline measurement took place before the COVID-19 related restrictions, whereas for all the participants the follow-up measurement took place during the restrictions, which in Finland led to temporal closure or restricted access to different exercise facilities and also paused

many of the group activities. Considering this, another positive observation is that the participating aged people were able to upkeep or even increase their PA levels during these restrictions. The results also indicate that physical activity applications can be effective among aged people under free-living PA without accompanied active PA counseling or specific PA programs to follow. This is an important finding, as for most of the potential users, taking an application or another digital wellness technology into personal use just by themselves is a much more accessible option than signing up for a PA intervention or program with PA counselors.

Regarding the demographic differences, the analyses show little differences in the changes in PA behavior between the sub-groups. Although the results give a cautious indication that the effectiveness might be better among those not in a relationship, people aged 70 years or older, and females, in general, the effectiveness of a physical activity application to promote PA behavior change among aged people does not seem to considerably depend on the demographic background (gender, age, education, marital status) of the user. For the designers and developers this indicates that modifying the physical activity applications for different demographics of aged people may not be worthwhile, but we speculate that application modifications could be more in place for users of different PA backgrounds, as this likely has more effect on the user's needs towards PA solutions. Further research on this is warranted.

In terms of BCSSs and behavior change, the results indicate that physical activity applications can be utilized as BCSSs to support or persuade PA behavior change across different demographics; also for longer-term, as implied by the change in the PA levels after 12 months of use. Hence, from a more practical standpoint, such technologies could be used in different PA promotion programs and interventions or be taken into use by (aged) people looking for solutions to support their PA or PA behavior change. They could also be utilized to support PA during exceptional times, such as those when there are restrictions or home confinements in force.

5   Limitations & Future Research

This study has some limitations that should be acknowledged. First, those participants not tracking their PA before taking the application into use could have had a less accurate view on their baseline PA, whereas by using the application this view might have gotten more accurate. Hence, the latter measurement of PA could have been more accurate. Second, on average, the participants seemed to represent a rather physically active share of the aged population, which limits the generalization. For future research, it would be valuable to acquire more participants with lower PA levels to minimize non-participation bias. Third, due to the lack of control group, it is not certain whether the changes in PA resulted mainly from using the application or because of taking part in the study. We also cannot rule out the possible influence of some other uncontrolled factors, such as the COVID-19 pandemic and the resulting restrictions, which have potentially influenced some participants' PA levels around the follow-up. Hence, further related research is warranted. Future research with different digital wellness technologies and focus on other changes besides behavior would be valuable. As long-term research on related topics is much called for, we plan to continue the follow-ups to complement earlier study findings (e.g., Carlsson et al. 2020; Kari et al. 2021; Makkonen et al. 2020).

Acknowledgements

References

Carlsson, C., Kari, T., Makkonen, M., Frank, L., Walden, P. (2020). Sustainable physical activity programs for young elderly – A fuzzy AHP approach. In Proceedings of the 33rd Bled eConference "Enabling Technology for a Sustainable Society", pp. 423-438. University of Maribor, Bled, Slovenia.

Carlsson, C., Walden, P. (2017). Digital coaching to build sustainable wellness routines for young elderly. In Proceedings of the 30th Bled eConference "Digital Transformation – From Connecting Things to Transforming Our Lives", pp. 57-70. University of Maribor, Bled, Slovenia.

Changizi, M., Kaveh, M. H. (2017). Effectiveness of the mHealth technology in improvement of healthy behaviors in an elderly population—A systematic review. mHealth, 3, 51.

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., ... Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. Medicine & Science in Sports & Exercise, 35, 1381-1395.

Ekelund, U., Sepp, H., Brage, S., Becker, W., Jakes, R., Hennings, M., Wareham, N. J. (2006). Criterion-related validity of the last 7-day, short form of the International Physical Activity Questionnaire in Swedish adults. Public Health Nutrition, 9, 258-265.

Fogg, B. J. (2003). Persuasive technology: using computers to change what we think and do. Morgan Kaufmann Publishers, San Francisco, CA.

Gordon, M., Althoff, T., Leskovec, J. (2019). Goal-setting and achievement in activity tracking apps: a case study of MyFitnessPal. In Proceedings of the World Wide Web Conference, pp. 571-582. ACM, New York, NY.

Hoogendijk, E. O., Afilalo, J., Ensrud, K. E., Kowal, P., Onder, G., Fried, L. P. (2019). Frailty: implications for clinical practice and public health. The Lancet, 394, 1365-1375.

Hurtig-Wennlöf, A., Hagströmer, M., Olsson, L. A. (2010). The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility. Public Health Nutrition, 13, 1847-1854.

IPAQ group. (2005a). International Physical Activity Questionnaire. https://sites.google.com/site/theipaq/home/, accessed 10.2.2021.

IPAQ group. (2005b). Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). https://sites.google.com/site/theipaq/scoring-protocol/, accessed 10.2.2021.

Kari, T. (2014). Can exergaming promote physical fitness and physical activity? a systematic review of systematic reviews. International Journal of Gaming and Computer-Mediated Simulations, 6, 59-77.

Kari, T., Kettunen, E., Moilanen, P., Frank, L. (2017). Wellness technology use in everyday life: a diary study. In Proceedings of the 30th Bled eConference "Digital Transformation – From Connecting Things to Transforming Our Lives", pp. 279-294. University of Maribor, Bled, Slovenia.

Kari, T., Koivunen, S., Frank, L., Makkonen, M., Moilanen, P. (2016a) Critical Experiences During the Implementation of a Self-tracking Technology. In Proceedings of the 20th Pacific Asia Conference on Information Systems, pp. 1-16. AIS, Chiayi, Taiwan.

Kari, T., Makkonen, M., Carlsson, J., Frank, L. (2021). Using a Physical Activity Application to Promote Physical Activity Levels Among Aged People: A Follow-Up Study. In Proceedings of the 54th Hawaii International Conference on System Sciences, pp. 1242-1251. University of Hawai'i at Manoa, Hawaii.

Kari, T., Piippo, J., Frank, L., Makkonen, M., Moilanen, P. (2016b). To gamify or not to gamify? gamification in exercise applications and its role in impacting exercise motivation. In Proceedings of the 29th Bled eConference "Digital economy", pp. 393-405. University of Maribor, Bled, Slovenia.

Kari, T., Rinne, P. (2018). Influence of digital coaching on physical activity: motivation and behaviour of physically inactive individuals. In Proceedings of the 31st Bled eConference "Digital Transformation – Meeting the Challenges", pp. 127-145. University of Maribor, Bled, Slovenia.

Kari, T., Salo, M., & Frank, L. (2019). Role of Situational Context in Use Continuance After Critical Exergaming Incidents. Information Systems Journal, 30, 596-633.

Karppinen, P., Oinas-Kukkonen, H., Alahäivälä, T., Jokelainen, T., Keränen, A. M., … Savolainen, M. (2016). Persuasive user experiences of a health Behavior Change Support System: A 12-month study for prevention of metabolic syndrome. International Journal of Medical Informatics, 96, 51-61.

Kettunen, E., Kari, T., Makkonen, M., Frank, L., Critchley, W. (2020). Young elderly and digital coaching: a quantitative intervention study on exercise self-efficacy. In Proceedings of the 33rd Bled eConference "Enabling Technology for a Sustainable Society", pp. 469-484. University of Maribor, Bled, Slovenia.

Kettunen, E., Kari, T., Moilanen, P., Vehmas, H., Frank, L. (2017). Ideal types of sport and wellness technology users. In Proceedings of the 11th Mediterranean Conference on Information Systems, 12 pages. AIS, Genoa, Italy.

Kirwan, M., Duncan, M., Vandelanotte, C. (2013). Smartphone apps for physical activity: a systematic review. Journal of Science and Medicine in Sport, 16, e47.

Koivisto, J., Hamari, J. (2019). The rise of motivational information systems: a review of gamification research. International Journal of Information Management, 45, 191-210.

Larsen, R. T., Christensen, J., Juhl, C. B., Andersen, H. B., Langberg, H. (2019). Physical activity monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis. European Review of Aging and Physical Activity, 16, 7.

Lee, P. H., Macfarlane, D. J., Lam, T. H., Stewart, S. M. (2011). Validity of the international physical activity questionnaire short form (IPAQ-SF): A systematic review. International Journal of Behavioral Nutrition and Physical Activity, 8, 115.

Loos, E., Zonneveld, A. (2016). Silver gaming: serious fun for seniors? In International Conference on Human Aspects of IT for the Aged Population, pp. 330-341. Springer, Cham.

Makkonen, M., Frank, L., Kari, T., Moilanen P. (2012). Examining the usage intentions of exercise monitoring devices: The usage of pedometers and route trackers in Finland. In Proceedings of the 25th Bled eConference, pp. 439-453. University of Maribor, Bled, Slovenia.

Makkonen, M., Kari, T., and Frank, L. (2020). Applying UTAUT2 to Explain the Use of Physical Activity Logger Applications Among Young Elderly. In Proceedings of the 33rd Bled eConference "Enabling Technology for a Sustainable Society", pp. 567-582. University of Maribor, Bled, Slovenia.

Moilanen, P., Salo, M., Frank, L. (2014). Inhibitors, enablers and social side winds: explaining the use of exercise tracking systems. In Proceedings of the 27th Bled eConference, pp. 23-37. University of Maribor, Bled, Slovenia.

Muellmann, S., Forberger, S., Möllers, T., Bröring, E., Zeeb, H., Pischke, C. R. (2018). Effectiveness of eHealth interventions for the promotion of physical activity in older adults: A systematic review. Preventive Medicine, 108, 93-110.

Oinas-Kukkonen, H. (2013). A foundation for the study of behavior change support systems. Personal and Ubiquitous Computing, 17, 1223-1235.

Oduor, M., Oinas-Kukkonen, H. (2021). Committing to change: a persuasive systems design analysis of user commitments for a behaviour change support system. Behaviour & Information Technology, 40, 20-38.

Rockmann, R. (2019). Don't Hurt Me… No More? An Empirical Study on the Positive and Adverse Motivational Effects in Fitness Apps. In Proceeding of the European Conference on Information Systems, 17 pages. AIS, Stockholm & Uppsala, Sweden.

Romeo, A., Edney, S., Plotnikoff, R., Curtis, R., Ryan, J., Sanders, I., ... Maher, C. (2019). Can smartphone apps increase physical activity? systematic review and meta-analysis. Journal of Medical Internet Research, 21, e12053.

Seifert, A., Schlomann, A., Rietz, C., Schelling, H. R. (2017). The use of mobile devices for physical activity tracking in older adults' everyday life. Digital Health, 3, 1-12.

Schmidt, B., Benchea, S., Eichin, R., Meurisch, C. (2015). Fitness tracker or digital personal coach: how to personalize training. In Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 1063-1067. ACM, Osaka, Japan.

Stephenson, A., McDonough, S. M., Murphy, M. H., Nugent, C. D., Mair, J. L. (2017). Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. Journal of Behavioral Nutrition and Physical Activity, 14, 105.

Stockwell, S., Schofield, P., Fisher, A., Firth, J., Jackson, S.E., Stubbs, B., Smith, L. (2019). Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: A systematic review and meta-analysis. Experimental Gerontology, 120, 68-87.

Sullivan, A. N., Lachman, M. E. (2017). Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Frontiers in Public Health, 4, 289.

THL – The Finnish Institute for Health and Welfare. (2019). Physical activity in the adult population in Finland – the FinHealth Study. Report 48/2019.

United Nations. (2019). World population ageing 2019. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf, accessed 9.2.2021.

Van Poppel, M. N., Chinapaw, M. J., Mokkink, L. B., Van Mechelen, W., Terwee, C. B. (2010). Physical activity questionnaires for adults: a systematic review of measurement properties. Sports Medicine, 40, 565-600.

Wang, J. B., Cataldo, J. K., Ayala, G. X., Natarajan, L., Cadmus-Bertram, L. A., ... Pierce, J. P. (2016). Mobile and wearable device features that matter in promoting physical activity. Journal of Mobile Technology in Medicine, 5, 2-11.

Wellmo. (2019). Mobile health platform. https://www.wellmo.com/platform/, accessed 9.2.2021.

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. Biometrics Bulleting, 1, 80-83.

World Health Organization. (2020). WHO guidelines on physical activity and sedentary behaviour. https://www.who.int/publications/i/item/9789240015128, accessed 9.2.2021.

Yerrakalva, D., Yerrakalva, D., Hajna, S., Griffin, S. (2019). Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: Systematic review and meta-analysis. Journal of Medical Internet Research, 21, e14343

REAL-WORLD REINFORCEMENT LEARNING: OBSERVATIONS FROM TWO SUCCESSFUL CASES

PHILIPP BACK

Aalto University School of Business, Department of Information and Service Management, Helsinki, Finland; e-mail: philipp.back@aalto.fi

Abstract Reinforcement Learning (RL) is a machine learning technique that enables artificial agents to learn optimal strategies for sequential decision-making problems. RL has achieved superhuman performance in artificial domains, yet real-world applications remain rare. We explore the drivers of successful RL adoption for solving practical business problems. We rely on publicly available secondary data on two cases: data center cooling at Google and trade order execution at JPMorgan. We perform thematic analysis using a pre-defined coding framework based on the known challenges to real-world RL by Dulac-Arnold, Mankowitz, & Hester (2019). First, we find that RL works best when the problem dynamics can be simulated. Second, the ability to encode the desired agent behavior as a reward function is critical. Third, safety constraints are often necessary in the context of trial-and-error learning. Our work is amongst the first in Information Systems to discuss the practical business value of the emerging AI subfield of RL.

1   Introduction

Reinforcement Learning (RL) is a machine learning technique that enables artificial agents to learn optimal strategies for sequential decision-making problems (Sutton & Barto, 2018). No direct supervision is provided; the agent learns by trial-and-error while interacting with a (virtual) environment and receiving feedback on its actions.

Over the past decade, RL has achieved superhuman performance in a series of artificial domains. In a seminal paper by Google DeepMind (Mnih et al., 2015), an RL system learned how to play different Atari games, such as Breakout and Space Invaders, on a human or even superhuman level. This breakthrough started an RL frenzy in the scientific community and was presumably a major reason for Google to acquire DeepMind for $650 million in 2014. The next milestone came with AlphaGo, the first computer program to master the ancient Chinese board game of Go (Silver et al., 2017); a feat that experts believed was still decades away as the number of possible game states exceeds the number of atoms in the known universe. Over 200 million people watched online as AlphaGo beat the world's best Go player, Lee Sedol, 4 to 1 during The Google DeepMind Challenge. In the aftermath of the game, Lee Sedol described his opponent as following: *"I thought AlphaGo was based on probability calculation and that it was merely a machine. But when I saw this move, I changed my mind. Surely, AlphaGo is creative"* (Silver, Hubert, Schrittwieser, & Hassabis, 2018).

Despite its many opportunities, RL also presents significant challenges that have prevented wide-spread adoption in the real world (Dulac-Arnold et al., 2019). Google's DeepMind unit, conqueror of Atari, Go, and StarCraft II has been losing an estimated $1 billion over the past three years while trying to transfer its impressive work from artificial domains to real-world products - so far with little success.

Interestingly, Information Systems (IS) research has largely ignored the emerging AI subfield of RL. The research landscape remains dominated by the computer science community who continues to present algorithmic improvements that are usually evaluated on artificial problems. IS often treats AI technologies as mere **tools** for certain applications, such as decision making, supply chain management, market predictions, or innovation, rather than as objects of research themselves (Shmueli & Koppius, 2011). Specific AI subfields, such as machine learning, big data analytics, or predictive modeling, remain under-researched despite their rising importance to

management (Nascimento, da Cunha, de Souza Meirelles, Scornavacca Jr, & de Melo, 2018). Case in point: RL has already been discussed in IS as a **tool** for music recommendation (Liebman, Saar-Tsechansky, & Stone, 2019), but not yet as novel AI subfield **(object)**. However, amidst the ongoing RL hype, IS would serve practitioners well by treating RL as on object of study itself. Already 30 years ago, King (1984) pointed to a gap between inflated claims on what AI-based expert systems may be able to do, and what has actually been delivered. More recently, Nascimento et al. (2018) confirmed that claims made by AI vendors and the media seem ahead of what is supported by research findings. By critically examining RL's potential business value, IS would continue its long history of educating practitioners about the possibilities and challenges of novel AI technology. To quote Noel Sharkey, emeritus professor of Artificial Intelligence and Robotics at the University of Sheffield: *"[...] the wrong idea of what robotics can do and where AI is at the moment it's very, very dangerous."* (Delcker, 2018).

To this end, we conduct a qualitative exploratory study on two successful cases of real-world RL adoption: data center cooling by Google and trade order execution by JPMorgan. The research question that this paper seeks to address is: *What factors drive successful real-world adoption of Reinforcement Learning for practical business problems?* By studying "how others did it", we hope to provide decision-makers with a better understanding of RL's opportunities and pitfalls; and how to overcome them.

2   Theoretical Background

2.1   Reinforcement Learning Overview

RL is one of the three machine learning paradigms, alongside supervised and unsupervised learning (Sutton & Barto, 2018). An RL agent learns by interacting with its (virtual) environment (Figure 1). At each time step t, the agent observes the environment state s_t and chooses an action a_t from the set of available actions. The environment moves to the next state s_{t+1} and the agent receives a reward r_{t+1} based on the transition (s_t, a_t, s_{t+1}). The goal of an RL agent is to learn an optimal strategy (policy) which maximizes the expected cumulative reward. No direct supervision is provided to the agent; it is never directly told the best course of action. Rather, the agent has to learn from the consequences of its actions via trial-and-error. The only guidance comes from the numerical reward, a reinforcement signal

that encodes how good it is to take an action in a certain state. The agent must reason about the long-term consequences of its actions although the immediate reward associated with its action might be negative. For example, a stock trading agent may learn to accept small daily losses in exchange for a significant payoff when the market sentiment changes. Thus, RL is particularly well suited for sequential decision-making problems that include long-term versus short-term reward trade-offs[1].
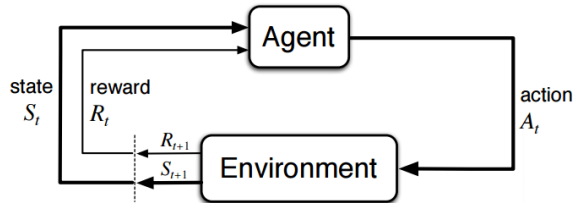
Figure 1: Reinforcement Learning framework
Source: Sutton & Barto (2018)

2.2   Known Challenges to Real-World Reinforcement Learning

Dulac-Arnold et al. (2019) identified nine challenges that – if present in an underlying real-world problem - must be addressed to productize RL:

1. training off-line from limited (historic) logs of the system's behavior
2. learning on the real system from limited (historic) data samples
3. high-dimensional continuous state and action spaces that become computational infeasible to search over (curse of dimensionality)
4. safety constraints that should never or at least rarely be violated
5. tasks that may only be partially observable (non-stationary/stochastic)
6. reward function design (unspecified, multi-objective, or risk-sensitive)
7. system operators who desire explainable policies and actions
8. inference that must happen in real-time at the system's control frequency[2]
9. large and/or unknown delays in the system actuators, sensors, or rewards

---

[1] Given the scope and limitation of this study, we refrain from a more detailed technical overview of RL and kindly refer the interested reader to Sutton & Barto (2018).
[2] The task can be run neither faster nor slower than real-time. This limits the quick generation of massive amounts of training, as well as slow, computationally expensive modeling approaches.

3   Methodology

To examine the drivers of successful RL adoption for practical business problems, we conducted a qualitative exploratory study on two cases: data center cooling by Google and trade order execution by JP Morgan. We selected the two cases based on the simple fact that they are – to the best of our knowledge – some of the only instances where RL has been successfully deployed in practice. We intentionally did not consider cases of self-driving cars and user-recommendation systems, as they are either not yet production-ready (Osiński et al., 2020), or use simpler forms of RL, such as contextual bandits (Amat, Chandrashekar, Jebara, & Basilico, 2018).

3.1   Data Collection

We rely on publicly available material for exploring the use of RL at Google and JPMorgan. Our sources include research publications, blog posts, newspaper articles, and published interviews with company representatives. We chose this form of qualitative secondary analysis (Heaton, 2008) because both case companies are reluctant to give interviews on the inner workings of their RL systems that would go beyond what they already chose to disclose. Nevertheless, useful insights can be gained from a careful analysis of various publicly available sources.

3.2   Data Analysis

We performed a thematic analysis using a pre-defined coding framework. Thematic analysis is a method for systematically identifying, organizing, and offering insights into patterns of meaning (themes) across a data set (Braun & Clarke, 2012). This method allowed us to identify what is common to the way in which the topic – what drives successful RL adoption for practical business problems - was described across multiple data items, and to make sense of those commonalities. As **themes**, we used the nine challenges to real-world RL (Section 2.2) by Dulac-Arnold et al. (2019). After becoming familiar with the data, we consolidated all sources in text format and systematically searched for the pre-defined themes. For example, the excerpt *"We send simulated orders to the exchange, we simulate how they execute, we simulate market impact, and then we feed the reward and batches of execution back to the agent's brain […]"* (QuantMinds365, 2018) related to "data efficiency" (theme 2) and "reward function design" (theme 6). Finally, we synthesized our findings in three key observations.

4 Findings

Here we present our findings to answer the research question of w*hat factors drive successful real-world adoption of Reinforcement Learning for practical business problems*. We start with a general overview of optimal trade execution at JPMorgan and data center cooling at Google. Next, we present the findings of our thematic analysis, i.e. how RL has been productized in each case. Finally, we highlight the business value that RL offered over competing solutions.

Overall, the thematic analysis of secondary data sources produced observations on four out of the nine themes: training from fixed logs (1), learning from limited samples (2), safety constraints (4), and reward function design (6).

4.1 Case 1: Optimal Trade Execution (JPMorgan)

Banks and financial service firms have long been using algorithms to make equity trading more efficient. In 2017, JPMorgan, one of the leading global financial services firm with assets of $2.6 trillion, announced LOXM, an AI-based limit-order placement engine that takes efficient trade execution to new heights. LOXM can execute equity trades at maximum speed and at optimal prices, and allows clients to offload large equity positions without causing market swings (Terekhova, 2017).

Optimal trade execution is a non-trivial problem: client needs differ, market conditions vary, and legal regulations must be met. An AI agent must learn to operate in the environment of bid/ask prices, and monitor the liquidity on both sides of the order book (QuantMinds365, 2018).

JPMorgan addressed these challenges by using RL. They constructed a market simulator from billions of past trades to provide the RL agent with a learning environment (Nevmyvaka, Feng, & Kearns, 2006; Vyetrenko et al., 2019). This simulator approach eliminated the need to learn directly from limited historic samples (theme 2). The artificial environment receives orders from the agent, simulates market impact, and sends rewards back to the agent who then updates its understanding of what actions are good/bad. In an interview, Vaslav Glukhow, Head of EMEA e-Trading Quantitative Research at JPMorgan, explained the process as following: *"In this approach we have an action, and the action is how much to place,*

what price, and for how long. It makes sense for the agent to be intelligent about the quantity that it asks for, and it needs to be smart in terms of how long it needs to place that order – if it gives the order for too long it will lose the opportunity and it will need to execute at a higher price. All these things need to be taken into consideration and the agent needs to be aware of the consequences of each action." (QuantMinds365, 2018). The agent is rewarded (theme 6) for being efficient in the market in the form of a scalar reward signal that encodes how well the agent splits large orders into smaller, more efficient "child" orders compared to executing the large "parent" order at once (Vyetrenko & Xu, 2019). LOXM makes use of RL's ability to balance long- and short-term rewards; the total rewards are not necessarily the sum of local rewards.

Since LOXM's depute in 2017, RL has proven its real-world worth for optimal trade execution. According to JPMorgan, the system provides significant savings and far outperforms both manual and existing automated trading methods (Terekhova, 2017).

4.2   Case 2: Data Center Cooling (Google)

Cooling is a critical component of data center operations. Servers produce considerable amount of heat, and high temperatures may lead to lower IT performance or equipment damage (Lazic et al., 2018). Dealing with excess heat is one of the biggest, most expensive factors when running a modern data center. Past solutions to the temperature problem have included moving data centers to cooler climates, or even situating them at the bottom of the ocean.

In 2018, Google announced that it has handed control over the cooling of several of its behemoth data centers to an AI that optimizes effective power management. The system learns how to adjust fans, ventilation, and other equipment to lower power consumption (Knight, 2018). Two years earlier, in 2016, Google had already presented an earlier version that made recommendations to human data center managers, who would then decide whether to implement them (Evans & Gao, 2016). The new system is managing cooling all by itself, although a human can still intervene. *"It's the first time that an autonomous industrial control system will be deployed at this scale, to the best of our knowledge",* said Mustafa Suleyman, head of applied AI at DeepMind, an AI company that was acquired by Google in 2014 (Knight, 2018).

Google uses RL to find optimal control policies for their complex, large-scale data center systems. Google does not explicitly disclosure how the RL learning environment has been constructed (Lazic et al., 2018), only that historic control data has insufficient information to directly train an RL agent (theme 2). However, earlier work has shown that neural networks can model Google's data center dynamics with 99.6% accuracy (Kava, 2014). Thus, it can be assumed that Google is using the latter model as a simulator to construct the RL environment. To ensure safe operation already during training (theme 4), Google used historic control logs (theme 1) to limit each control variable to a safe range. In the absence of such data, the safe range could be initialized conservatively and gradually expanded (Lazic et al., 2018).

According to Google, deep RL has proven highly effective at operating cooling systems: the system consistently achieves a 40% reduction in the amount of energy used for cooling (Evans & Gao, 2016).

5   Discussion

In both applications - optimal trade execution and data center cooling - the goal was to find an optimal strategy for a sequential decision-making problem. Despite very different domains, both applications share certain characteristics that allowed RL to be applied in practice. In the following, we discuss some of the key challenges in RL (Dulac-Arnold et al., 2019) and how Google and JPMorgan overcame them.

5.1   Observation 1: Learning Directly from Historic Data is Still Difficult

Standard supervised learning is teaching by example, whereas RL is teaching by experience. The agent gathers experience by interacting through trial-and-error with its environment; like a virtual playground. The learning by trial-end-error framework can be applied to almost any sequential decision-making problem, but this generality comes at a price: RL is hugely sample inefficient, meaning it requires a lot of training data. Imagine trying to learn a new board game, but instead of studying the rule book or recalling your experience from other board games, you simply take random moves while only receiving the final game score as feedback. What quickly strikes humans as a ludicrous endeavor is precisely how many RL problems are framed. Indeed, given enough time and computational power, the trial-and-error approach should converge to a (near-) optimal strategy, but what is enough? For AlphaGo it took 5

billion games of self-play and around $35 million in computing costs to beat the best human Go player.

Since RL agents require a lot of experience, and experience is gathered by interacting with an environment, the way in which a (real-world) problem can be represented as an RL environment is critical. Most of RL's successes have been with artificial systems that can be simulated, such Atari games (Mnih et al., 2015), Go (Silver et al., 2017), DOTA (Berner et al., 2019), and StarCraft II (Vinyals et al., 2019). The reason for this is that simulators allow us to generate unlimited training data - an advantage that can hardly be overstated in the context of sample inefficient learning. An RL agent might not be able to gather sufficient experience from a set of finite training data, but with a simulated environment it can continue to interact and learn until a (near-) optimal policy is found. Moreover, simulators eliminate much of the messiness that an agent may face in the real world and thus provide stable benchmarks against which new algorithms can be compared. Simulated environments therefore also play an important role in algorithmic research.

Unfortunately, out-of-the-box simulators rarely exist for real-world applications. The two herein presented applications addressed this issue by improving the fidelity of the simulations to a point where the gap between simulations and the real world was so small that things learned in simulation were directly transferable to the real world. JPMorgan used billions of past trades to construct a market simulator that closely reflected the true market dynamics (Vyetrenko et al., 2019). Similarly, Google managed to simulate data center dynamics with 99.6% accuracy (Kava, 2014). In contrast to training directly on a (limited) historic dataset (theme 2), simulators offer unlimited opportunities to learn. Assuming that the simulator accurately reflects the real system, it then becomes possible to transfer the learned optimal strategy directly from the artificial environment to the real world. The existence of an environment simulator, or the ability to construct one from historic data or first principles, is therefore a common characteristic of successful real-world RL applications.

Observation 1: If the tasks in a field can be accurately simulated, RL may dramatically improve the state of the art in that field over the next few years.

5.2   Observation 2: It's All about Rewards

The desired behavior of an RL agent has to be encoded in a reward function. Without direct expert guidance, the reward function is the only source of feedback that tells the agent what action in which state was good. Unfortunately, reward function design is notoriously difficult as it must capture *exactly* the desired behavior. An ill-defined reward function will cause the RL agents to break in surprising, counterintuitive, and sometimes amusing ways. One of the most (in-) famous examples of faulty reward functions is an RL agent learning to play the CoastRunners game (Clark & Amodei, 2016). The goal of the game - as understood by most humans - is to finish a boat race quickly and (preferably) ahead of others. Players are not directly rewarded for progressing around the course but earn points by hitting targets laid out along the racetrack. The human who designed the reward system simply assumed that the overall game score would implicitly reflect the goal of finishing the race. However, the targets were positioned so that the RL agent could gain a high score without ever having to finish the race. Instead, the agent finds an isolated lagoon where it can turn in a large circle and repeatedly knock over three targets, timing its movement so as to always knock over the targets just as they repopulate. Despite repeatedly catching on fire, crashing into other boats, and going the wrong way on the track, the agent manages to achieve a higher score using this strategy than is possible by completing the race in the normal (human) way (Clark & Amodei, 2016). The counterintuitive strategy leads to scores that are on average 20% higher than what is achieved by human players.

Reward function design becomes even more complex when systems have multi-dimensional costs, or when product owners cannot even articulate clearly what they want to minimize (Dulac-Arnold et al., 2019). In both herein presented cases, however, it was possible to encode the desired agent behavior as a scalar reward signal (theme 6): minimize trading costs and energy consumption. Neither JPMorgan nor Google had to perform any tricks to encode a complex behavior or multi-dimensional objective as a scalar reward. The natural representation of the desired system behavior thus greatly contributed to the suitability of RL.

Observation 2: An agent's desired behavior must be able to be expressed exactly as a reward function.

5.3   Observation 3: Safety First

Most physical systems can damage themselves and their environment if operated incorrectly. When neither a simulator nor sufficient training data are available, and when the RL agents has to gather experience by interacting directly via trial-and-error with the real system, safety becomes important; not only during the final system operation, but also during the exploratory learning phase. For example, a self-driving car cannot be allowed to explore all possible actions, including crashing full speed into a wall, until it has eventually mastered to stay within its lane. (Un)fortunately, safety violations will likely be very rare in historic logs and thus do not provide sufficient opportunity to learn from them (Dulac-Arnold et al., 2019). Furthermore, safety constraints are often assumed ("it's just common sense!") and are not even specified explicitly. Moreover, it may not be possible to impose strict limits by describing the action or state space directly; instead, hazard-avoiding behavior must be learned.

Safety constraints were crucial in both herein discussed applications of RL. Google's data center represents a large-scale system that can suffer catastrophic damage when operated incorrectly. JPMorgan had to consider legal, regulatory, and ethical constraints that may prohibit certain trading activities. As discussed in Section 5.2, it is notoriously difficult to express fine-grained behaviors in the reward function. Constrained RL is thus a viable alternative to prohibit the agent from learning, or even exploring, certain actions.

Observation 3: Trial-and-error learning makes safety considerations during training and control highly important.

6   Conclusion and Further Directions

Reinforcement Learning has achieved superhuman performance in many artificial domains, yet real-world applications remain rare. We conducted a qualitative study to explore the drivers of successful RL adoption for practical business problems. We used pre-defined themes based on the known challenges to real-world RL by Dulac-

Arnold et al. (2019) to perform a thematic analysis on secondary data that described how Google and JPMorgan successfully deployed RL for data center cooling and optimal trade execution. Three success drivers emerged: simulatability of the problem dynamics, a natural representation of the desired agent behavior as a reward function, and compliance with safety constraints during trial-and-error learning. Our work is amongst the first in IS to approach RL as a separate object of study. By exploring two successful real-world RL applications, we hope to shine some light on the practical business value of this emerging AI subfield.

Notwithstanding some useful insights, this study is limited by the use of publicly-available secondary data. For example, we were only able to identify four out of the nine themes (challenges) in our dataset. Especially the Google case could have benefited from more details. This study is thus a mere first step towards a more complete understanding of the practical business value of RL. For the future, we hope to collect primary data through interviews with key people who were involved in the design and deployment of the herein described RL solutions at Google and JPMorgan. Such data would allow us to directly analyze the technical details of the RL systems, rather than using high-level (textual) descriptions. Our study is further limited by having only two case companies. It will be interesting to extend our study to future use cases of RL once this method has become more widely adopted. Finally, it would be interesting to tackle a new application area with RL, and to document how the known challenges are addressed during the design process.

References

Amat, F., Chandrashekar, A., Jebara, T., & Basilico, J. (2018). Artwork personalization at Netflix. Proceedings of the 12th ACM Conference on Recommender Systems, 487–488. Vancouver, Canada.

Berner, C., Brockman, G., Chan, B., Cheung, V., Dębiak, P., Dennison, C., … Hesse, C. (2019). Dota 2 with large scale deep reinforcement learning. ArXiv Preprint ArXiv:1912.06680.

Braun, V., & Clarke, V. (2012). Thematic analysis. In APA handbook of research methods in psychology, Vol 2 (pp. 57–71). Washington, DC: American Psychological Association.

Clark, J., & Amodei, D. (2016). Faulty Reward Functions in the Wild. Retrieved September 20, 2020, from https://openai.com/blog/faulty-reward-functions/

Delcker, J. (2018). Europe divided over robot 'personhood.' Retrieved November 10, 2020, from https://www.politico.eu/article/europe-divided-over-robot-ai-artificial-intelligence-personhood/

Dulac-Arnold, G., Mankowitz, D., & Hester, T. (2019). Challenges of real-world reinforcement learning. ArXiv Preprint ArXiv:1904.12901.

Evans, R., & Gao, J. (2016). DeepMind AI reduces Google data centre cooling bill by 40%. Retrieved September 20, 2020, from https://deepmind.com/blog/article/deepmind-ai-reduces-google-data-centre-cooling-bill-40

Heaton, J. (2008). Secondary analysis of qualitative data: An overview. Historical Social Research/Historische Sozialforschung, 33–45.

Kava, J. (2014). Better data centers through machine learning. Retrieved August 21, 2020, from https://blog.google/inside-google/infrastructure/better-data-centers-through-machine/

King, W. R. (1984). Editor's comment: Decision support systems, artificial intelligence, and expert systems. MIS Quarterly, 8(3), iv–v.

Knight, W. (2018). Google just gave control over data center cooling to an AI. Retrieved September 5, 2020, from https://www.technologyreview.com/2018/08/17/140987/google-just-gave-control-over-data-center-cooling-to-an-ai/

Lazic, N., Lu, T., Boutilier, C., Ryu, M. K., Wong, E. J., Roy, B., & Imwalle, G. (2018). Data center cooling using model-predictive control. Advances in Neural Information Processing Systems. Montreal, Canada.

Liebman, E., Saar-Tsechansky, M., & Stone, P. (2019). The right music at the right time: Adaptive personalized playlists based on sequence modeling. MIS Quarterly, 43(3).

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., … Ostrovski, G. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529.

Nascimento, A. M., da Cunha, M. A. V. C., de Souza Meirelles, F., Scornavacca Jr, E., & de Melo, V. V. (2018). A literature analysis of research on artificial intelligence in management information system (MIS). AMCIS.

Nevmyvaka, Y., Feng, Y., & Kearns, M. (2006). Reinforcement learning for optimized trade execution. Proceedings of the 23rd International Conference on Machine Learning, 673–680. Pittsburgh, USA.

Osiński, B., Jakubowski, A., Zięcina, P., Miłoś, P., Galias, C., Homoceanu, S., & Michalewski, H. (2020). Simulation-based reinforcement learning for real-world autonomous driving. 2020 IEEE International Conference on Robotics and Automation (ICRA), 6411–6418. Paris, France.

QuantMinds365. (2018). The Latest in LOXM and why we shouldn't be using single stock algos. Retrieved July 3, 2020, from https://informaconnect.com/the-latest-in-loxm-and-why-we-shouldnt-be-using-single-stock-algos/

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. MIS Quarterly, 553–572.

Silver, D., Hubert, T., Schrittwieser, J., & Hassabis, D. (2018). AlphaZero: Shedding new light on chess, shogi, and Go. Retrieved September 18, 2020, from https://deepmind.com/blog/article/alphazero-shedding-new-light-grand-games-chess-shogi-and-go

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., … Bolton, A. (2017). Mastering the game of go without human knowledge. Nature, 550(7676), 354–359.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). Cambridge: MIT Press.

Terekhova, M. (2017). JPMorgan takes AI use to the next level. Retrieved August 19, 2020, from https://www.businessinsider.com/jpmorgan-takes-ai-use-to-the-next-level-2017-8?r=US&IR=T

Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., … Georgiev, P. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature, 575(7782), 350–354.

Vyetrenko, S., Byrd, D., Petosa, N., Mahfouz, M., Dervovic, D., Veloso, M., & Balch, T. H. (2019). Get real: Realism metrics for robust limit order book market simulations. ArXiv:1912.04941.

Vyetrenko, S., & Xu, S. (2019). Risk-sensitive compact decision trees for autonomous execution in presence of simulated market response. ArXiv Preprint ArXiv:1906.02312

ONLINE PROCTORING: ADDING HUMAN VALUES TO THE EQUATION

MARLIES VAN STEENBERGEN & IRENE VAN DER SPOEL

HU University of Applied Sciences Utrecht, Research Centre for Learning and Innovation, Utrecht, The Netherlands; e-mail: marlies.vansteenbergen@hu.nl, irene.vanderspoel@hu.nl

Abstract The COVID-19 pandemic led to an accelerated implementation of digital solutions, such as online proctoring. In this paper we discuss how the use of an ethical matrix may influence the way in which digital solutions are applied. To initiate an ethical discussion, we conducted an online workshop with educators, examiners, controllers, and students to identify risks and opportunities of online proctoring for various stakeholders. We used the *Ethical Matrix* to structure the meeting. We compared the outcome of the workshop with the outcomes of a proctoring software pilot by examiners. We found that the two approaches led to complementary implementation criteria. The ethical session was less focused on making things work and more on transparency about conditions, processes, and rights. The ethical session also concentrated more on the values of all involved rather than on fraud detection effectiveness.

1   Introduction

The COVID-19 pandemic has accelerated the already existing trend of increased digitalization in government, commerce, and education. Studies show, for instance, that within education the use of technology enhanced learning has jumped forward in 2020 (van der Spoel et al., 2020). Educational institutions feel pressed to introduce digital solutions such as online proctoring. Despite concerns regarding human values such as privacy, distributive justice, autonomy, well-being, reputation, and reliability, the pressure for fast action may be so strong that it carries with it the risk of unintended negative consequences or backfiring (Stibe & Cugelman, 2016) and ethical blindness (Palazzo et al., 2013). To mitigate this risk, it is important to integrate explicit ethical discussion in the design and implementation process (Van den Hoven, 2017). One tool to support this, is the ethical matrix (Mepham, 2000; van der Stappen & van Steenbergen, 2020). The ethical matrix is a tool that stimulates a closer look at potential risks and opportunities of digital innovations.

At the start of the first corona lockdown in March 2020, Dutch institutes of higher education were pressured to find alternative ways of taking exams. Having entire classes sit an exam in large halls under surveillance of a human proctor was no longer an option. Many courses turned to alternative ways of examination, such as having students write essays. But for some courses, the only viable option turned out to be taking the exam online, with the students sitting the exam from their homes, using their own devices. To prevent fraud during the exam, many institutions turned to online proctoring software. This involves recording the sitting and analysing the recordings afterwards for deviations that might indicate irregular behaviour. The use of this type of software immediately raised questions about privacy and potential unjust accusations. But other human values might be impacted as well.

In this paper we address the following research question: *How does the use of the ethical matrix influence the formulation of implementation criteria for proctoring software*? To answer this question, we conducted a case study concerning the implementation of online proctoring software to enable online examination. We carried out a pilot test with teachers evaluating the proctoring software in parallel with conducting a workshop with various stakeholders in which we applied the ethical matrix to the case of online proctoring software. From both the pilot and the workshop we collected implementation criteria, which we then compared.

After providing a theoretical background in the next section, we describe the research method we applied in this study in section 3. In section 4 we present and discuss our results, followed by conclusions in section 5.

2   Theoretical Background

To integrate ethical consideration in the process of design and implementation of digital solutions, we combine the Value Sensitive Design approach (Friedman, Kahn & Borning 2006; Friedman & Hendry, 2019) with the ethical matrix (Mepham, 2000; Mepham et al., 2006; van der Stappen & van Steenbergen, 2020).

2.1   Value Sensitive Design

Value Sensitive Design (VSD) originates from the nineties of the last century (Friedman & Hendry, 2019). It is "a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process" (Friedman, Kahn & Borning, 2006, p. 349). Human value is defined as "what is important to people in their lives, with a focus on ethics and morality (Friedman & Hendry, 2019, p. 4). In VSD not only the values of the actual users of a technological artefact are considered, but also the values of parties that may indirectly be impacted by the artefact. For example, bystanders, future generations, or individuals who cannot or will not use a service. The values of these stakeholders, as well as potential tensions between these values, are investigated from a conceptual, empirical, and technical perspective. At the conceptual level, the relevant stakeholders and values are identified and defined, based on existing literature and knowledge. At the empirical level, the actual perception of these values by the various types of stakeholders is studied by employing methods such as interviews, focus groups or experiments, leading to further elaboration of the values into norms. At the technical level, the values and norms are translated into technical design or implementation criteria or requirements. The three perspectives are iteratively employed. Over the years VSD has been applied to various domains, including the design of browsers (Friedman, Howe & Felten (2002), wind turbines and wind parks (Oosterlaken, 2015) and AI systems (Umbrello & van de Poel, 2021).

In their overview of 20 years of VSD, Friedman & Hendry (2019) discuss 17 instruments and techniques that have been used over the years by various VSD projects. These techniques are either unique to VSD or existing techniques that were adapted to use in VSD. Friedman & Hendry indicate that the list is not exhaustive and that new or newly adapted techniques and instruments are likely to be added over time. They provide, among others, as heuristics for the VSD research and design process to seek an iterative and integrative approach during the entire design process, to use a variety of empirical values-elicitation methods and to continue to elicit stakeholder values throughout the design process as well as apply a value sensitive evaluation process through the deployment phase (Friedman & Hendry, 2019). The ethical matrix as used in this study is an existing instrument adapted to the design and implementation of digital solutions, that can span the entire design, implementation, and deployment process (van der Stappen & van Steenbergen, 2020).

2.2   Ethical matrix

The ethical matrix originates from agriculture and was developed to support rational ethical evaluation of biotechnological innovations in agriculture and food production (Mepham, 2000; Mepham et al., 2006). It was developed to support non-ethicists in discussing the ethical implications of biotechnical innovations. In the rows of the original matrix the relevant stakeholder groups in biotechnology are distinguished (producers, consumers, treated organisms and biota). The columns distinguish the three fundamental ethical principles of autonomy (deontology), fairness (Rawls) and well-being (utilitarianism). When a biotechnical innovation is under consideration, the ethical matrix is used to discuss the impact of the innovation regarding each of the principles on each of the stakeholders. This impact is captured in the cells of the matrix. Figure 1 presents the original ethical matrix.

The ethical matrix is developed for innovation in the food industry. Since its introduction it has been applied and adapted for various other fields (Vinnari, Vinnari & Kupsala, 2017; Schroeder & Palmer, 2003; (Kaiser, Millar, Thorstensen, & Tomkins, 2007; Kermisch & Depaus, 2018; Chatfield, 2018), among which digitalization in education (van der Stappen & van Steenbergen, 2020).

Respect for	**Wellbeing**	**Autonomy**	**Fairness**
Producers	Satisfactory income and working conditions	Managerial freedom	Fair trade laws
Consumers	Safety and acceptability	Choice	Affordability
Treated organisms	Welfare	Behavioural freedom	Intrinsic value
Biota	Conservation	Biodiversity	Sustainability

Figure 1: A generic ethical matrix example (Mepham et al., 2006)

In our study we use the adapted version of the ethical matrix as described in van der Stappen & van Steenbergen (2020). In this adaptation the stakeholders are the direct and indirect stakeholders that are identified in the conceptual perspective of VSD. The ethical principles of the original are replaced by the values as conceptualized in VSD (fig.2). In the cells the potential positive and negative impact of the digital solution on the values of the stakeholders is recorded.

	<Value>	<Value>	…
<Stakeholder>	<Impact>		
<Stakeholder>			
…			

Figure 2: VSD-adapted ethical matrix for digital innovation (van der Stappen & van Steenbergen, 2020)

This version of the ethical matrix can be used to structure and capture a discussion among stakeholders about the potential positive and negative impacts of an intended digital innovation. Examples of its use in this manner are the design of an App supporting students performing preventive health checks (van der Stappen & van Steenbergen, 2020; van Steenbergen et al., 2019) and the design of an App supporting internship coaching to students (van der Stappen & van Steenbergen, 2020).

3   Research Method

To investigate whether the use of the ethical matrix contributes to more value-driven implementation decisions, we conducted a case study at an institution of higher education. The case study concerned the implementation of online proctoring software to enable online examination. The proctoring software concerned makes recordings of the exam sitting of students at home through a webcam and by recording keystrokes. Before starting the exam, the student can be asked to turn her camera around to show the room in which she is taking the exam. During the exam, the software records the students' screens, as well as the students themselves. The images and recordings of the sitting are analysed by an AI algorithm. Any divergent behaviour is reported for further inspection by the examiner. For this to work, the students must install a specific web browser as well as an online proctoring plug-in.

Before implementing the selected online proctoring software, a pilot was conducted with 24 participants (20 teachers, 2 IT professionals, 2 members of the exam committee) who conducted an exam using the proctoring software. The aim of the pilot was to test the usability and effectiveness of the software. Each of the participants answered 12 questions. These included an overall grade for the software, any problems experienced by the participants and the degree of usability and effectiveness the participants attributed to the software. The results of the pilot were translated into implementation criteria which were categorized into requirements, advice, and considerations.

To initiate an ethical discussion about the use of online proctoring and to create awareness about potential undesired consequences, an online workshop was conducted with 10 participants (1 teacher, 1 member of the exam committee, 2 IT professionals, 1 education logistics employee, 1 Digital Learning Environment manager, 2 privacy officers, 2 students). The workshop was led by one of the authors. The aim of the workshop was to identify risks and opportunities of online proctoring for various stakeholders. The ethical matrix was used to structure the discussion. The workshop started with an ethical matrix that already contained the main stakeholders and values. These were identified from literature and earlier discussions with experts (conceptual perspective of VSD). As starting point for the stakeholders, we identified the primary people involved in the processes of

preparing, conducting, and evaluating examinations, both on site (as done before the pandemic) and online using the online proctoring software.

The values were selected from lists of values relevant to digitalization (Friedman, Kahn & Borning, 2006; Royakkers et al., 2018), which were compared to views expressed in online posts and publications about the use of online proctoring. The values thus extracted were discussed with IT experts, a teacher, and a student from a different institute of higher education. In the workshop the stakeholders and values identified were validated and the potential impact of the online proctoring software on the identified values for the identified stakeholders was discussed. This was done via identifying potential harms and benefits of using the proctoring software. After the workshop, the results were laid down in a report which was validated by the participants. After validation by the participants, the authors translated the results into implementation criteria. These criteria, too, were categorized into requirements, advice, and considerations.

The implementation criteria of both the pilot and the workshop were combined into one list of 39 criteria. From the list four types of criteria emerged: criteria concerned with facilitation (4), instruction and procedures (22), fraud and reputation (10) and logistics (3).

To analyse the contribution of the ethical matrix, we compared the criteria that resulted from the workshop with the criteria that resulted from the pilot.

4 Results

The average grade given by the pilot participants to the proctoring software was a 7. Problems reported concerned mainly technical problems with installing the required browser or plug-in. Most of the participants concluded that use of the software would be feasible, if necessary, though a few participants doubted its usefulness to detect all fraud. One participant expressed concerns about privacy and other ethical considerations. Based on the pilot 17 implementation criteria were formulated.

In the ethical workshop we started with a matrix containing the stakeholders student, examiner, surveillant, educational institute, programme manager and IT department and the values equality, well-being, reputation, autonomy, privacy, sustainability, and trustworthiness. In the workshop the stakeholder of housemate was added, while

the value of sustainability did not generate response from the participants and was subsequently removed from the matrix. Table 1 contains descriptions of the values as well as examples of impact on one type of stakeholder, the student. In a similar manner impacts on the other stakeholders were formulated.

Table 1: Values impacted by online proctoring software

Value	Description	Potential impact on student (examples)
Equality	Equal opportunity to successfully complete the exam	Differences in housing, physical disabilities, differences in available internet connection or hardware.
Well-being	Material and immaterial contentment	Unease or stress from being observed and recorded, worries about identification requirements
Reputation	How one is regarded by others	Unjust accusation of fraud
Autonomy	The degree to which persons can make their own choices in line with their being	Uncertainty about consequences of refusing online proctoring, mandatory installation of specific software
Privacy	The right to keep certain parts of ones live (such as ideas, data, or personal circumstances) to oneself	Exposure of personal living sphere, risk of data breach
Trustworthiness	The value of the exam result, the reliability of the proctoring	Fear of exam result being considered less trustworthy by outside world, lack of trust in fraud detection process

Based on the workshop, 25 implementation criteria were formulated to mitigate the potential negative impacts.

We divided the criteria from both sources into three categories: requirements, advice, and considerations. Requirements are criteria that are considered hard prerequisites for implementation. They are not negotiable. Advice includes criteria that are strongly recommended, but not mandatory to proceeding. Considerations

are criteria that are considered beneficial but are left to the implementation project to decide on. Examples of each category, one originating from the ethical matrix workshop and one originating from the pilot can be found in Table 2.

Table 2: Examples of implementation criteria

Type	Criterium	Source
Requirement	Students are given explicit and clear instructions for installing all required software	Pilot
Requirement	Students without suitable hardware (laptop) are provided with a laptop by the institute	Ethical matrix
Advice	Have students check all equipment beforehand	Pilot
Advice	Think about how to support examiners who also need to act as surveillant, because of an expected increase in workload	Ethical matrix
Consideration	Concerns are about the privacy aspects of the mandatory browser	Pilot
Consideration	The reputation of students may be damaged if they are unjustly accused of fraud and records of the accusation are kept.	Ethical matrix

We categorized the criteria into four categories: criteria concerned with facilitation, with instruction and procedures, with fraud and reputation, and with logistics. Table 3 shows the distribution of criteria from the two sources over the categories.

Table 3: Distribution of criteria over categories

Category	Number of criteria from pilot	Number of criteria from workshop
Facilitation	-	4
Instruction and procedures	12	12
Fraud and reputation	4	7
Logistics	2	2

A total of 43 implementation criteria were derived from the pilot and workshop together, with an overlap of 4 criteria that emerged from both the pilot and the workshop. Leaving 39 distinct criteria.

Analysis of the two sets of criteria reveals a clear difference in focus between the pilot and the workshop. As expected, the criteria from the pilot were more function-oriented, whereas the criteria from the workshop were more value-oriented.

Thus, only the workshop led to criteria regarding the facilitation of students who do not have access to the required hardware or to a suitable space to take the exam (related to the value of equality) and the facilitation of examiners who experience a sudden increase of workload because of the application of online proctoring software (related to the value of well-being).

As for the category of instruction and protocol we found that the criteria from the pilot are focused on providing clear instructions to both students and employees regarding all phases of the examination process, ranging from timely preparation and testing of the technology beforehand to sitting the exam as well as the careful closure of the sitting. The criteria from the ethical workshop are focused on augmenting the protocol with protective measures for students, such as safe online identification, informed consent, right of inspection, dealing with physical disabilities and technical incidents during the exam sitting (related to the values of well-being and privacy). In addition, the workshop led to criteria concerning the long-term effects and feasibility of the online proctoring solution (related to the value of autonomy).

In the category of fraud and reputation, the criteria from the pilot dealt with the fraud analysis effectiveness. The criteria from the workshop dealt with the risk to the reputation of both students (incorrect signalling by the algorithm of potential fraud) and institute (mistakes in the process, reduced perceived value of exam result, privacy breach).

Finally, in the category of logistics, the pilot led to criteria concerning the suitability of online proctoring software for various types of exam, whereas the workshop focused on the feasibility of the entire process of online proctoring (value of well-being).

5   Conclusion

In this study we investigated whether the use of the ethical matrix as adapted by Van der Stappen & van Steenbergen (2020) enriched the outcomes of a functional pilot concerning the formulation of implementation criteria of online proctoring software. We expected that the explicit focus of the matrix on the values of various stakeholders would generate additional criteria. The analysis of the two lists of criteria generated from the pilot on the one hand and the workshop using the ethical matrix on the other hand, confirmed that the two approaches lead to different types of criteria.

We conclude that the pilot and the ethical session are complementary. The pilot led to implications focused on function, whereas the ethical session provided insight into value-oriented requirements. We believe that in educational institutes value and function are equally important. By allocating a workshop to formulating ethical requirements and considerations early in the process, the importance of both function and values can be considered during the implementation. The ethical matrix appears to be a very useful instrument in facilitating and structuring discussions on values by non-ethicists such as educators and students.

Our study concerns only one case which of course limits its potential for generalization. We believe, however, that the results are promising. Increased application of the ethical matrix in a diversity of contexts will hopefully lead to more comparative analyses in the vein of our study. Besides providing increasing insight in the effects of applying the ethical matrix, we are hopeful that it will also contribute to implementations that are more sensitive to the values of all stakeholders concerned. We intend to study how the ethical matrix can also be used to test this, by applying it again after having conducted online proctoring for some time, as proposed in Van der Stappen & Van Steenbergen (2020).

We believe that the use of the ethical matrix might add the dimension of impact to the widely accepted dimensions of functional and non-functional requirements in digital application.

Acknowledgements

References

Chatfield, K. (2018). An Ethical Matrix for Traditional and Complementary Medicine. https://doi.org/10.1007/978-3-030-05300-0_3

Friedman, B., & Hendry, D. G. (2019). Value Sensitive Design: Shaping Technology with Moral Imagination. The MIT Press.

Friedman, B., Howe, D. C., & Felten, E. (2002, January). Informed consent in the Mozilla browser: Implementing value-sensitive design. In Proceedings of the 35th annual hawaii international conference on system sciences (pp. 10-pp). IEEE.

Friedman, B., Kahn, P. H., & Borning, A. (2006). Value Sensitive Design and Information Systems. In D. F. Galletta & P. Zhang (Eds.), Human-computer interaction and management information systems: Foundations (pp. 348–372). https://doi.org/10.1002/9780470281819.ch4

Mepham, B. (2000). A framework for the ethical analysis of novel foods: The ethical matrix. Journal of Agricultural and Environmental Ethics, 12(2), 165–176. https://doi.org/10.1023/A:1009542714497

Mepham, B., Kaiser, M., Thorstensen, E., Tomkins, S., & Millar, K. (2006). Ethical Matrix Manual.

Kaiser, M., Millar, K., Thorstensen, E., & Tomkins, S. (2007). Developing the Ethical Matrix as a decision support framework: GM fish as a case study. Journal of Agricultural and Environmental Ethics, 20, 65–80. https://doi.org/10.1007/s10806-006-9023-8

Kermisch, C., & Depaus, C. (2018). The Strength of Ethical Matrixes as a Tool for Normative Analysis Related to Technological Choices: The Case of Geological Disposal for Radioactive Waste. Science and Engineering Ethics, 24(1), 29–48. https://doi.org/10.1007/s11948-017-9882-6

Oosterlaken, I. (2015). Applying value sensitive design (VSD) to wind turbines and wind parks: An exploration. Science and engineering ethics, 21(2), 359-379.

Royakkers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. Ethics and Information Technology, 20(2), 127-142.

Schroeder, D., & Palmer, C. (2003). Technology assessment and the "ethical matrix." Poiesis & Praxis, 1(4), 295–307. https://doi.org/10.1007/s10202-003-0027-4

Stibe, A., & Cugelman, B. (2016, April). Persuasive backfiring: When behavior change interventions trigger unintended negative outcomes. In International conference on persuasive technology (pp. 65-77). Springer, Cham.

Umbrello, S., & van de Poel, I. (2021). Mapping value sensitive design onto AI for social good principles. AI and Ethics, 1-14.

Van den Hoven, J. (2017). Ethics for the digital age: Where are the moral specs? In Informatics in the Future, pages 65–76. Springer International Publishing, Cham.

Van der Spoel, I., Noroozi, O., Schuurink, E., & van Ginkel, S. (2020). Teachers' online teaching expectations and experiences during the Covid19-pandemic in the Netherlands. European journal of teacher education, 43(4), 623-638.

Van der Stappen, E., & van Steenbergen, M. (2020). The Ethical Matrix in Digital Innovation Projects in Higher Education.

Vinnari, M., Vinnari, E., & Kupsala, S. (2017). Sustainability Matrix: Interest Groups and Ethical Theories as the Basis of Decision-Making. Journal of Agricultural and Environmental Ethics, 30(3), 349–366. https://doi.org/10.1007/s10806-017-9670-y

Health Information Literacy: The Saving Grace During Traumatic Times

Maedeh Ghorbanian Zolbin, Khadijah Kainat & Shahrokh Nikou

Åbo Akademi University, Faculty of Social Sciences, Business and Economics, Turku, Finland; e-mail: maedeh.ghorbanianzolbin@abo.fi, khadijah.kainat@abo.fi, shahrokh.nikou@abo.fi

Abstract When it comes to engaging with health information in their daily lives, people face different challenges. In the context of COVID-19, the aim of this study is to determine whether health information literacy can assist people in making informed health-related decisions. An empirical study was conducted to investigate such an effect. Building on a dataset composed of 155 respondents, the research model was examined through structural equation modelling. The results showed that health information literacy – as an individual ability – not only influences health decision making but also has a direct impact on the awareness of the challenges imposed by the current pandemic situation. In addition, the results show that too much information leads to information fatigue, and consequently negatively impacts decision making. The findings of this paper show that the concept of health information literacy should be understood and developed separately from the health literacy concept. Theoretical contributions and practical implications are discussed.

1   Introduction

Being abreast of health and engagement in the health-related decision-making process is highly expected in the digital age and in digital health systems (Brabers et al., 2017). Some reasons – such as noticeable changes in the healthcare domain, a revolution in information and communication technologies (ICTs), the rapid distribution of health information, the worldwide expansion of the internet, and the availability of a tremendous amount of health-related information – have led to a higher level of patient confidence in managing their own health (Cullen, 2005; Weinhold & Gastaldi, 2015). To increase the quality and safety of the healthcare systems, the role of patient decision making, as a crux of patient-centred care, seems very critical (Godolphin, 2009). As opined by Seymour (2018), patient involvement in decision making enhances the attainment of favourable health outcomes (Seymour, 2018). However, this is not possible if patients lack an understanding of healthcare practices and issues. When patients lack knowledge and understanding, it becomes difficult for them to make appropriate health decisions independently, since they have nothing to offer (Rodríguez et al., 2013). This emphasises the significance of health information literacy (hereinafter HIL) skills, which enable individuals to enhance their understanding of healthcare issues. Individuals need to be able to understand healthcare processes, needs, and requirements in order to make informed health decisions (Cui & Chang, 2020). Additionally, the COVID-19 pandemic is exposing dysfunction and fragility in healthcare services. As such, it is critical to have the ability and required skills to find, evaluate, and use health-related information from the internet in order to make independent health-related decisions. In the case of the COVID-19 pandemic situation and in order to control the consequences, it is vital for individuals to develop a solid foundation of information and knowledge about the situation (Al-Dossary et al., 2020). Additionally, in the past, rational decisions were made based on all the available information. As such, one of the most critical challenges involved in making a decision was a lack of access to adequate information, while today, there is often far more information available than is required. With the growth in information technology and the internet, it is not surprising that patients now have access to an unlimited amount of health information. However, this overload of information adversely affects the process of independent decision making (Buchanan & Kock, 2001). The issue of information overload in the context of health has become exacerbated, owing to the expanding availability of relevant information, especially through online sources (Khaleel et al.,

2020). Once patients become overloaded from processing health-related content on the internet, they are likely to feel fatigued (Cao & Sun, 2018), which negatively affects their decision-making process. It seems that to make appropriate health decisions and to be able to take care of themselves within the digital health environment, patients need to have special skills and abilities, such as HIL (Krist et al., 2017).

All in all, this study seeks to investigate how people's health-related decision making, as a consequence of the global pandemic imposed by COVID-19, is influenced by their HIL skills and their awareness of COVID-19–related issues. The research question guiding this study is, "*To what extent do people's HIL skills impact their health-related decision making, and what is the role of information overload and information fatigue in the process of making appropriate decisions?*".

2 Literature Review and Hypothesis Developments

2.1 Health Information Literacy and Patient Decision Making

The combination of health literacy and information literacy leads to a new concept known as health information literacy. HIL refers to "the set of abilities needed to recognise a health information need, identify information sources and use them to retrieve relevant information, assess the quality of the information and its applicability to a specific situation, analyse, understand, and use the information to make appropriate health decisions" (Shipman, 2009, p. 294). HIL is very important for people since it helps them to understand how well they can take care of themselves and make decisions regarding issues concerning their health (Eriksson-Backa et al., 2012). HIL skills may become increasingly important when the internet is the main source of finding information. The authors highlight the role of HIL skills in making appropriate health decisions. It is acknowledged that information-literate people are able to make appropriate decisions regarding issues concerning their health (Cui & Chang, 2020). Additionally, Krist et al. (2017) highlighted the positive role of being information literate in making independent health decisions; therefore, we propose the following:

H1: *Health information literacy has a positive effect on patient decision making*

2.2 COVID-19 Awareness and Health Information Literacy

As pointed out by Shehata (2020), dealing with COVID-19 requires HIL skills. In fact, when people are more health information literate, they can find and acquire more information and knowledge about COVID-19. The higher the level of HIL skills, the easier it is to gain awareness in the context of coronavirus. In other words, lack of HIL skills leads to a lesser awareness of COVID-19–related symptoms. The main reason is that such individuals are not able to find relevant information from reliable sources, such as reliable medical webpages (McCaffery et al., 2020). Consequently, a better grasp of the existing situation leads to better health decision making; hence, we propose the following:

H2: *Health information literacy has a positive effect on COVID-19 awareness*

H3: COVID-19 awareness has a positive effect on patient decision making

2.3 Information Overload and Health Information Literacy

In the process of decision making, an ever-increasing amount of information is one of the main challenges (Falschlunger et al., 2016), especially when it comes to health information. The constant increase in the amount of information makes it difficult to organise and find high-quality information effectively. According to Ruff (2002), "Once capacity is surpassed, additional information becomes noise and results in a decrease in information processing and decision quality. Having too much information is the same as not having enough". Information overload refers to too much data, which is received too fast for a person of average cognitive ability to absorb and process (Zhang et al., 2020). This concept is defined by Wilson (2001, p. 113) as "a perception on the part of the individual (or observers of that person) that the flow of information associated with work tasks is greater than can be managed effectively and a perception that overload in this sense creates a degree of stress for which his or her coping strategies are ineffective". Therefore, people need special skills, such as HIL skills, to deal with massive amounts of information (Kurelović et al., 2016). Jiang and Beaudoin (2016) emphasised the importance of HIL skills in coping with information overload. They argued that health information literate people are able to combat overload. These studies also highlighted that individual

who are not adroit at dealing with large amounts of information may experience overload. The higher the level of HIL skills in an individual, the better she or he can cope with information overload; hence, we propose the following:

H4: *Health information literacy has a negative effect on information overload*

2.4   Information Fatigue and Information Overload

Information fatigue is defined as the tendency for information users to back away from seeking more information when they become overwhelmed with too many pieces of content, too many contacts, too many information sources (such as websites), and too much time spent keeping up with these connections (Bright et al., 2015). Fatigue is more than feeling tired and drowsy. Thomas (1998) stated that when patients are inundated with a massive amount of health information, they may make more mistakes and misunderstand communication. In fact, an overwhelming amount of information leads to anxiety, exhaustion, and other symptoms of information fatigue. Once patients become overloaded from processing too much health-related content on the internet, they are likely to feel fatigued (Cao & Sun, 2018); hence, we propose the following:

H5: *Information overload has a positive effect on information fatigue*

2.5   Information Fatigue and Patient Decision Making

Prior studies have indicated that in the context of finding health-related information, information overload results in information fatigue, which has a diverse effect on patient decision making (Cao & Sun, 2018), since exhausted individuals will not be able to focus well and make informed decisions. Human behaviours can be altered owing to the exhaustion that results from overload. People can make hasty decisions or shut down when they have information fatigue, or their ability to make informed decisions can ultimately be depleted. As such, this may negatively impact their health-making decisions; hence, we propose:

H6: Information fatigue has a negative effect on patient decision making

2.6   Patient Decision Making

The growing number of studies that seek to identify effective factors relating to patients' decision-making abilities indicated that this topic is of interest to professionals and scholars in various research fields (Brabers et al., 2017). Patients' ability to make health decisions results in various positive outcomes, such as improved health outcomes, mortality reductions, improved satisfaction with the care experience, and reduced costs. In other words, the active engagement of patients in the process of making health decisions fosters their health and well-being and makes it possible for patients to live longer (Krist et al., 2017). The core dimension of a quality modern health service is patient-centred care. Active participation in the process of making health decisions demands well-informed patients. They should be able to access up-to-date information about their care, and the potential outcomes of their treatment. Additionally, a layer of vigilance and protection against errors can be created by making informed decisions (Sketcher-Baker, 2017). Patient decision making can be affected by various factors, and this study aims to focus on some of these. In summary, we argue that patient decision making is impacted not only by traditional social media factors such as information overload and information fatigue but also by individual-level HIL and contextual (e.g. COVID-19 awareness) related factors that equally impact the health-related decision–making process (Figure 1).
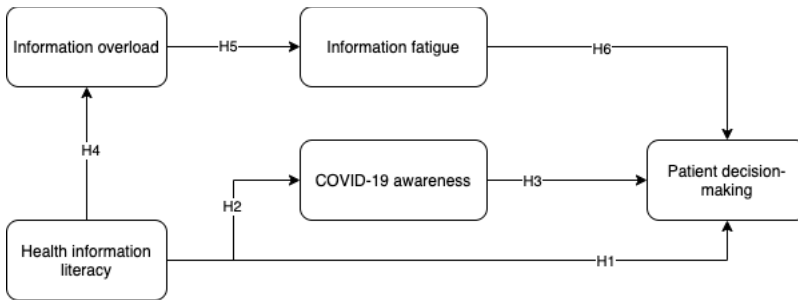
Figure 1: Proposed Conceptual Model

3   Research Methodology

3.1   Instrument and Data Collection

All items for measuring constructs were derived from previously validated measures, and if needed, some changes were made to fit the study context. Items for measuring COVID-19 awareness (5 items) were derived from Alea et al. (2020, p. 134-136) and McCaffery et al. (2020). Items for measuring information overload (12 items) and information fatigue (8 items) were derived from Whelan et al. (2020) and Norman and Skinner (2006), respectively. Items for measuring HIL (7 items) were derived from Norman and Skinner (2006). Finally, we used 5 items from Seo et al. (2016) to measure patient decision making. We used an online survey to collect data during the time of the COVID-19 pandemic: March 2021. We sent more than 230 invitations, using different channels, such as authors' personal networks and social media platforms. In total, we received 155 complete responses.

4   Results

4.1   Descriptive Results

The respondents' ages fell within the ranges of 18–25 years (12.3%), 26–35 years (47.7%), 36–45 years (20.6%), 46–55 years (18.1%), and over 55 years (1.3%). The sample comprised 78 females, 67 males, and 10 subjects who chose not to reveal their gender. The educational level of the respondents was as follows: high school diploma (n = 6), bachelor's degree (n = 32), master's degree (n = 85), and PhD degree (n = 29), and 3 indicated other educational attainments. With regard to the current occupational status of the respondents, most were students (n = 69), 12 respondents were employed, 13 were non-employed, 54 were self-employed, and 7 held some other occupation as their occupation. In the sample, there were 64 Finnish and 91 non-Finnish respondents. Moreover, 144 respondents lived in Finland, and 11 lived in another country. When we asked the respondents to indicate the social networking sites (SNSs) they used to access health-related information, they reported the following: the use of Facebook was mentioned by 91, Instagram by 70, Twitter by 41, and Telegram by 39 respondents, whereas the least-used SNSs were Snapchat (n = 11) and TikTok (n = 12). We also asked the participants to indicate how much time they spend per day searching and sharing health-related information on SNSs. They reported (less than 30 minutes, n = 73), (from 30

minutes to one hour, n = 42), (from one to two hours, n = 19), and (more than two hours, n = 9), and interestingly, 12 respondents indicated that they do not use SNSs for health-related information.

4.2   Measurement and Conceptual Results

PLS-SEM (partial least squares structural equation modelling) was used to investigate path relationships in the proposed conceptual model. The results showed that many items had sufficient factor loadings above the threshold value of 0.70. In total, we used 37 items to measure the five constructs and retained 27 items for further analysis. All internal reliability and validity assessments, i.e. Cronbach's alpha (α), and composite reliability (CR), for all constructs were in the range of the recommended values of 0.70 and 0.70, respectively (Hair et al., 2019) (see Table 1). The highest CR value was for information fatigue (0.89) and the lowest was for patient decision making (0.82). Also, the highest value for Cronbach's alpha (α) was for information fatigue (0.90), and the lowest value was for patient decision making (0.87). As shown in Table 2, all the average value extracted (AVE) values were above the recommended threshold of 0.50 (Hair et al., 2019), such that the highest was for patient decision making (0.80) and the lowest was for HIL (0.68).

Table 1: Descriptive Statistics

Construct	Item	Loadings	Mean	Std.	(α)	CR	AVE
COVID_19 awareness	COV_A 1	0.87	4.63	0.68			
	COV_A 2	0.86	4.48	0.79			
	COV_A 3	0.87	4.36	0.86	0.89	0.85	0.75
	COV_A 4	0.86	4.34	0.96			
	COV_A 5	0.88	4.48	0.87			
Patient decision making	DECM3	0.88	2.93	1.36			
	DECM4	0.90	2.81	1.42	0.87	0.82	0.80
	DECM5	0.89	1.92	1.30			
	HIL1	0.81	3.46	1.31	0.88	0.87	0.68

Health information literacy	HIL2	0.86	3.32	1.29			
	HIL3	0.84	3.39	1.25			
	HIL5	0.83	3.01	1.27			
	HIL6	0.74	3.36	1.28			
	HIL7	0.87	3.39	1.18			
Information fatigue	INFA2	0.83	3.29	1.30			
	INFA4	0.86	3.12	1.29			
	INFA5	0.88	2.95	1.36	0.90	0.89	0.79
	INFA6	0.92	3.06	1.33			
	INFA7	0.92	3.06	1.30			
	INFA8	0.93	3.01	1.28			
Information overload	INOV1	0.91	3.06	1.21			
	INOV12	0.77	2.94	1.26			
	INOV2	0.91	3.02	1.30			
	INOV3	0.88	3.03	1.32	0.84	0.85	0.73
	INOV4	0.91	3.11	1.36			
	INOV5	0.89	2.97	1.33			
	INOV9	0.73	2.91	1.21			

4.3   Discriminant Validity

To establish the discriminant validity, we examined the AVE scores, and all AVE values were lower than the shared variance for all model constructs; therefore, the discriminant validity was established in this research (Fornell & Larcker, 1981).

Table 2: Discriminant validity (Fornell and Larcker criterion)

Constructs	CAWA	HIL	INFA	INOV	PDM
COVID 19 challenges	**0.87**				
Health information literacy	0.32	**0.83**			
Information fatigue	-0.01	-0.20	**0.89**		
Information overload	-0.05	-0.18	0.57	**0.85**	
Patient decision making	0.15	0.53	-0.27	-0.22	**0.89**

4.4   Structural Results

The structural results showed that the independent variable, i.e. patient decision making, has been explained by a variance of 38%. The predictor variables, i.e. information fatigue and COVID-19 awareness, have been explained by variance of

36% and 12%, respectively. The SEM results showed that HIL had a direct and strong positive effect ($\beta = 0.57$; $t = 6.001$; $p = 0.001$) on patient decision making; thus, H1 was accepted in the model. The results also provided theoretical support for H2, where it was expected that HIL positively affects COVID-19 awareness. The SEM results showed that this path was significant ($\beta = 0.35$; $t = 4.211$; $p = 0.001$). Moreover, the results showed that HIL had a significant and – as we expected – negative impact on information overload ($\beta = -0.19$; $t = 2.039$; $p = 0.05$). This indicates that the respondents with higher HIL are more capable of distinguishing between true and fake information; thus, they are less impacted by indecisiveness in regard to health-related information overload. Hence, H4 was supported in the model. However, the results did not provide theoretical support for H3, where we postulated that COVID-19 awareness affects patient decision making. The SEM results showed that, as we predicted, information overload had a positive and significant effect ($\beta = 0.60$; $t = 9.144$; $p = 0.001$) on information fatigue, thus providing theoretical support to accept H5. Similarly, the SEM results showed information fatigue ($\beta = -0.17$; $t = 2.006$; $p = 0.05$) had a significant influence on patient decision making, but the effect, as expected, was negative. Thus, H6 was supported by the model.
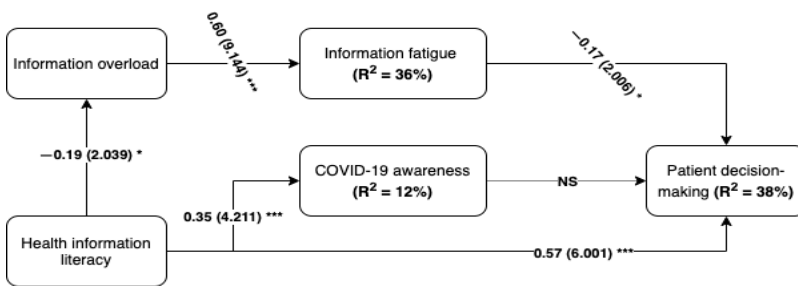
Figure 2: Structural model results. Note*: *p < 0.05. **p < 0.01. ***p < .001

5   Discussions and Conclusion

This study focuses on the relationship between health information literacy and patient decision making in the context of the COVID-19 pandemic, examining how patient utilises digital health services. The COVID-19 pandemic has severely limited patients' access to health information, as the traditional methods used to acquire information – like face-to-face visits with healthcare providers – are no longer an

option. Therefore, digital health services and digital health service platforms seem far more suitable (Gálvez et al., 2020). However, using these kinds of digital services to find the required information demands special skills. As highlighted by Seo et al. (2016), essential skills – such as the ability to seek, find, evaluate, and use health information from digital platforms – may affect patients' decision-making preferences. The results of this study are consistent with Seo et al.'s (2016) findings. In the present research we show that individuals with higher HIL skills may have a considerable advantage when seeking health information from digital sources, and may therefore be able to make more informed health decisions, when faced with the limitations (e.g. having face-to-face interactions with doctors) imposed by COVID-19 restrictions. In addition, the results reveal that a higher level of HIL skill leads to an improvement in individuals' consciousness regarding COVID-19. HIL skills augment people's knowledge regarding the seriousness of the epidemic and the dangers it poses, along with possible preventive strategies. Since most scientific and reliable information regarding COVID-19 is uploaded to digital platforms, we argue that to take advantage of the available information, HIL is the saving grace during this traumatic time.

However, such pure knowledge and skills do not in themselves lead to the ability to make an independent and appropriate health decision during challenging times such as the COVID-19 pandemic, as these kinds of decisions demand a higher level of medical and pharmaceutical knowledge. In addition, digital health platforms may act like a double-edged sword. While these platforms provide individuals with noteworthy health information, the huge amount of supplied health information can lead to health information overload. The results of this research show that patients' decision making can be retrogressed by the exhaustion that results from being overloaded with information from online channels. These results are consistent with those obtained by Cao and Sun (2018). Based on our own findings, HIL skills empower people to control their uptake of the available information and to cope with the overload of health information. This means that higher HIL skills lead to a lesser overload of health information. This result is consistent with the results obtained by Jiang and Beaudoin (2016). This study provides some practical implications. too. For example, a patient's decision-making ability – when viewed in the light of high HIL skills – reduces the burden on the healthcare providers and shares the risk of decision making between patients and healthcare professionals. Additionally, the ability to make informed health decisions reduces the time patients

spend accessing physical health services, such as emergency rooms, which benefits both governments and patients. So, empowering individuals with HIL skills is an acceptable solution to facilitate an informed decision-making process in regard to digital health systems. This paper has some limitations. The findings may not be applicable in another context, as we studied decision making in the context of health-related issues during the COVID-19 pandemic. Thus, we cannot claim that the results can be generalised to other contexts. In addition, this study does not consider age or education distribution in its multi-group analysis. Moreover, we conceptualised that fatigue, as a result of information overload, negatively impacts health-related decision making, while recognising that other variable, such as individual characteristics or peer pressure, can also affect fatigue.

References:

Al-Dossary, R., Alamri, M., Albaqawi, H., Al Hosis, K., Aljeldah, M., Aljohan, M., Almadani, N., Alrasheadi, B., Falatah, R., & Almazan, J. (2020). Awareness, Attitudes, Prevention, and Perceptions of COVID-19 Outbreak among Nurses in Saudi Arabia. International Journal of Environmental Research and Public Health, 17(21), 8269.

Alea, L. A., Fabrea, M. F., Roldan, R. D. A., & Farooqi, A. Z. (2020). Teachers' Covid-19 Awareness, Distance Learning Education Experiences and Perceptions towards Institutional Readiness and Challenges. International Journal of Learning, Teaching and Educational Research, 19 (6), 127-144.

Brabers, A. E., Rademakers, J. J., Groenewegen, P. P., Van Dijk, L., & De Jong, J. D. (2017). What role does health literacy play in patients' involvement in medical decision-making? PLOSONE, 12 (3).

Bright, L. F., Kleiser, S. B., & Grau, S. L. (2015). Too much Facebook? An exploratory examination of social media fatigue, Computers in Human Behaviour, 44, 148-155.

Buchanan, J., & Kock, N. (2001). Information overload: A decision making perspective. In Multiple criteria decisions making in the new millennium (pp. 49-58). Springer, Berlin, Heidelberg.

Cao, X., & Sun, J. (2018). Exploring the effect of overload on the discontinuous intention of social media users: An S-O-R perspective. Computers in Human Behaviour, 81, 10–18.

Cui, X., & Chang, C. T. (2020). Distribution Dynamics, Regional Differences, and Convergence of Elderly Health Levels in China. Sustainability, 12 (6), 2288.

Cullen, R. (2005). Empowering patients through health information literacy training. Library Review, 54 (4), 231- 244.

Eriksson-Backa, K., Ek, S., Niemelä, R., & Huotari, M.-L. (2012). Health information literacy in everyday life: A study of Finns aged 65–79 years. Health Informatics Journal, 18(2), 83–94.

Falschlunger, L., Lehner, O., & Treiblmaier, H. (2016). InfoVis: The Impact of Information Overload on Decision Making Outcome in High Complexity Settings. SIGHCI 2016 Proceedings. 3.

Fornell, C. G., & Larcker. D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50.

Gálvez, A., Romero, B., Trigo, S., & Serrano, M. (2020). Older people, dependency, and vulnerability in the coronavirus pandemic: emergence of social and health integration. Enfermería Clinica, 31, S18-S23. DOI: 10.1016/j.enfcli.2020.05.004.

Godolphin, W. (2009). Shared Decision-Making. Healthcare Quarterly, 12(Spec No Patient), e186-e190. doi:10.12927/hcq.2009.20947.

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

Jiang, S., & Beaudoin, C. E. (2016). Health literacy and the internet: An exploratory study on the 2013 HINTS survey. Computers in Human Behaviour, 58, 240–248.

Khaleel, I., Wimmer, B. C., Peterson, G. M., Zaidi, S. T. R., Roehrer, E., Cummings, E., & Lee, K. (2020). Health information overload among health consumers: A scoping review. Patient Education and Counselling, 103(1), 15-32.

Krist, A. H., Tong, S. T., Aycock, R. A., & Longo, D. R. (2017). Engaging patients in decision-making and behaviour change to promote prevention. Information Services & Use, 37 (2), 105-122.

Kurelović, E. K., Tomljanović, J., & Davidović, V. (2016). Information overload, information literacy and use of technology by students. International Journal of Educational and Pedagogical Sciences, 10(3), 917-921.

McCaffery, K. J., Dodd, R. H., Cvejic, E., Ayre, J., Batcup, C., Isautier, J. M. J., Copp, T., Bonner, C., Pickles, K., Nickel, B., Dakin, T., Cornell, S., & Wolf, M. S. (2020). Health literacy and disparities in COVID-19–related knowledge, attitudes, beliefs, and behaviours in Australia. Public Health Research & Practice, 30 (4): e30342012.

Norman, C. D., & Skinner, H. A. (2006). eHEALS: the eHealth literacy scale. Journal of medical Internet research, 8(4), e27.

Rodríguez, V., Andrade, A. D., García-Retamero, R., Anam, R., Rodríguez, R., Lisigurski, M., Shari, J., & Ruiz, J. G. (2013). Health Literacy, Numeracy, and Graphical Literacy Among Veterans in Primary Care and Their Effect on Shared Decision Making and Trust in Physicians. Journal of Health Communication, 18(sup1), 273–289.

Ruff, J. (2002). Information Overload: Causes, Symptoms and Solutions. (pp. 1-13). Harvard Graduate School of Education's Learning Innovations Laboratory (LILA).

Seo, J., Goodman, M. S., Politi, M., Blanchard, M., & Kaphingst, K. A. (2016). Effect of Health Literacy on Decision-Making Preferences among Medically Underserved Patients. Medical Decision Making, 36(4), 550–556.

Seymour, J. (2018). The Impact of Public Health Awareness Campaigns on the Awareness and Quality of Palliative Care. Journal of Palliative Medicine, 21, S30-S36.

Shehata, A. (2020). Health Information behaviour during COVID-19 outbreak among Egyptian library and information science undergraduate students. Information Development, 026666692097618. doi:10.1177/0266666920976181.

Shipman, J. P., Kurtz-Rossi, S., & Funk, C. J. (2009). The Health Information Literacy Research Project. Journal of Medical Library Association, 97 (4), 293–301.

Sketcher-Baker, K. (2017). Guide to Informed Decision-making in Health Care. Second edition, Australia. From: https://www.health.qld.gov.au/__data/assets/pdf_file/0019/143074/ic-guide.pdf.

Thomas, S. P. (1998). Editorial: Information Fatigue Syndrome - Is there an Epidemic? Issues in Mental Health Nursing, 19 (6), 523–524.

Weinhold, I., & Gastaldi, L. (2015). From shared decision making to patient engagement in health care processes: The role of digital technologies. In Challenges and opportunities in health care management (pp. 185-196). Springer, Cham.

Whelan, E., Islam, N., & Brooks, S. (2020). Is boredom proneness related to social media overload and fatigue? A stress–strain–outcome approach. Emerald Insight, 30 (3), 869-887.

Wilson, T. D. (2001). Information overload: implications for healthcare services. Health Informatics Journal, 7 (2), 112-117.

Zhang, X., Ding, X., & Ma, L. (2020). The influences of information overload and social overload on intention to switch in social media. Behaviour & Information Technology, 1–14

CREATING A TAXONOMY OF BUSINESS MODELS FOR DATA MARKETPLACES

MONTIJN VAN DE VEN,[1] ANTRAGAMA EWA ABBAS,[2]
ZENLIN KWEE[2] & MARK DE REUVER[2]

[1] Eindhoven University of Technology, Department of Industrial Engineering and Innovation Sciences, The Netherlands; e-mail: m.r.v.d.ven@tue.nl
[2] Delft University of Technology, Faculty of Technology, Policy and Management, The Netherlands; e-mail: a.e.abbas@tudelft.nl, z.roosenboom-kwee@tudelft.nl, g.a.dereuver@tudelft.nl

Abstract Data marketplaces can fulfil a key role in realizing the data economy by enabling the commercial trading of data between organizations. Although data marketplace research is a quickly evolving domain, there is a lack of understanding about data marketplace business models. As data marketplaces are vastly different, a taxonomy of data marketplace business models is developed in this study. A standard taxonomy development method is followed to develop the taxonomy. The final taxonomy comprises of 4 meta-dimensions, 17 business model dimensions and 59 business model characteristics. The taxonomy can be used to classify data marketplace business models and sheds light on how data marketplaces are a unique type of digital platforms. The results of this research provide a basis for theorizing in this rapidly evolving domain that is quickly becoming important.

1   Introduction

As an organization may not always possess the required data to carry out or improve their processes and services, they may wish to purchase these data from other organizations. A data marketplace can enable data purchase by providing a digital platform through which individuals and organizations can exchange data (Stahl et al., 2016; Schomm et al., 2013). In contrast to most other platforms, where data is utilized to improve services or manage customer relationships, on data marketplaces data is actually the product itself (Spiekermann et al., 2018).

Despite the potential benefits of data marketplaces, in practice, very little data is shared or traded via these platforms (Koutroumpis et al., 2020). In general, little research has been conducted on data marketplaces (Thomas & Leiponen, 2016) and data marketplace business models in particular (Fruhwirth et al., 2020; Spiekermann, 2019). As a foundation for research on a novel and diverse phenomenon, a first step is developing a taxonomy, because it can be used to classify data marketplace business models (Lambert, 2015). Two taxonomies of data marketplace business models are currently available in the literature, i.e. those proposed by Fruhwirth et al. (2020) and Spiekermann (2019) respectively.

The existing taxonomies (Fruhwirth et al., 2020; Spiekermann, 2019), however, overlook two main areas which this study aims to address. Firstly, the two studies mostly focus on the classification of multilateral data marketplaces, while in practice data trading often happens via bilaterally negotiated contracts (Koutroumpis et al., 2017). Secondly, the studies view data marketplace business models from a single firm perspective. However, data marketplaces take part in a network of stakeholders involving data analysts, application vendors, algorithm developers, data providers, consultants, licensing entities, and platform providers (Muschalle et al., 2012; Thomas & Leiponen, 2016).

To address the above two under-researched areas, this study develops a taxonomy from a multi-stakeholder perspective on business models. We define a business model as the way a network of stakeholders creates and captures value (Bouwman et al., 2008). This multi-stakeholder perspective allows us to understand the business model for the data ecosystem as a whole. Moreover, we define a data marketplace as a digital system where data is traded as an economic good, that connects data

buyers and data sellers, and facilitates data exchange and financial transactions (Koutroumpis et al., 2020; Stahl et al., 2016). In this way, the term data marketplace is broadly interpreted, to go beyond the already studied multilateral data marketplaces.

The remainder of this paper is structured as follows: in Section 2, the taxonomy development process is described. Subsequently, Section 3 presents the developed taxonomy on the basis of the identified business model dimensions. Lastly, Section 4 provides a conclusion of the research and discusses the scientific contribution, practical relevance and limitations of this study.

2 Taxonomy development process

To develop the taxonomy, we follow the taxonomy development method by Nickerson et al (2013). Meta-characteristics of the taxonomy are defined first. Next, the thirteen ending conditions suggested by Nickerson et al. (2013) were employed. After that, multiple iterations are conducted to refine the taxonomy.

2.1 Meta-characteristics

Meta-characteristics function as overarching characteristics of the object of interest (Nickerson et al., 2013). We use the four business model domains of the STOF ontology (i.e. Service, Technology, Organization and Finance domains as in Bouwman et al. (2008)) as the meta-characteristics of the taxonomy, as the STOF approach takes service as a unit of analysis and employs a multi-stakeholder perspective on business models (Bouwman et al., 2008). This perspective is well-suited for data marketplaces because a network of business actors are involved in and around data marketplaces (Muschalle et al., 2012; Thomas & Leiponen, 2016).

2.2 Literature search

We collected dimensions and characteristics from existing literature. A literature search was conducted to discover existing knowledge about the object of interest (Webster & Watson, 2002). Google Scholar was consulted to find relevant academic sources, using the search string "Data marketplaces" AND ("Business models" OR "Digital platform" OR "Digital marketplace" OR "Data trading" OR "Data

economy"). This string resulted in a total of 359 articles. The articles were scanned based on their title, abstract and relevance, which resulted in a preliminary selection of 17 articles. After making this pre-selection of articles, the full text of the articles was read, which resulted in the exclusion of seven articles that did not explicitly discuss dimensions or characteristics of data marketplace business models. Based on cross-reference of the selected articles, we added four additional articles that presented topic-relevant business model taxonomies to the list. The literature review resulted in a final set of 14 articles as presented in Table 1.

Table 1: Overview of classifications relevant to data marketplace business models

Author(s) (Year)	Type	Citations (dated 14 April 2020)
Schomm et al. (2013)	Dimensions of data providers and data marketplaces	73
Stahl et al. (2014a)		14
Stahl et al. (2014b)		16
Stahl et al. (2017)		12
Stahl et al. (2016)	Classification of electronic marketplaces	30
Koutroumpis et al. (2017)	Market designs for data marketplaces	19
Muschalle et al. (2012)	Pricing models for data marketplaces	74
Fricker and Maksimov (2017)		8
Spiekermann (2019)	Taxonomy of data marketplace business models	9
Fruhwirth et al. (2020)		1
Bock and Wiener (2017)	Taxonomy of digital business models	22
Täuscher (2016)	Taxonomy of marketplace business models	6
Täuscher and Laudien (2018)		153
Hartmann et al. (2014)	Taxonomy of data-driven business models	131

2.3 Selection of empirical cases

To account for the practical relevance of the taxonomy, we conducted desk research between May and July 2020 to build a database of empirical cases of data marketplaces. Sixty-five websites of data marketplaces that were mentioned in existing studies of data marketplaces were included in the database (Carnelley et al.,

2016; Koutroumpis et al., 2020, 2017; Prlja, 2019; Spiekermann, 2019; Stahl et al., 2016). The data discovery platform datarade.ai, a website that provides an overview of over 1,800 data providers and 200 data platforms, was consulted. In total, the search in the repository of datarade.ai resulted in the discovery of an additional set of 187 data marketplaces. To complement the database with cases that were not considered in the existing studies or part of the datarade.ai database, we used the search engine Google to further conduct a desk research. The keywords "data marketplace", "data market" and "data trading platform" were applied during the search. From this search, fifteen data marketplaces were added to the database.

Four criteria were applied to the companies that resulted from the desk research to ensure the relevance of the empirical cases. Firstly, data marketplaces that turned out to be shut down, after inspecting the website, were excluded from the database. Secondly, the companies that did not fit this study's definition of a data marketplace were excluded. This implied that data marketplaces that only provided open data, such as governmental organizations and NGOs, were excluded as these platforms adopt non-commercial business models (Carnelley et al., 2016). Thirdly, data marketplaces that did not have an English version of their website were excluded. Lastly, data marketplaces that were still in the construction phase were excluded. The application of these four criteria to the cases led to the exclusion of 89 cases. Therefore, the final database consisted of 178 cases of data marketplaces.

To analyse the business models of existing data marketplaces, a sample was taken from the database of cases. The empiricist philosophy of classification prescribes to build a taxonomy based on the consideration of many characteristics (Lambert, 2015). Therefore, the cases of data marketplaces in the database were first segmented into groups based on the similarity of their characteristics. The website of datarade.ai categorized data marketplaces based on the type of data traded on the platform. This variable was selected as the leading sampling variable to explore the variation between cases in the database. Based on the available information on datarade.ai and an inspection of the case's website, 138 cases could be labelled by type of data traded on the platform. The remaining 40 cases in the database were labelled based on the classification of the cases in the existing studies (Fruhwirth et al., 2020; Spiekermann, 2019) and through the manual inspection of the companies' website.

The segmentation of data marketplaces by type of data traded on the platform reveals that some segments of data marketplaces in the database were overrepresented. This was especially the case for audience data marketplaces, that constituted over 60% of the cases (N=112). Audience data is combined data about a certain target group of customers, which is much sought after by marketeers. To compensate for the overrepresentation, instead of random sampling, a disproportionate stratified sample of N=40 cases was taken from the database (Daniels, 2011). The final sample of 40 data marketplaces consisted of ten data marketplaces on which any type of data is traded (25% of the sample), four financial and alternative data marketplaces (10%), nine audience data marketplaces (22.5%), six sensor and mobility data marketplaces (15%), four geo data marketplaces (10%) and seven health and personal data marketplaces (17.5%) (available here: https://doi.org/10.4121/14679564.v1).

2.4   Design iterations

Our design phase started with a conceptual-to-empirical approach (Nickerson et al., 2013). In these design iterations, the concepts derived from the literature were compared to the sample of empirical cases. Information on the business models of the cases was collected from publicly available sources such as company websites and news articles. The discovered information fragments were coded using the dimensions and characteristics from the literature review as a guideline (See Table 2). After each case, newly identified characteristics were added to the dimensions of the taxonomy. After two conceptual-to-empirical design iterations, two empirical-to-conceptual iterations were conducted, which resulted in the addition of two dimensions to the taxonomy: *enterprise data marketplace* and *data processing and analytics tools*. After every design iteration, the ending conditions were checked. After two conceptual-to-empirical iterations and two empirical-to-conceptual iterations, both the objective and subjective ending conditions were met. Finally, to test the usefulness of the taxonomy, three mini-case studies were conducted on empirical cases of data marketplaces that were not part of the sample, i.e. Wibson, QueXopa and Advaneo respectively. The taxonomy was found to be useful, as the business models of the cases could be classified based on public information about the cases.

Table 2: Coding examples for the value proposition dimension

Characteristic	Case	Quote
Easy data access and/or tooling	Open:Factset Marketplace	*"FactSet creates data and technology solutions for investment professionals around the world, providing instant access to financial data and analytics that investors use to make crucial decisions."*
	Knoema	*"Knoema is a cloud-based data technology platform that makes data accessible and delivers intelligent data tools to enable data access and discovery."*
Secure data sharing	DAWEX	*"With Dawex Global Data Marketplace providers can highlight the value of their data while retaining full control over the distribution and configuration of usage rights."*
	Snowflake	*"Unlike other data marketplaces, Snowflake Data Marketplace leverages Snowflake's Secure Data Sharing technology, which means no data transfer and no need to squeeze data through APIs or use cloud storage."*
High quality and unique data	Amazon DSP	*"Use exclusive Amazon audiences to reach your ideal audience on and off Amazon."*
	Datax	*"Quality business data for better sales leads – Any campaign is only as good as the data it's built on – so make sure yours is the best.*

3 Taxonomy of Data Marketplace Business Models

The final taxonomy consists of 4 meta-dimensions, 17 dimensions and 59 characteristics and is presented in Table 3. In the following sections, the data marketplace business model dimensions are discussed per meta-dimension (STOF).

Table 3: Taxonomy of data marketplace business models

	Dimension	Characteristics			
Service domain	Value proposition	Easy data access and/or tooling	Secure data sharing	High quality and unique data	All services in a single platform
	Enterprise data marketplace	Yes		No	

	Data processing and analytics tools	Yes			No		
	Marketplace participants	B2B		C2B		Any	
	Industry domain	Any data	Geo data	Financial & Alternative data	Health & Personal data	Audience data	Sensor & Mobility data
	Geographic scope	Global		Regional		Local	
	Time frame	Static	Up-to-date	(Near) real-time		Multiple	
Technology domain	Platform architecture	Centralized			Decentralized		
	Data access	API	Download	Specialized software		Multiple options	
	Data source	Self-generated	Customer provided data	Acquired data		Multiple sources	
Organizatio n domain	Matching mechanism	One-to-one	One-to-many	Many-to-one		Many-to-Many	
	Platform sponsor	Private	Consortium		Independent		
Finance domain	Revenue model	Commissions	Subscriptions	Usage fees		Asset sales	
	Pricing model	Freemium	Pay-per-use	Flat fee tariff	Package based pricing	Multiple	
	Price discovery	Set by buyers	Negotiation	Set by marketplace provider		Set by external sellers	
	Smart contract	Yes			No		
	Payment currency	Fiat money			Cryptocurrency		

3.1 Service domain

The **value proposition** is a statement that indicates the proposed value that an enterprise intends to deliver to the customer (Bouwman et al., 2008). It often describes how customers can benefit from using the service and how the enterprise aims to set itself apart from the competition. Some data marketplaces offer an **enterprise data marketplace** as an additional service. An enterprise data marketplace functions as a private data marketplace that enables organizations to share data within the company or with external partners, such as suppliers and customers, that are invited by the focal organization. The **data processing and analytics tools** characteristic indicates whether a data marketplace offers additional tooling on top of the data, to perform analytics activities on proprietary data or data bought via the platform. The **marketplace participants** dimension describes the type of participants that are allowed to register and exchange data on the marketplace. While most data marketplaces allow the exchange of any type of data on their marketplace, some data marketplaces focus their data offering towards a specific **industry domain**. The **geographic scope** describes the regions in which the data marketplace is operating and available to users (Täuscher & Laudien, 2018; Täuscher, 2016). The **time frame** dimension describes whether or not the data needs frequent updates to maintain the relevancy of the data (Schomm et al., 2013).

3.2 Technology domain

Data marketplaces may adopt two types of **platform architectures**: centralized or decentralized (Koutroumpis et al., 2017). In the centralized approach, data providers offer their data products via a predefined centralized location on the platform, such as a cloud repository. In decentralized platforms, the data products remain at the data provider and the data is traded using distributed ledger technologies such as blockchain. Platform providers may provide **access to the data** in a number of different ways (Schomm et al, 2013). The **data source** dimension describes the origin where the data was gathered or collected by the data marketplace platform (Hartmann et al., 2014).

3.3 Organization domain

The **matching mechanism** determines the number of parties on each side of the platform (Koutroumpis et al., 2017). Besides multilateral data marketplaces, three more types of data marketplaces exist: bilateral data marketplaces (one-to-one matching), dispersal data marketplaces (one-to-many matching), and harvest data marketplaces (many-to-one matching). The **platform sponsor** can be a private individual or a group, a consortium of buyers or sellers, or an individual or a group that is independent of other market players (Stahl et al., 2017, 2016).

3.4 Finance domain

The **revenue** dimension describes the main source of revenue for a data marketplace (Spiekermann, 2019; Täuscher & Laudien, 2018; Täuscher, 2016). The **pricing model** specifies how the final price for the data good or service is composed (Fruhwirth et al., 2020; Schomm et al., 2013; Spiekermann, 2019; Täuscher & Laudien, 2018; Täuscher, 2016). A **price discovery** function allows buyers and sellers on the marketplace to determine a transaction price which they both agree on (Bakos, 1998). Data marketplaces may implement **smart contracts** to enhance transparency and to enforce trust among marketplace participants (Fruhwirth et al., 2020). The **payment currency** dimension explicates which currencies are accepted for the payments made by marketplace participants (Fruhwirth et al., 2020).

4 Discussion and Conclusion

The developed taxonomy of data marketplace business model has two key scientific contributions. First, the results of the study contribute to the scarce knowledge about data marketplaces and their respective business models (Thomas & Leiponen, 2016). This study adopts a multi-stakeholder perspective on data marketplace business models by emphasizing the roles in the data ecosystem. The taxonomy provides an overview of contemporary knowledge about data marketplace business models and exposes new business model alterations that have emerged in practice. A second contribution made by this study is related to the interpretation of a data marketplace. Existing taxonomies (Fruhwirth et al., 2020; Spiekermann, 2019) focus on studying one type of data marketplaces: multilateral data marketplaces (Koutroumpis et al., 2017). In our study, data marketplaces are more broadly

interpreted as digital systems for trading data as an economic good, that connect buyers and sellers, and facilitate data exchange and financial transactions. This allows us to identify additional business model dimensions, which are not part of existing taxonomies: enterprise data marketplace, data processing and analytics tools, geographic scope, matching mechanism and platform sponsor. By eliciting how data marketplace business models differ, we provide a basis for fine-grained theory development, which is often lacking in platform studies (De Reuver et al., 2018).

The developed taxonomy can guide decision-makers who are exploring the options of setting up a data marketplace or to join an existing data marketplace. An improved understanding about data marketplace business models may help to achieve commercialization, that will make data more accessible and exploitable to individuals, businesses and authorities.

Although we took a systematic approach, subjectivity in assessing the cases may pose a limitation. We dealt with this by conducting multiple iterations and re-interpretations of the data. Further, not all data marketplace companies disclose sufficient information about all of their business model characteristics. Therefore, not all empirical cases could be classified into all of the conceptually derived dimensions. This was especially the case for financially related dimensions such as revenue partners and cost categories (Täuscher & Laudien, 2018). Lastly, as in any taxonomy development study, our study is limited to the current set of phenomena that exist in practice. Hence, future research may update our taxonomy in light of fundamentally new data marketplace types.

Data marketplaces pose a foundation for the data economy: they enable firms to access external data to drive their business and to profit from selling their own data. The EU is investing heavily in data marketplaces in the years to come (European Commission, 2020). At the same time, ambiguity pertains over what constitutes a viable data marketplace business model. Our taxonomy takes a broad and multi-stakeholder perspective to data marketplaces, going beyond the single-firm multilateral perspective of extant taxonomies. We argue that such a broad conceptual basis is needed to advance scholarly understanding of ecosystems in the data economy and to unlock the potential of trading data for a functioning data economy.

Acknowledgments

References

Bakos, Y. (1998). The emerging role of electronic marketplaces on the Internet. Communications of the ACM, 41(8), 35-42.

Bock, M., & Wiener, M. (2017). Towards a Taxonomy of Digital Business Models-Conceptual Dimensions and Empirical Illustrations. In ICIS.

Bouwman, H., de Vos, H., & Haaker, T. (Eds.). (2008). Mobile service innovation and business models. Springer Science & Business Media.

Carnelley, P., Schwenk, H., Cattaneo, G., Micheletti, G., & Osimo, D. (2013). Europe's data marketplaces—current status and future perspectives,'. European Data Market SMART, 63.

Daniel, J. (2011). Sampling essentials: Practical guidelines for making sampling choices. Sage Publications.

de Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: a research agenda. Journal of Information Technology, 33(2), 124-135.

European Commission. (2020). A European Strategy for Data [Press release]. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN

Fricker, S. A., & Maksimov, Y. V. (2017). Pricing of data products in data marketplaces. In International Conference of Software Business (pp. 49-66). Springer, Cham.

Fruhwirth, M., Rachinger, M., & Prlja, E. (2020). Discovering Business Models of Data Marketplaces. In Proceedings of the 53rd Hawaii International Conference on System Sciences.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). Big data for big business? A taxonomy of data-driven business models used by start-up firms. Cambridge Service Alliance, 1-29.

Koutroumpis, P., Leiponen, A., & Thomas, L. D. (2017). The (unfulfilled) potential of data marketplaces (No. 53). ETLA Working Papers.

Koutroumpis, P., Leiponen, A., & Thomas, L. D. (2020). Markets for data. Industrial and Corporate Change, 29(3), 645-660.

Lambert, S. (2015). The importance of classification to business model research. Journal of Business Models, 3(1).

Muschalle, A., Stahl, F., Löser, A., & Vossen, G. (2012). Pricing approaches for data markets. In International workshop on business intelligence for the real-time enterprise (pp. 129-144). Springer, Berlin, Heidelberg.

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.

Prlja, E. (2019). Discovering Business Models of Data Marketplaces [Master's thesis]. Graz University of Technology, Graz, Austria.

Schomm, F., Stahl, F., & Vossen, G. (2013). Marketplaces for data: an initial survey. ACM SIGMOD Record, 42(1), 15-26.

Spiekermann, M., Tebernum, D., Wenzel, S., & Otto, B. (2018). A metadata model for data goods. In Multikonferenz Wirtschaftsinformatik (Vol. 2018, pp. 326-337).

Spiekermann, M. (2019). Data marketplaces: Trends and monetisation of data goods. Intereconomics, 54(4), 208-216.

Stahl, F., Schomm, F., & Vossen, G. (2014a). Data Marketplaces: An Emerging Species. In DB&IS (pp. 145-158).

Stahl, F., Schomm, F., & Vossen, G. (2014b). The data marketplace survey revisited (No. 18). ERCIS Working Paper.

Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. Vietnam Journal of Computer Science, 3(3), 137-143.

Stahl, F., Schomm, F., Vomfell, L., & Vossen, G. (2017). Marketplaces for Digital Data: Quo Vadis?. Computer and Information Science, 10(4).

Täuscher, K. (2016). Business models in the digital economy: an empirical study of digital marketplaces. Fraunhofer MOEZ. Fraunhofer Center for International Management and Knowledge Economy, Städtisches Kaufhaus Leipzig, Neumarkt, 9-19.

Täuscher, K., & Laudien, S. M. (2018). Understanding platform business models: A mixed methods study of marketplaces. European Management Journal, 36(3), 319-329.

Thomas, L. D., & Leiponen, A. (2016). Big data commercialization. IEEE Engineering Management Review, 44(2), 74-90.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly, xiii-xxiii

Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers

Tiina Kemppainen,[1,2] Lauri Frank,[2]
Markus Makkonen[2] & Oona-Iina Hyvönen[2]

[1] University of Jyvaskyla, School of Business and Economic, Finland; e-mail:
tikemppa@jyu.fi
[2] University of Jyvaskyla, Faculty of Information Technology, Finland; e-mail:
tiina.kemppainen@hotmail.com, lauri.frank@jyu.fi, markus.v.makkonen@jyu.fi,
hyvonen.oona@gmail.com

Abstract This qualitative study investigates the barriers to responsible consumption in e-commerce from the online shoppers' viewpoint. The purpose of the study is to increase our understanding of what prevents young adults from making responsible purchases in online stores in the context of fashion retail. The data were collected by interviewing ten Finnish fashion shoppers aged 23-27 years. The findings show that responsible consumption is perceived as complex and challenging. The study identified barriers related to online stores and consumers themselves. Online store implementation (product availability, information and transparency, and pricing) is vital in facilitating online shoppers' responsible purchasing decisions. However, consumers' personal consumption patterns and habits, and problems related to time use and responsibility assessment, can also be constraints on responsible consumption. Future studies are encouraged to investigate how online solutions, such as user interfaces, online tools, and apps, could better assist consumers in overcoming the identified barriers.

1 Introduction

The popularity of online shopping has grown tremendously with the spread of information and communications technologies (ICTs). The popularity of e-commerce is explained by its ease, availability, and breadth of selection. However, while e-commerce provides multiple benefits for today's consumers, from an ecological perspective it has also contributed to negative changes in consumption habits and purchasing behavior (Abukhader & Jonson 2004). The efficiency and affordability of online shopping have led to more shopping being done without real need. Many online stores offer consumers free shipping and returns, as well as an extended payment period, which has resulted in products being easily ordered and inappropriate products being returned. Other environmental challenges posed by e-commerce relate to individual deliveries and small one-off purchases as well as packaging materials. Simultaneously, the state of the planet and overconsumption have become significant issues, and consumers are becoming more environmentally conscious (Demarque et al. 2015). The negative repercussions of e-commerce have become important topics of discussion in business and academia.

This qualitative study investigates the barriers to responsible consumption in e-commerce from the online shoppers' viewpoint. The study aims to increase our understanding of what prevents young adults (aged 23-27) from making responsible purchases in online stores in the context of fashion retail. Previous research has suggested that there is no significant difference in terms of consumers' purchasing decision process between online and brick-and-mortar stores (Katawetawaraks & Wang 2011). In addition, most studies approach responsible consumption at a general level without shopping channel (online/offline) specification. Despite a bulk of papers investigating responsible consumption from different perspectives, a specific understanding of online environments is missing. Therefore, as the importance of online shopping grows continuously, there is a need for studies that focus on responsible consumption in online channels. Furthermore, previous research has shown that consumers' attitudes towards responsibility and their behavior can be contradictory: responsibility is valued, but consumers do not necessarily behave responsibly (Belk, Devinney & Eckhardt 2005; Carrington, Neville & Whitwell 2014). Investigations on the barriers to responsible consumption help to better understand why this is the case. The findings of this study are also

T. Kemppainen, L. Frank, M. Makkonen & O.-I. Hyvönen:
Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers

325

useful in designing online store user interfaces that support consumers' responsible purchasing behavior.

This paper discusses the previous studies on responsible consumption in the e-commerce context in Section 2 below. Section 3 discusses the data collection and analysis, and Section 4 presents the findings of the empirical study. Section 5 discusses the contributions and managerial implications of this study and gives some suggestions for future research.

2 Theoretical background: responsible consumption and e-commerce

Responsible and sustainable consumption have become common research topics in the 21st century. The concepts are intrinsically linked, as sustainability can be seen as an objective pursued through responsible action. The study fields have become deeply blurred and similar issues are discussed within both fields (Bansal & Song 2017). The core idea of responsible consumption is to reduce the impact of goods or services on the environment by various eco-friendly activities. In line with Ulusoy (2016), this study defines responsible consumption as consumption that has a less negative impact on the environment than consumption that does not take into account the foundations of sustainable development.

Extant studies on responsible consumption do not usually distinguish between different service channels; hence responsible consumption in e-commerce channels has not received special attention. However, factors affecting responsible consumption and other responsible behavior have been widely explored. These studies have highlighted situational factors. According to Stern et al. (1995), favorable situational factors increase and encourage green consumption. One can assume that the situational factors affecting responsible purchase behavior, in particular, can be very different in online and offline environments. In addition, the impact of demographic and psychographic factors on consumer responsibility have been addressed. Straughan and Roberts (1999) highlight the psychographic factors: environmental awareness, opinions, and attitudes influence responsible purchasing intentions. According to previous research, factors that encourage responsible consumption include, for instance; knowledge (Young et al. 2010; Joshi & Rahman 2016), a positive attitude towards responsibility (Straughan & Roberts 1999; Demarque et al. 2015; Kumar, Manrai & Manrai 2017), marketing and campaigns,

and brand image (Joshi & Rahman 2016). Consumer sentiment, especially environmental concerns, has a positive and direct impact on responsible goals and behavior (Straughan & Roberts 1999). Identified barriers to responsible consumption include, for instance, high prices (Peattie 2010; Joshi & Rahman 2016), lack of information (Demarque et al. 2015), and awareness (Luthra, Govindan & Mangla 2017), poor product availability and quality (Joshi & Rahman 2016), and lack of policy support (Blok et al. 2015).

The fashion industry is often emphasized in the debate on responsibility because its enormous environmental challenges are increasingly being recognized. The fashion industry has been associated with significant negative social and ecological impacts, such as employee exploitation, excessive use of renewable materials, and waste generation (Wang et al. 2019). In addition, fashion retail is also an important segment of e-commerce: its share of online shopping is significant. For example, in 2017, 50% of Finnish consumers bought clothes and accessories online, and the number is expected to grow every year (Paytrail 2018). Online shopping for clothing and shoes in particular is a burden on the environment due to high return rates (Chen & Chai 2010; Demarque et al. 2015). For these reasons, the empirical part of this study focuses on fashion retail: the online shopping for clothing, accessories, and footwear.

3   Data collection and analysis

As the purpose of this study was to identify and better understand the barriers to responsible consumption from the consumer viewpoint, a qualitative research approach was chosen. Young adults aged 23 to 27 were selected as the target group for the study because they are characterized by traits associated with the Generation Z. Those born around the mid-1990s were born into the digital world with internet and electronics. They are typically experienced online shoppers and are often aware of responsibility-related themes. Dabija et al. (2019) note that applying green strategies and sustainable principles in business activities is vital when trying to appeal to the Generation Z consumers.

T. Kemppainen, L. Frank, M. Makkonen & O.-I. Hyvönen:
Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers

327

The data collection was carried out with semi-structured interviews in 2020. The number of interviewees was determined on the basis of the saturation of the material obtained from the interviews: interviews were conducted until they no longer produced new information relevant to the study. Ten people, four males and six females (referred to as participants 1-10: P1-P10) were interviewed (for details, see Appendix 1). The interviews were conducted via the Google Meet video call service with an average duration of 39 minutes. The participants were asked to describe their e-commerce shopping and responsible consumption behaviors regarding fashion wear. The discussed themes included, for instance: how participants define responsibility in consumption, how responsibility guides them as consumers and influences their purchase behavior in online stores, and what are the problems or benefits of online shopping in terms of responsible consumption. The exact form or order of the questions was not defined in advance, which provided an opportunity for additional questions and discussion. All interviews were recorded and transcribed.

The data was processed by inductive coding (Thomas 2003). The findings emerged from the analysis of the transcribed data. All the mentioned barriers to responsible consumption were first collected from the material, after which they were classified into groups according to their content. Finally, as the result of several classification rounds, the identified barriers were formed into two main groups, including online store characteristics, and personal characteristics and resources. The findings are discussed more closely in the next section.

4 Findings

The findings show that responsible consumption is often perceived as complex and challenging. It is hampered both by the information and products offered by online stores and by consumers' own personal characteristics and resources. The data analysis revealed three barriers related to the online stores (poor availability, lack of information and transparency, and high prices) and three personal barriers (poor knowledge and challenges in responsibility assessment, existing consumption habits, and lack of time) to responsible consumption in e-commerce. These often intertwined barriers are discussed here below.

4.1   Online store characteristics

Poor availability was one of the main factors influencing the participants' responsible purchasing decisions. According to the interviewees, the supply of responsible products is still weak, and some manufacturers do not offer responsible products at all. It was concluded that when a need for a particular product type arises, there are only a few responsibly produced options available. If one desires a specific brand, there might not be any responsible alternatives.

> "Supply. Because if you want something specific, there may not be responsible alternatives." – Male, 23 (P4)

However, online stores were seen as a better option in terms of supply than brick-and-mortar stores. The online product range is more extensive. Particularly specialty stores (that only sell products that are produced responsibly) make it easier to browse products from different manufacturers. Many ecofriendly brands have also focused their sales online, making products unavailable offline. With the growth and spread of e-commerce, the selection and availability of responsible brands have increased significantly.

Lack of information and transparency. The interviewees were aware of the environmental burden caused by consumption. Better awareness had led them to consider more of their purchases, and many admitted that they had begun to pay more attention to the backgrounds of products with increased awareness. However, obtaining information was not considered easy. The lack of information and transparency of company operations was repeated in the interviewees' reports.

> "I hope it would not be left to the online store user—that we have to play a detective and look for information. Backgrounds should be very openly presented in the online store." – Female, 25 (P9)

Although production transparency has been increased, and companies' sustainability efforts and goals have been communicated more openly, it was unclear to the participants how the companies' operations (e.g., production) have actually changed: whether companies are behaving more responsibly than in the past. In particular,

T. Kemppainen, L. Frank, M. Makkonen & O.-I. Hyvönen:
Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers

329

the different eco-certificates and the increased openness of fast fashion chains made the participants suspicious.

> "I don't think it has changed very much from the past, although transparency and awareness are now much increased and promoted. Those are just trendy words." – Female, 25 (P9)

Instead of certificates and superficial information, interviewees looked for concrete examples and explanations.

> "I think those certifications might be a bit out there at the moment. If the production methods and chains are clearly explained, it is much better than a single certificate." – Male, 27 (P6)

> "I'd be interested to know more about it than just that 'this charge will offset the carbon footprint'. It would be nice to know exactly what you pay extra for and whether it truly makes a big difference." – Female, 25 (P2)

However, information retrieval and comparison were found to be easier and more trustworthy in online than offline environments. In the online environment, information about product backgrounds and producers' practices can also be obtained from different external sources.

High prices. The importance of price was apparent for the participants' willingness to make a responsible purchasing decision—many considered high price to be a considerable barrier to responsible or more responsible consumption.

> "Many times, if there is a product in which all the material and workmanship is Finnish, then the price is often so much higher that it may not be possible to buy such a product." – Female, 26 (P5)

The interviewees emphasized the importance of brand image and value-for-money when considering the price. The products of responsible brands were perceived to be more expensive yet of better quality than fast-fashion chains' responsible options. Purchasing responsible fast fashion at a cheaper price did not seem sensible in the long run.

4.2   Personal characteristics and resources

Poor knowledge and challenges in responsibility assessment were identified as important obstacles to responsible consumption. When consumers are unaware of the effect of their purchases, they also do not feel guilty about their irresponsible purchasing decisions.

> "I think it's just that people don't necessarily know or think about these issues that much." – Female, 26 (P5)

The assessment of the environmental impact was considered a challenging task. Whereas looking for information about products and companies and their background can be difficult, the fact that a consumer must also evaluate the information and assess its reliability adds to the burden.

Existing consumption habits. The individual's current consumption habits were also identified as a barrier to responsible consumption. The phenomenon of responsible shopping was considered rather new and many of the interviewees had only recently woken up to consider their consumption habits.

> Responsibility is a whole new thing. Probably many are used to shopping, and not thinking about the effects at all. I haven't always thought about it as I do now." – Female, 26 (P10)

Moreover, responsibility and the disadvantages caused by consumption are not always the main things in mind when shopping. The interviewees admitted that their hedonism restricts responsible purchasing decisions. While responsibility is considered important, it can be easily forgotten in a purchase situation.

> "When I don't need or want anything, I start to think that 'yes, we should have more solidarity with each other and reduce consumption'. But as soon as you need something, such ideas disappear. Your greed and desire for pleasure are obstacles." – Female, 25 (P9)

T. Kemppainen, L. Frank, M. Makkonen & O.-I. Hyvönen:
Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers

331

When compared to brick-and-mortar stores, online stores were seen as a more irresponsible shopping channel. This was explained by people's consumption habits in online shopping environments.

> "I find it [e-commerce] less environmentally friendly. Just because of the way people behave in online shopping: they order a lot and then return." – Female, 23 (P1)

The interviewees concluded that it is easier to make impulsive and unplanned purchases in an e-commerce environment because online stores create various incentives for the consumer (e.g., free shipping) to increase the total amount of purchases. Many of the interviewees concluded that they tend to buy less at once in brick-and-mortar stores. Moreover, the responsibility of consumption in general became an essential theme in the interviews, as some of the interviewees considered consumption irresponsible regardless of the channel.

> "It's complicated. If you want to be truly responsible, then you don't buy anything; that's the most responsible activity of all." – Male, 27 (P6)

However, completely stopping or significantly reducing consumption was perceived as challenging, as it would require making significant changes in one's everyday life.

Lack of time. Finally, urgency and lack of time were considered noteworthy barriers to responsible consumption, which was seen as time-consuming. While it was deemed important, the participants explained that they might not behave in a responsible manner because they do not have the time to look for information about products, or to look for a responsible alternative for a certain product. Aspirations to responsibility are not always reflected in actions, as the following quote shows.

> "I do think about these [ecological] issues. But actions may be different, you don't always behave as you would like to." – Female, 23 (P1)

It was pointed out that it may take a lot of time to clarify the background of the products, even if there is some information provided in the online store. In the context of online shopping, the customer also needs to consider the total time taken by the process, including the time required for the order to arrive.

"Getting acquainted takes a lot of time, and then time just runs out. You go: 'I can't do this anymore'." – Male, 27 (P6)

Online shopping was generally perceived as easy and effortless. Responsible consumption in online channels, on the other hand, was perceived to be complex, as it often requires awareness and the willingness to search for information.

5 Discussion: conclusions and suggestions for future research

This study investigated the barriers to responsible consumption in e-commerce from young adults' viewpoint in the context of fashion retail. The findings show that while the online store implementation (including product availability, pricing, and information related to responsibility) is essential in facilitating responsible purchasing decisions, consumers' personal consumption patterns and habits, as well as everyday challenges, also play a vital role in responsible consumption. While responsibility is considered necessary, responsible purchase decisions are often considered to require time and effort. Incorporating responsibility into existing consumption habits can be challenging, and it is easy to ignore and forget, especially in urgent purchase decisions.

Overall, the findings of this study are in line with previous studies not focused on e-commerce. They demonstrate that responsible consumption is of interest to many, but there is a gap between ideals and responsible actions (Belk, Devinney & Eckhardt 2005; Carrington, Neville & Whitwell 2014). Responsibility is easily avoided due to its complications related to the pricing and product availability (Chen & Chai, 2010). The findings suggest that pricing of responsible products can be a challenging task from an online store's perspective. Even though high prices are considered purchase barriers, overly low prices, in contrast, can be perceived as suspicious, signaling that everything is not truly done responsibly in the production chain. Many of the identified barriers are linked to information, knowledge, and understanding related to responsibility (e.g., Young et al. 2010; Demarque et al. 2015; Joshi & Rahman 2016). Although responsibility-related information has increased in online stores, background information regarding products and company operations is still considered limited. Furthermore, shops and manufacturers communicate their responsibility measures in very different ways, making the responsibility assessment

T. Kemppainen, L. Frank, M. Makkonen & O.-I. Hyvönen:
Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers

333

difficult. Comparing the different options, including companies, their eco-labels and eco-actions, and products, is a challenge for consumers.

The findings suggest that consumers need common standards and more concrete and easy-to-understand measures and indicators to support their responsible purchasing decisions in online environments. In general, solutions that make responsible shopping more straightforward and less time-consuming are needed. There is a need for studies and solutions that consider how product backgrounds, material choices, and manufacturing information, for instance, can be better informed and communicated to consumers. Future studies should investigate how user interfaces, online tools, and apps could help consumers in their responsibility assessments. Studies on other contexts and different consumer groups are also welcomed in order to better understand whether similar barriers exist within different retail categories or consumer groups, for example younger and older consumers.

Finally, although this study identified barriers to responsible consumption with e-commerce, similar problems exist with brick-and-mortar stores. Hence, when considering the obstacles, online stores can have many competitive advantages over offline stores. Even though the availability of products and responsibility-related information were characterized as deficient in online stores, online stores typically offer more options for responsible consumption. In addition, information retrieval and product comparisons are often easier online. Hence, as the importance of responsible consumption and online shopping will grow continuously in the future, this should create opportunities for new online businesses that are either fully specialized in selling responsible products or acknowledge the foundations of sustainable development in their services. It also provides excellent opportunities for online solutions and systems that educate and assist both consumers and businesses in their pursuit of responsible consumption in all kinds of services.

References

Abukhader, S. M. & Jonson, G. (2004). E-commerce and the environment: A gateway to the renewal of greening supply chains. International Journal of Technology Management, 28 (2), 274-288.

Bansal, P. & Song, H. (2017). Similar but not the same: Differentiating corporate sustainability from corporate responsibility. Academy of Management Annals, 11 (1), 105-149.

Belk, R., Devinney, T. & Eckhardt, G. (2005). Consumer ethics across cultures. Consumption Markets & Culture, 8 (3), 275-289.

Blok, V., Long, T. B., Gaziulusoy, A. I., Ciliz, N., Lozano, R., et al. (2015). From best practices to bridges for a more sustainable future: Advances and challenges in the transition to global sustainable production and consumption: Introduction to the ERSCP stream of the Special volume. Journal of Cleaner Production, 108, 19-30.

Carrington, M. J., Neville, B. A. & Whitwell, G. J. (2014). Lost in translation: Exploring the ethical consumer intention–behavior gap. Journal of Business Research, 67 (1), 2759-2767.

Chen, T. B. & Chai, L. T. (2010). Attitude towards the environment and green products: Consumers' perspective. Management Science and Engineering, 4 (2), 27-39.

Dabija, D., Bejan, B. M. & Dinu, V. (2019). How Sustainability oriented is generation Z in retail? A literature review. Transformations in Business & Economics, 18 (2), 140.

Demarque, C., Charalambides, L., Hilton, D. J. & Waroquier, L. (2015). Nudging sustainable consumption: The use of descriptive norms to promote a minority behavior in a realistic online shopping environment. Journal of Environmental Psychology, 43, 166-174.

Joshi, Y. & Rahman, Z. (2016). Predictors of young consumer's green purchase behaviour. Management of Environmental Quality: An International Journal, 27 (4), 452-472.

Katawetawaraks, C. & Wang, C. (2011). Online shopper behavior: Influences of online shopping decision. Asian Journal of Business Research, 1 (2), 66-74.

Kumar, B., Manrai, A. K. & Manrai, L. A. (2017). Purchasing behaviour for environmentally sustainable products: A conceptual framework and empirical study. Journal of Retailing and Consumer Services, 34, 1-9.

Luthra, S., Govindan, K. & Mangla, S. K. (2017). Structural model for sustainable consumption and production adoption—A grey-DEMATEL based approach. Resources, Conservation and Recycling, 125, 198-207.

Paytrail 2018. Finnish E-commerce Report 2017. Available in:
https://www.paytrail.com/en/report/finnish-ecommerce-2017?hsCtaTracking=ffa39c05-b328-4556-b789-303ad39fce14%7C678c1de2-64db-4c7c-a221-0ea4530e6cc6.

Peattie, K. (2010). Green consumption: behavior and norms. Annual Review of Environment and Resources, 35, 195-228.

Stern, P. C., Dietz, T. & Guagnano, G. A. (1995). The new ecological paradigm in social-psychological context. Environment and Behavior, 27 (6), 723-743.

Straughan, R. D. & Roberts, J. A. (1999). Environmental segmentation alternatives: a look at green consumer behavior in the new millennium. Journal of Consumer Marketing, 16 (6), 558-575.

Thomas, D. R. (2003). A general inductive approach for qualitative data analysis. School of Population Health, University of Auckland.

Ulusoy, E. (2016). Experiential responsible consumption. Journal of Business Research, 69 (1), 284-297.

Wang, H., Liu, H., Kim, S. J. & Kim, K. H. (2019). Sustainable fashion index model and its implication. Journal of Business Research, 99, 430-437.

Young, W., Hwang, K., McDonald, S. & Oates, C. J. (2010). Sustainable consumption: green consumer behaviour when purchasing products. Sustainable Development, 18 (1), 20-31

T. Kemppainen, L. Frank, M. Makkonen & O.-I. Hyvönen:
Barriers to Responsible Consumption in e-Commerce: Evidence from Fashion Shoppers

335

Appendix 1: List of participants

	Age	Sex	Position	Interview length (min)
P1	23	Female	Student & part-time employee	42
P2	25	Female	Student	37
P3	25	Male	Student & full-time employee	45
P4	23	Male	Student	34
P5	26	Female	Student & full-time employee	37
P6	27	Male	Student & full-time employee	42
P7	25	Male	Student & part-time employee	38
P8	25	Female	Student & part-time employee	35
P9	25	Female	Full-time employee	40
P10	26	Female	Student & full-time employee	37

User Information Satisfaction Among Female Refugees and Immigrants as Assessed by the Level of Information Literacy on Social Media

Khadijah Kainat, Maedeh Ghorbanian Zolbin, Gunilla Widén & Shahrokh Nikou

Åbo Akademi University, Faculty of Social Sciences, Business and Economics, Turku, Finland; e-mail: khadijah.kainat@abo.fi, maedeh.ghorbanianzolbin@abo.fi, gunilla.widen@abo.fi, shahrokh.nikou@abo.fi

Abstract Female refugees and immigrants face various challenges in accessing, using, and sharing information during their integration process. In the context of COVID-19, this study aims to identify the user information satisfaction of female refugees and immigrants living in Finland and Sweden. Using a dataset comprising 232 respondents, the research model was examined through structural equation modelling. The results show that information overload in social media streams has an impact on information fatigue and consequently on the information avoidance behaviour of the target group. The results also show that information literacy helps to avoid social media information overload, in addition to its direct effect on user information satisfaction. Being familiar with the perceived COVID-19 challenges also positively impacts user information satisfaction. In view of the fact that European countries are receiving an ever-increasing number of refugees and immigrants, the findings of this study provide both theoretical and practical contributions.

1   Introduction

When entering a new country, refugees and immigrants face various integration challenges such as socio-cultural barriers, lack of language skills, health and psychological problems, employment problems, and family issues (Bronstein, 2019). They also face multiple information challenges in accessing, sharing, and using the information, in addition to challenges such as information overload (hereinafter IO), lack of information, misinformation, culturally nuanced information due to the language barrier, lack of literacy skills, lack of social networks, socio-cultural differences, and personal preferences, which are essential for managing everyday information needs in a host country (Lloyd et al., 2013). As such, it can be argued that information is vital for successful integration. Some authors, such as Swar et al. (2017), have stated that social media IO occurs when seekers have more information than can be assimilated, or in other words, the ability of information seekers to process and handle the information is insufficient to successfully cope with the amount of information they receive. The increasing use of social media platforms (e.g. Facebook, Instagram, Twitter) has fed people with excessive information content in the form of messages, tweets, wall posts, and constant status updates on a daily basis (Bright et al., 2015). This is resulting in people quitting or reducing the use of social media to avoid excessive information, which is known as social media information fatigue (hereinafter IF) (Cao & Sun, 2018). Moreover, social media fatigue, caused by IO, can be linked to the inherent problems associated with information avoidance (hereinafter IA) by social media users. It can be difficult for people to check each message on social media; they may skim or skip irrelevant information or even avoid some information (Guo et al., 2020). Mostly, people are seeking information that is in line with their interests, needs, or existing attitudes, and they avoid the information if these are not in accordance with their point of view – whether this is done consciously or unconsciously (Narayan et al., 2011). Moreover, a less-studied concept in the context of refugees and immigrants is information literacy (hereinafter IL), which has been proven to influence user satisfaction. Some authors, such as Aljanabi and Hadban (2018), have asserted that mastering certain skills and acquiring a robust IL level give individuals a sense of pleasure and confidence which, in turn, can foster their satisfaction. In the context of the COVID-19 pandemic, this paper aims to focus on information seeking in female refugees and immigrants to (i) investigate how such individuals react to information challenges, and (ii) examine their level of information satisfaction. Khan and Eskola (2020) highlighted female refugees and immigrants in particular face unique information challenges during the integration process and need special

attention to fully understand their information practices and the inherent problems, along with the strategies they use to overcome these challenges. In earlier literature (e.g. Liebig & Tronstad, 2018), female refugees and immigrants are often overlooked or studied in combination with their male counterparts. Thus, this study takes the liberty of explicitly focusing on female refugees and immigrants, in order to analyse their user information satisfaction. More importantly, in light of the pandemic brought about by the spread of COVID-19, this study aims to investigate how the current situation impacts the user information satisfaction of this target group. The COVID-19–related challenges in this paper refer to, for example, searching for information to manage stress from day to day when quarantined at home. The question guiding this study is, "*What factors influence the user information satisfaction of female refugees and immigrants, and what roles do information literacy and the perception of COVID-19–related challenges play in this context?*"

2 Literature Review and Hypotheses Developments

The research extends the (C–A–C) framework (Hilgard, 1980), which includes three components of the mind or consciousness: cognition–affect–conation. According to Huitt (1999), cognition refers to the process of knowing and understanding, and affect refers to the emotional interpretation of perception, information, or knowledge. The effect can be used to address questions such as, "How do I feel about this knowledge or information?". With conation, questions such as "Why is this information important?" can be asked, and it refers to the intentional and personal motivation of behaviour. In the following sections, we discuss this in more detail.

2.1 Social Media Information Overload

In the literature, with the specific focus of understanding the experiences of refugees and immigrants, there is little research related to the information challenges faced by this group. It is important to understand the challenges refugees and immigrants face in accessing, using, and sharing the information to provide them with better integration services and facilities when faced with living in a new country (Quike, 2011). Among the many challenges – such as socio-cultural and linguistic barriers to information, lack of information, and misinformation – information overload is one of the main challenges faced by this group and has the potential to affect their settlement experiences in a new country. Some researchers use IO to indicate the

possibility of being bombarded with a large amount of unwanted information, some of which may be relevant (Zhang et al., 2016). The IO situation can lead to undesirable consequences such as stress and anxiety (Guo et al. 2020), may weaken the decision quality (Eppler & Mengid, 2004), and may cause the individual concerned to withdraw from the tasks at hand (Cao & Yu, 2019). Social media is one of the main sources of IO. People seek, use, and share different types of information on social media every day. Information overload is defined as the tendency of people to withdraw from social media use when they become overwhelmed with too much content, too many sites, too many friends and contacts, and too much time spent keeping up with these connections (Yu et al., 2018). People cut down on their use of social media for a variety of reasons, such as lack of interest, energy, time, and knowledge, privacy concerns, anxiety, personal reasons, and fear of academic failure (Turan, 2013), but IO remains one of the main reasons for quitting or withdrawing from social media (Cao & Yu, 2019). Moreover, Dai et al. (2020) showed that IO leads to social media fatigue; hence:

H1: *Social media IO has a significant effect on the social media IF experienced by female refugees and immigrants*

2.2   Social Media Information Fatigue

Researchers have found that social media fatigue is an important driver of individuals' decisions to quit social networking platforms (Ravindran et al., 2014). When people tire of spending too much time and energy on social media, they may practise IA behaviours to escape from the negative emotions and fatigue. Dia et al. (2020) found positive relationship between social media fatigue and IA; hence:

H2: *Social media IF has a significant effect on the social media IO of female refugees and immigrants*

2.3   Social Media Information Avoidance

Information avoidance refers to human behaviour in which users ignore some information intentionally in order to save their time and energy or reduce stress (Dia et al., 2020). Information overload is linked to IA in different aspects such as when people actively ignore information because of external pressures of fears and uncertainty. Savolainen (2007) defined IA as when people avoid unnecessary or

negative information by filtering it and withdrawing from it in their everyday life. Similarly, Fisher et al. (2005) asserted that certain coping behaviours – i.e. monitoring (actively seeking solutions to one's problems) and blunting (avoidance of threatening information) – are essential aspects of information behaviour. It makes sense that when people feel less stressed, they feel more satisfied with the information. As such, it can be argued that an individual's ability to avoid irrelevant and unwanted information is directly linked to their user information satisfaction; hence:

H3: *Social media IA has a significant effect on the user information satisfaction of female refugees and immigrants*

2.4   Information Literacy

Information literacy is a required skill for evaluating the information retrieved and shared via social media (Pinto et al., 2020). It is a useful way of ensuring the effective use of information accessible through social media, considering the excessive amount of information available and the potential risks that it carries on the internet. Information literacy is linked to IA. For example, McCloud et al. (2013) identified that if an individual is having difficulties in finding information or comprehending information, he/she is more likely to avoid information. In another study (Karim et al., 2019), it was revealed that higher education and higher IL self-efficacy can reduce the propensity for IA. However, IL is crucial for understanding the user information satisfaction among refugees and immigrants in a social media context. IL skills relating to the use of digital and social media resources help refugees and immigrants to overcome the inherent obstacles in order to satisfy their everyday information needs and to better integrate into the host society (Martzoukou & Burnett, 2018). The need for social media IL during the current pandemic is even more crucial – so as to avoid fake news, myths, and rumours about COVID-19, to make sense of the right information, to identify the most relevant information, and to use the information in the right way (Fujii et al., 2020). Previous studies lack an understanding of the variations in IL that exist among refugees and immigrants within a social media context and how IL impacts IA; hence, we propose:

H4: *Information literacy has a significant effect on social media IA, such that the higher the level of literacy in a female refugee or immigrant is, the higher the likelihood she will be able to deal with unwanted and irrelevant information*

H5: *Information literacy has a significant effect on the user information satisfaction, such that the more literate a female refugee or immigrant is, the more likely she will be satisfied with the use of information*

2.5   Perceived COVID-19 Challenges

The literature on COVID-19–related factors – and how this affects people's level of user information satisfaction – is scarce. This is even more evident in the context of refugees and immigrants. Alea et al. (2020) explored the impact of perceived COVID-19 challenges in the higher education context and examined how an awareness of the challenges improves the learning experience. In this study, we aim to investigate how the challenges associated with the COVID-19 pandemic situation impact the level of user information satisfaction among female refugees and immigrants. We argue that the more familiar this target group is with the challenges of the situation, the more they will be able to control the information they receive from different sources and consequently the more satisfied they will be with the information obtained. Hence, we propose the following:

H6: *Perceived COVID-19 challenges have a significant effect on the level of user information satisfaction among female refugees and immigrants*

2.6   User Information Satisfaction

From a refugee and immigrant perspective, social media and ICT have normally been at the forefront of improvements in key factors concerning well-being, such as reducing isolation and stress, increasing social networking, improving feelings of agency, and establishing oneself in a new country (Udwan et al., 2020). However, refugees and immigrants are also exposed to some negative impacts of social media use, such as risky sexual behaviour, cyberbullying, and feelings of envy about others' lives (Anderson et al., 2020). In one study of Syrian refugees in the Netherlands (Udwan et al., 2020), the two sides of social media are presented – where it provides connectivity with family and friends on the one hand, and yet has an emotionally draining effect on the other (Udwan et al., 2020). The authors found that the use of social media impacts people differently and that individuals with a refugee and immigrant background hold different views on social media use and level of user satisfaction. Few studies have shown that users of social media adopt the IA strategy to deal with information overload (e.g. Lee et al., 2017; Sasaki et al., 2016). Lee et al.

(2017) suggested that when individuals are highly overloaded with news information from social media, they are likely to access only certain selective news sources. Further, Sasaki et al. (2016) found that when Twitter users face the challenge of IO, they avoid viewing all received tweets rather than reduce their number of friends to reduce the total number of received tweets. User information satisfaction decreases if people are overloaded with unwanted information or less relevant information, but the ability to cope with IF – and consequently to have an appropriate IA strategy in place to cope with IO – may lead to enhanced satisfaction because people then limit their information to only matter that is relevant and wanted. In summary, we argue that user information satisfaction among those with a refugee or immigrant background is not only impacted by traditional social media factors but also by individual-level (information literacy) and contextual (COVID-19) related factors which equally impact their level of information satisfaction (see Figure 1).
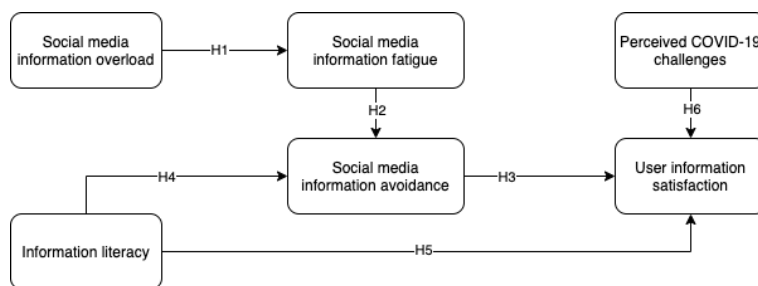
Figure 1: Proposed Conceptual Model

3   Research Methodology

3.1   Instrument and Data Collection

The items used to measure the constructs were all derived from previously validated measures, and if needed, were modified slightly to fit the study context. Alea et al.'s (2020, p. 134-136) items were used to measure COVID-19 challenges (3 items). Information overload (12 items), IF (8 items), and IA (7 items) were derived from Whelan et al. (2020), Bright et al. (2015), Dai et al. (2020), and Shin and Lin (2016), respectively. Items for information literacy (6 items) and user information satisfaction (5 items) were derived from Pinto et al. (2020) and Roberts (1999). We used an English language online survey to collect data between February and March,

2021. We sent more than 200 personal invitations, in addition to posting the survey link using different channels, such as authors' personal networks and social media platforms. We received 336 complete responses in total. As the main focus was on females with either a refugee or immigrant background living in Finland or Sweden, the final dataset consisted of 232 respondents who matched this criterion.

4 Results

4.1 Descriptive Results

The respondents' ages fell within the ranges of 18–25 years (10%), 26–35 years (56%), 36–45 years (30%), and 46–55 years (4%). The sample, as mentioned, consisted of 232 female respondents, who lived either in Finland (n = 73) or in Sweden (n = 159). Of the 232 respondents, 151 (65%) were immigrants and 81 (35%) were refugees. With regard to the current occupational status of the respondents, most were students (n = 124), followed by 48 respondents who indicated they were employed. Additionally, 42 respondents were non-employed, and 14 were self-employed. When we asked the respondents to indicate the social networking sites (SNSs) they used to access information, we found interesting results. The use of Facebook was mentioned by 210 respondents, and Instagram by 203 respondents. The least-used SNSs were Twitter (n = 58), TikTok (n = 57), Telegram (n = 16), and Snapchat (n = 62). We also asked the respondents to indicate how long they spend per day searching for and sharing information on SNSs. The findings were as follows: less than 30 minutes = 10 respondents; between 30 minutes and one hour = 72 respondents; between one and two hours = 89 respondents; and more than two hours = 61 respondents.

Table 1: Descriptive Statistics

Construct	Item	Loadings	(α)	CR	AVE
Perception of COVID- 19 challenges	COV_CH1	0.89			
	COV_CH2	0.84	0.79	0.86	0.69
	COV_CH3	0.76			
Social media information fatigue	Info-Fat2	0.79			
	Info-Fat3	0.73			
	Info-Fat4	0.81	0.92	0.94	0.68
	Info-Fat5	0.82			
	Info-Fat6	0.89			

	Info-Fat7	0.88			
	Info-Fat8	0.85			
	IL1	0.81			
	IL2	0.87			
Information literacy	IL3	0.81	0.92	0.94	0.71
	IL4	0.83			
	IL5	0.89			
	IL6	0.82			
	Inf_AV2	0.79			
	Inf_AV3	0.83			
Social media information avoidance	Inf_AV4	0.83	0.89	0.92	0.65
	Inf_AV5	0.84			
	Inf_AV6	0.83			
	Inf_AV7	0.73			
	Inf_OV10	0.83			
	Inf_OV12	0.77			
	Inf_OV2	0.81			
	Inf_OV3	0.76			
Social media information overload	Inf_OV4	0.79	0.94	0.95	0.64
	Inf_OV5	0.82			
	Inf_OV6	0.71			
	Inf_OV7	0.81			
	Inf_OV8	0.84			
	Inf_OV9	0.87			
	Info-SAT1	0.77			
	Info-SAT2	0.81			
User information satisfaction	Info-SAT3	0.81	0.82	0.87	0.58
	Info-SAT4	0.71			
	Info-SAT5	0.71			

The results showed that all factor loadings (except for 5 items) were above the threshold value of 0.70. In total, we used 41 items to measure the six constructs and retained 37 items for further analysis. All internal reliability and validity assessments, i.e. Cronbach's alpha (α), and composite reliability (CR), for all constructs were in the range of the recommended values of 0.70 and 0.70, respectively (Hair et al., 2019). The highest and the lowest CR values were for social media IO (0.95) and for perceived COVID-19 challenges (0.86). Also, the highest Cronbach's alpha (α) value was for social media IO (0.94), and the lowest value was for perceived COVID-19

challenges (0.79) (see Table 1). To examine whether all measures within each construct were theoretically related to one another as predicted, the convergent validity was assessed using the average variance extracted (AVE) values. As shown in Table 2, all the AVE values were above the recommended threshold of 0.50. Regarding the AVE values, the highest was for IL (0.71) and the lowest was for user information satisfaction (0.58). To establish the discriminant validity, the Fornell–Larcker criterion was used. We examined the AVE scores, and all AVE values were found to be lower than the shared variance for all model constructs; in this way, the discriminant validity was established (Fornell & Larcker, 1981).

Table 2: Discriminant validity (Fornell and Larcker criterion)

Constructs	PCCH	SMIA	IL	SMIO	UIS	SMF
Perception of COVID-19 challenges	**0.83**					
Social media information avoidance	0.15	**0.81**				
Information literacy	0.12	0.43	**0.84**			
Social media information overload	0.42	0.29	0.21	**0.81**		
User information satisfaction	0.23	0.36	0.47	0.48	**0.76**	
Social media information fatigue	0.28	0.51	0.21	0.69	0.38	**0.82**

PLS-SEM (partial least squares structural equation modelling) was used to evaluate and examine the path relationships in the model. The structural results showed that the independent variable, i.e. information use satisfaction has been explained by variance of 35%. The predictor variables, i.e. social media IF and social media IA, have been explained by variance of 47% and 37%, respectively. The SEM results showed that, as we predicted, social media IO had a positive and significant effect ($\beta = .69$; $t = 14.697$; $p = .001$) on social media IF, thus providing theoretical support to accept H1. Similarly, social media IF ($\beta = .44$; $t = 7.491$; $p = .001$) positively affected social media IA; thus, H2 was supported. Socia media IA significantly affected ($\beta = .18$; $t = 2.977$; $p = .005$) user information satisfaction, hence supporting H3 in the model. The analysis revealed interesting results when the path relationships from IL to IA and user information satisfaction were assessed. The path from IL to IA was positive ($\beta = .69$; t = 14.697; p = .001), hence supporting H4 in the model.

The direct path between IL to user information satisfaction was also significant (β = .38; t = 6.001; p = .001), indicating that the level of IL of the respondents plays a major role in their information satisfaction. Therefore, H5 was supported by the model. Finally, as we proposed and predicted, the respondent's perception towards COVID-19 challenges had a direct and positive effect (β = .15; t = 2.759; p = .005) on user information satisfaction, hence supporting H6 in the model.
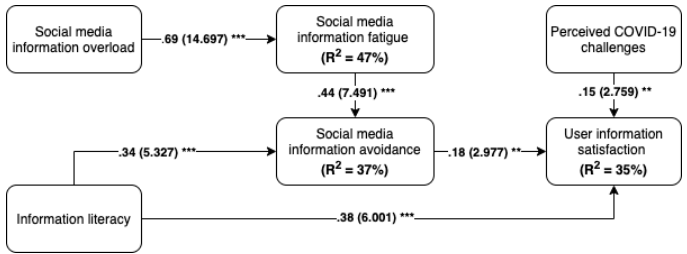
Figure 2: Structural model results. Notes: **p < .01. **p < .005. ***p < .001*

5   Discussions and Conclusion

The research model proposed in this research extends the (C–A–C) framework (Hilgard, 1980). The novelty of our proposed research model lies in the integration of the additional contextual construct, i.e. the perceived COVID-19 challenges, and the individual-level construct, i.e. literacy skills. To the best of our knowledge, these two variables have never been used to assess the user information satisfaction of females with a refugee or immigrant background. The results confirmed the effect of traditional constructs associated with social media use, such as IO to IF (Kim & Park, 2015) and IF to IA (Lee et al., 2017), and consequently IO to user information satisfaction (Dai et al., 2020). More importantly, the situational and contextual factors such as the perception of challenges associated with COVID-19 and the IL skills of the respondents were found to impact the level of user information satisfaction. As such, we contribute to the literature by showing that females with a refugee or immigrant background living in Finland and Sweden, while using different social media platforms and networking sites excessively to obtain and share information, nevertheless use information avoidance as a possible strategy for coping with information overload during the COVID-19 pandemic. This paper provides some practical implications too. For example, all respondents who participated in this research were female, living in Finland or Sweden, and with either

a refugee or immigrant background. As such, the findings might be applicable to policymakers to assist them in developing strategies for information management and information access in society. In addition, we showed that IF will likely result in IO behaviour among individuals (more specifically, females with the social status of being refugees or immigrants) who experience negative or unpleasant emotions when interacting with a huge amount of information. It is important that the providers of information in the social media platforms implement measures that could prevent users from experiencing negative satisfaction. However, this paper has some limitations. For example, we cannot claim that the results can be generalised, and our findings are applicable only to the context of this research. Moreover, we cannot solidly confirm the residency of the participants, as the responses were collected via an online and self-reported survey. Moreover, we cannot confirm if respondents clearly understood that their reported time on the internet was solely linked to reporting information retrieval activities.

References

Alea, L. A., Fabrea, M. F., Roldan, R. D. A., & Farooqi, A. Z. (2020). Teachers' Covid-19 Awareness, Distance Learning Education Experiences and Perceptions towards Institutional Readiness and Challenges. International Journal of Learning, Teaching and Educational Research, 19 (6), 127-144.

Aljanabi, A. R. A., & AL-Hadban, W. K. H. M. (2018). Information literacy and learners' satisfaction: The mediating role of social networking technology usage. Information and Learning Sciences, 119(11), 618-634.

Anderson, S., & Marguerite, D. (2020). Refugees and social media in a digital society. The Journal of Community Informatics, 16, 26-44.

Bright, L. F., Kleiser, S. B., Grau, S. L. (2015). Too much Facebook? An exploratory examination of social media fatigue. Computers in Human Behaviour, 44, 148-155.

Bronstein, J. (2019). Reframing integration information marginalization and information resistance among migrant workers. Journal of Documentation, 76(1), 27-48.

Cao, X., & Sun, J. (2018). Exploring the effect of overload on the discontinuous intention of social media users: An S-O-R perspective. Computers in Human Behaviour, 81, 10–18.

Cao, X., & Yu, L. (2019). Exploring the influence of excessive social media use at work: A three-dimension usage perspective. International Journal of Information Management, 46, 83–92.

Dai, B., Ali, A., & Wang, H. (2020). Exploring information avoidance intention of social media users: a cognition–affect–conation perspective. Internet Research, 30(5), 1455-1478.

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature. The Information Society: An International Journal, 20, 325-344.

Fisher, K.E., Erdelez, S. & McKechnie, L. eds. (2005). Theories of information behaviour. Information Today, Inc.

Fornell, C. G., and D. F. Larcker. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 18(1), 39–50.

Fujii, M. S., Hüttmann, J., Kutscher, N., & Friedrichs-Liesenkötter, H. (2020). Participation?! Educational Challenges for Young Refugees in Times of the COVID-19 Pandemic. Media Education, 11(2), 37-47.

Guo, Y., Lu, Z., Kuang, H., & Wang, C. (2020). Information avoidance behaviour on social network sites: Information irrelevance, overload, and the moderating role of time pressure. International Journal of Information Management, 102067.

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

Hilgard, E. R. (1980). The trilogy of mind: cognition, affection, and conation. Journal of the History of the Behavioural Sciences, 16(2), 107-117.

Huitt, W. (1999). Conation as an important factor of mind. Educational Psychology Interactive. Available at: http://www.edpsycinteractive.org/topics/conation/conation.html. (Accessed 18 May 2021).

Karim, M., Widén, G., & Heinström, J. (2019). Influence of demographics and information literacy self-efficacy on information avoidance propensity among youth. In Proceedings of CoLIS, the Tenth International Conference on Conceptions of Library and Information Science, Ljubljana, Slovenia, June 16-19, 2019.

Khan, K.S., & Eskola, E. (2020). The cultural landscape of women refugees in Sweden - a road to information and integration. In Proceedings of ISIC, the Information Behaviour Conference, Pretoria, South Africa. Information Research, 25(4).

Kim, S., & Park, H. (2015). Empirical study on antecedents and consequences of users' fatigue on SNS and the moderating effect of habit. Journal of Information Technology Services, 14(4), 137-157.

Lee, S. K., Lindsey, N. J., & Kim, S. K. (2017). The effects of news consumption via social media and news information overload on perceptions of journalistic norms and practices. Computers in Human Behaviour, 75, 254-263.

Liebig, T. and Tronstad, K. R. (2018). Triple Disadvantage? A first overview of the integration of refugee women," OECD Social, Employment and Migration Working Papers, 216, OECD Publishing, Paris.

Lloyd, A., Mary, A. K., Kim, M. T., & Asim Q. (2013). Connecting with New Information Landscapes: Information Literacy Practices of Refugees. Journal of Documentation, 69 (1), 121– 144.

Martzoukou, K. and Burnett, S., (2018). Exploring the everyday life information needs and the socio-cultural adaptation barriers of Syrian refugees in Scotland. Journal of Documentation, 74 (5), 1104 -1132.

McCloud, R. F., Jung, M., Gray, S. W., & Viswanath, K. (2013). Class, race and ethnicity and information avoidance among cancer survivors. British Journal of Cancer, 108(10), 1949.

Narayan, B., Case, D. O., & Edwards, S. L. (2011). The role of information avoidance in everyday-life information behaviours. Proceedings of the American Society for Information Science and Technology, 48(1), 1–9.

Pinto, M., Caballero, D., Sales, D., & Fernández-Pascual, R. (2020). MOBILE-APPS questionnaire: Developing and validating a scale to measure the attitudes and perceptions of undergraduate students on mobile information literacy. Journal of Librarianship and Information Science, 52 (4), 1063-1072.

Ravindran, T., Yeow Kuan, A. C., & Hoe Lian, D. G. (2014). Antecedents and effects of social network fatigue. Journal of the Association for Information Science and Technology, 65(11), 2306–2320.

Roberts, E. S. (1999). In defence of the survey method: An illustration from a study of user information satisfaction. Accounting & Finance, 39 (1), 53-77.

Sasaki, Y., Kawai, D., & Kitamura, S. (2016). Unfriend or ignore tweets? a time series analysis on Japanese twitter users suffering from information overload. Computers in Human Behaviour, 64, 914-922.

Savolainen, R. (2007). Filtering and withdrawing strategies for coping with information overload in everyday contexts. Journal of Information Science, 33 (5), 611-621.

Shin, W., & Lin, T. T. C. (2016). Who avoids location-based advertising and why? Investigating the relationship between user perceptions and advertising avoidance. Computers in Human Behaviour, 63, 444-452.

Swar, B., Hameed, T., & Reychav, I. (2017). Information overload, psychological ill-being, and behavioural intention to continue online healthcare information search. Computers in Human Behaviour, 70, 416-425.

Turan, Z., Tinmaz, H., & Goktas, Y. (2013). The Reasons for Non-Use of Social Networking Websites by University Students. Comunicar, 21(41), 137–145.

Udwan, G., Leurs, K., & Alencar, A. (2020). Digital resilience tactics of Syrian refugees in the Netherlands: Social media for social support, health, and identity. Social Media + Society, 6(2), 2056305120915587.

Whelan, E., Islam, N., Brooks, S. (2020). Is boredom proneness related to social media overload and fatigue? A stress–strain–outcome approach. Emerald Insight, 30 (3), 869-887.

Yu, L., Cao, X., Liu, Z., & Wang, J. (2018). Excessive Social Media Use at Work Exploring the Effects of Social Media Overload on Job Performance. Information Technology & People, 31 (6): 1091–1112.

Zhang, S., Zhao, L., Lu, Y., & Yang, J. (2016). Do you get tired of socializing? An empirical explanation of discontinuous usage behaviour in social network services. Information & Management, 53(7), 904–914

Evaluating an Implementation Protocol for Digitization and Devices in Operating Rooms: A Case Study

Navin Sewberath Misser,[1] Joris Jaspers,[2]
Bas van Zaane[2], Hein Gooszen[3] &
Johan Versendaal[1]

[1] HU University of Applied Sciences Utrecht, University Medical Center Utrecht, Utrecht, The Netherlands; e-mail: Navin.SewberathMisser@hu.nl, Johan.Versendaal@hu.nl
[2] University Medical Center Utrecht, Utrecht, The Netherlands; e-mail: J.Jaspers@umcutrecht.nl, B.vanZaane@umcutrecht.nl
[3] Radboud University Medical Center, Nijmegen, The Netherlands; e-mail: Hein.Goozen67@gmail.com

Abstract Digitization of activities in hospitals receives more attention, due to Covid-19 related regulations. The use of e-health to support patient care is increasing and efficient ways to implement digitization of processes and other technological equipment are needed. We constructed a protocol for implementation and in this study, we evaluate this protocol based on a case to implement a device in the OR. We used various data sources to evaluate this protocol: semi-structured interviews, questionnaires, and project documents. Based on these findings, this protocol, including identified implementation activities and implementation instructions can be used for implementations of other devices. Implementation activities include setting up a project plan, organizational and technological preparation, maintenance, and training. In future research, these activities and instructions need to be evaluated in more complex projects and a flexible tool needs to be developed to select relevant activities and instructions for implementations of information systems or devices.

1   Introduction

Digitizing health care activities within hospitals to support hospital and patient care have been of increasing interest due to the Covid-19 pandemic and related regulations. The Covid-19 pandemic shows the need for rapid implementation of digitized processes, information systems or devices in hospitals (Meyer *et al.*, 2020; Rodriguez Socarrás *et al.*, 2020). Digitizing activities or processes generally require well-planned development activities and implementation of digitized processes require well-prepared implementation activities in order to reach identified goals and to improve adoption among users (Fennelly *et al.*, 2020). Edmondson (2001) describes the implementation of technological equipment as the integration of new technologies in day-to-day activities in an organization (Edmondson, Bohmer and Pisano, 2001). Technological equipment includes technological devices and (medical) information systems. To support implementation of technological devices and digitization in hospitals, such as telehealth, electronic health records, management information systems, we constructed a protocol for implementation with a focus on the Operating Room department (OR) in hospitals (Dutch Hospital Association, 2016). This protocol consists of implementation factors, implementation activities, and implementation instructions (Sewberath Misser *et al.*, 2020). These factors, activities and instructions are based on a systematic literature review and a survey completed by scrub nurses and circulating nurses (Sewberath Misser, Jaspers, *et al.*, 2018; Sewberath Misser, Zaane, *et al.*, 2018). The purpose of this study is to evaluate and refine this protocol for implementation and the research question for this study described as:

- To which extent is our protocol for implementation ready for use in practice, based on real life case studies?

To address this question, we describe the method and research instruments in the second section of this article. In the third section, we introduce a case and in section four, we evaluate our protocol for implementation based on implementation experiences and results. Finally, we will draw conclusions and describe possibilities for future research.

N. *Sewberath Misser, J. Jaspers, B. van Zaane, H. Gooszen & J. Versendaal:*
Evaluating an Implementation Protocol for Digitization and Devices in Operating Rooms: a Case Study

353

2   Method

In previous studies, we used focus groups with experts to evaluate this protocol for implementation. In this study, we address the research question by focusing on the evaluation of this protocol for implementation in actual projects. This study consisted of three stages: 1) setting up a study procedure, 2) data gathering, 3) data processing, and analysis.

2.1   Setting up a study procedure

We set up a study procedure consisting of sections regarding general information, procedures, research instruments and data analysis guidelines (Maimbo and Pervan, 2005; Yin, 2018). We selected a project for use of the protocol for implementation based on scope, implementation period and feasibility. Projects or cases entailed the implementation of a new device or digitization of a process in the OR, with a limited number of stakeholders during implementation. These cases needed to be implemented between March and April 2020. The selected case for this research involved using the protocol for a pilot study to introduce an exoskeleton for surgical supporting staff. A project leader was assigned to implement an exoskeleton in the OR for selected surgeries. The timeframe for data collection and reporting was extended up until December 2020.

2.2   Data gathering

In our study procedure, we considered and selected different instruments to gather data and to ensure quality and rigor: semi-structured interviews, questionnaires, and project documents.

1. Interview with a project leader. In a semi-structured interview, we focused on clearness, completeness, and ease of use of included factors, activities and instructions for implementation. The interview was digitally conducted with MS Teams due to Covid-19 measures.

2. Questionnaires. We composed questionnaires based on the technology assessment model, in which we focus on the intended use, perceived ease of use, and perceived usefulness. These questions could be scored on a lickert 5-points

scale and participants were able to add comments to clarify their responses (Heijden, 2004; Wu and Wang, 2005; Gagnon et al., 2012; Tantiponganant and Laksitamas, 2014). We developed two sets of questionnaires respectively for project leaders and users. In the questionnaire for project leaders, we focused on the use of the implementation protocol and the questionnaire for users had a focus on the implemented tool.

3. Project documents. Project documents created during and after completion of the project relating to the implementation of the device were used as data source.

2.3   Data processing and analysis

Collected questionnaires were processed in MS Excel and the interview with the project leader was video recorded and transcribed in MS Word. This interview was conducted in the Dutch language. Evaluation results based on this case are described according to the structure of the protocol for implementation. Following the analysis of these results, suggestions for refinement for the protocol for implementation are provided.

3   Case: implementation study of the Leavo Exoskeleton

An exoskeleton is a wearable, mechanical external structure that enhances or supports the power of a person. Exoskeletons can be either 'active' or 'passive'. Active exoskeletons enhance human power with use of for example electric motors, hydraulic actuators or other types of power. A passive exoskeleton is a mechanical structure using materials such as springs, belts or dampers to support a posture or a motion (Looze de *et al.*, 2016). The Leavo exoskeleton (see figure 1) can be classified as a passive exoskeleton, which supports chest and back. This wearable relieves back and spine muscles and which should reduce back pain and increase durability of people who frequently carry heavy items or keep static positions (Koopman *et al.*, 2019).

N. Sewberath Misser, J. Jaspers, B. van Zaane, H. Gooszen & J. Versendaal:
Evaluating an Implementation Protocol for Digitization and Devices in Operating Rooms: a Case Study

355
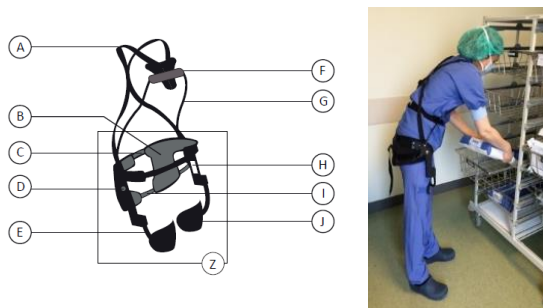
Figure 1: Leavo exoskeleton
Legend: A Suspender ; B Hip pads ; C Hip belt; D Smart joint ; E Leg structure; F Chest pad; G
Torso structure; H Label ; I Buck belt; J Leg pad; ; Z Hip assembly
Source: http://www.Leavo.nl

In the OR, scrub nurses and circulating nurses prepare surgeries by setting up surgical instruments prior to surgeries. These instruments are stored in metal instrument baskets, which vary in weight. Depending on the surgical discipline, it often occurs that scrub nurses keep static positions during a surgical procedure. For the purpose of this study, the hospital (client) acquired four exoskeletons for use by scrub and circulating nurses in the OR and the client defined the data collection period. The novelty of this study is that this exoskeleton was used for the first time in an OR-setting. The client and the human resources department (HR) recruited and assigned a project leader. The first author informed the project leader via e-mail about the study procedure, the protocol for implementation, and the data gathering process. In a briefing session, the implementation protocol was explained, as well as the study procedure. As part of this study, the project leader used the protocol for implementation of the device, to complete the questionnaire for project leaders, and to distribute and collect questionnaires for users. Together with the HR-department, the project leader recruited four users for this device. For the purpose of our study, we interviewed the project leader after completion of the implementation. The project leader completed a questionnaire and users of the exoskeleton completed two out of four distributed questionnaires.

4 Evaluation results

The protocol for implementation consists of five factors for implementation, with related implementation activities and instructions for implementation. The factors for implementation are: 1.) setting up a plan, 2.) organizational preparation, 3.) technological preparation, 4.) maintenance, and 5.) training and evaluation. In the next paragraphs, we describe evaluation results regarding of the use of this protocol based on the introduction of an exoskeleton.

4.1 Evaluating implementation factor: set up project plan

The first factor for implementation refers to setting up a project plan. The interview with the recruited project leader shows that implementation activities such as 1.1 identifying strategic and tactical topics, and 1.2 identify performance, were determined in previous stages of the implementation project. The activities 1.3 identifying stakeholders and 1.4 identifying risks evolved during the implementation process, as the number of stakeholders increased as the project progressed. Identified stakeholders were client, HR, researchers, users of the device. During the interview, the project leader stated that these activities and instructions were clearly described, complete, and ready for use. In table 1, implementation activities for the first implementation factor are described.

Table 1: Factor 1: set up a project plan and related activities

Id	Description of activities
1.1	Identify strategic and tactical topics
1.2	Identify performance
1.3	Identify stakeholders
1.4	Identify Risks
1.5	Identify activities for implementation

4.2 Evaluating implementation factor: organizational preparation

The project leader was responsible for the organizational preparation related to the introduction of this device. Together with stakeholders (client, HR and OR-team), three types of surgeries were selected to use this exoskeleton: vascular surgery, orthopedic surgery and cardiothoracic surgery. These surgeries were selected based

N. Sewberath Misser, J. Jaspers, B. van Zaane, H. Gooszen & J. Versendaal:
Evaluating an Implementation Protocol for Digitization and Devices in Operating Rooms: a Case Study

357

on the duration of surgeries, positioning of the scrub nurse during surgeries and usage of instruments. The project leader assembled an implementation team (see table 2, activity 2.1) by recruiting four scrub nurses to use an exoskeleton prior to and during surgeries. The project leader was able to foster team familiarity (activity 2.2), as she provided instructions how to use the device and as she responded to users' queries. After the introduction of the device, scrub nurses were able to identify the affected activities (activity 2.3) caused by the new device, such as preparatory activities to assemble and to wear the device. According to the project leader, existing checklists or procedures completed by scrub nurses or circulating nurses were not updated (activity 2.4). She stated that simulations or sessions to practice (activity 2.5) were scheduled to identify the performance of the device and to assess whether the project goals could be met. In the interview, the project leader expected a gradual increase in adoption of the device. She expected an increased use of the device, as the intention of this device was to provide support during lifting and static positions. In contrast to her expectation, her encouragement and guidance was needed to convince users to use the device. This encouragement was needed due to some technical difficulties and extra work (activities 2.6, 2.7 and 2.8). After completion of the project, scrub nurses completed questionnaires and they confirmed that the project leader was responsive and available for questions and guidance. This evaluation shows that identified activities and instructions, related to the implementation factor organizational preparation, are ready to be used in practice.

Table 2: Factor 2: Organizational preparation and related activities

Id	Description of activities
2.1	Assemble a multidisciplinary implementation team
2.2	Foster team familiarity
2.3	Identify affected activities and/or processes
2.4	Update checklists
2.5	Perform simulations
2.6	Identify and deploy activities to increase employees' engagement
2.7	Identify and deploy activities to increase employees' adoption
2.8	Communicate with stakeholders

4.3   Evaluating implementation factor: technological preparation

The third implementation factor, related activities, and instructions involve the technological preparation of the device and its environment. To prepare the device for use, the manufacturer of the exoskeleton tailored and adjusted each device to each users' body type (activity 3.1 in table 3). Ergonomic aspects for use were considered, according to the project leader (3.2) as the device supported static positions and heavy lifting (see figure 1). With reference to the information systems (IT) environment, no interfaces were needed and no electronic data was generated, as the exoskeleton is classified as a mechanical device (activities 3.3 and 3.5). As the project progressed, integration of the device in the existing working environment (activity 3.4) was increasingly relevant after introduction. During the course of the project, various troubleshooting challenges occurred: when lead aprons were used during surgeries to reduce effects of x-rays, the exoskeletons were difficult to adjust and wear. In simulations and during execution of regular activities, users had trouble with rotating movements when wearing the device (activity 3.6).

Table 3: Factor 3: Technological preparation and related activities

Id	Description of activities
3.1	Prepare equipment
3.2	Consider ergonomic aspects
3.3	Prepare interfaces with other information systems
3.4	Integrate device within existing environment
3.5	Manage generated data
3.6	Interpret screens and troubleshooting

4.4   Evaluating implementation factor: maintenance

As part of the implementation protocol, an activity setting up a maintenance plan (activity 4.1 in table 4) is included. In the interview, the project leader stated that she did not set up a maintenance plan for the exoskeleton. She addressed safety issues regarding use of the device during instructions. Updates of safety regulations were not addressed in this stage of the project.

N. Sewberath Misser, J. Jaspers, B. van Zaane, H. Gooszen & J. Versendaal:
Evaluating an Implementation Protocol for Digitization and Devices in Operating Rooms: a Case Study

359

Table 4:Factor 4: Maintenance and related activities

Id	Description of activities
4.1	Set up maintenance plan
4.2	Update safety (regulations)

4.5 Evaluating implementation factor: training

The final factor in the protocol for implementation refers to training activities (activity 5.1 in table 5), assessing skills (activity 5.2) and evaluating experiences (activity 5.3). Scrub nurses were trained to assemble, use, and disassemble the device. According to the project leader, attention and supervision was needed to adjust the exoskeleton properly, for optimal use of the device during observed surgeries. Reports regarding the use and functionality of the exoskeleton were gathered and reported to the client and the manufacturer. These reports mainly referred to the intended use of the device. Two scrub nurses completed a questionnaire to reflect on the implementation of the device.

Table 5. Factor 5: Training and evaluation, and related activities

Id	Description of activities
5.1	Train involved staff
5.2	Assess Skills
5.3	Evaluate experiences

4.6 Evaluation of the protocol: perceived ease of use and perceived usefulness

The questionnaire for project leaders focused on the perceived ease of use and perceived usefulness of the protocol for implementation. The project leader stated in a completed questionnaire that activities and instructions were clearly structured, clearly described, and ready for use. In the interview, the project leader suggested a more user-friendly layout for this protocol in general, because the appearance and structure of the used protocol had a scientific lay out. She proposed to omit referrals to scientific literature and proposed to simplify some sentences to improve user-friendliness. The project leader stated that different factors and activities were helpful to prepare and to introduce this new device. She also found that the protocol provides flexibility to adjust to this project or other implementation projects, by

choosing relevant activities and implementation instructions. With reference to usefulness of activities and related instructions, the project leader agrees fully with the statement that the use of a protocol can improve efficiency and increase adoption of new devices with users. Users indicated in completed questionnaires that they were not informed of the use of an implementation protocol. One user, with more than 20 years of experience as a scrub nurse, stated that the introduction of this device was performed better than previous implementations. This scrub nurse indicated that this implementation performance was caused by the project leaders' involvement, as she was available for questions and instructions.

5   Discussion

In hospital environments, specifically in OR's, surgeons and other involved staff such as scrub nurses and circulating nurses use information systems and technological devices to support or execute surgeries. However, possibilities for digitization of supporting activities remain a topic of interest and research continues (Fennelly *et al.*, 2020; Rodriguez Socarrás *et al.*, 2020; Scott *et al.*, 2020; Beiser *et al.*, 2021). The focus of this study was to evaluate an implementation protocol with a case to introduce an exoskeleton for use by scrub and circulating nurses. With reference to the first implementation factor 'set up a project plan' and activities, evaluation results show that the implementation stage of a project is preceded by several other project activities and project stages. Activities such as identifying strategic topics, performance and stakeholders (activities 1.1 – 1.3) were addressed in previous stages of the project and prior to implementation. Examples of stakeholders are project leader, client, and human resources. Based on these evaluation results, we propose a change in the descriptions of included activities. In the implementation stage of the project, focus should be on topics and performance criteria related to the *implementation* of the device. Regarding the second factor 'organizational preparation', various activities were deployed to recruit users. In practice, many potential users refused to participate, possibly caused by social pressure, fear of wearing a shield, or fear for an uncomfortable fit. Activities related to the third factor 'technological preparation' were addressed, with focus on the activities preparing equipment, considering ergonomic aspects and integration within the existing environment. The last factor for implementation, training, was operationalized by providing instructions and simulations. Training plans and assessment plans were not developed for this device. Based on these evaluation

N. Sewberath Misser, J. Jaspers, B. van Zaane, H. Gooszen & J. Versendaal:
Evaluating an Implementation Protocol for Digitization and Devices in Operating Rooms: a Case Study

361

results of this protocol, we consider two findings: 1.) implementation activities are sorted per factor and 2.) functionality and user-friendly design of a tool affect implementation success and adoption.

Finding 1: implementation activities are sorted per factor.

In the current protocol for implementation, activities and instructions are grouped according to theme or implementation factor. Results show that many activities are not performed sequentially and some executed activities need adjustment during the implementation process. For example, preparation activities involving technology, organization, and training are interconnected: when the manufacturer tailored the exoskeletons to the user's body type, users were instructed and users practiced with the device. Activities may need adjustment during the implementation process for example changes in stakeholders, implementation team, and communication activities.

Finding 2: functionality and user-friendly design affect implementation success and adoption.

Implementation of a device in an organization requires effort from involved stakeholders and users. Following the technology assessment model, we argue that functionality and user-friendly design should address a specific need of users within an organization. Considering these aspects during the development process of the tool, will affect adoption and implementation success (Gagnon *et al.*, 2012). Based on the results of this case, a proven technology or device from a specific sector might not be transferrable to another sector or context due to situational factors or other environmental aspects.

6   Conclusions, limitations and future research

In this study, we addressed the question to which extent a protocol for implementation was ready for use in practice. Therefore, we evaluated this protocol by using this protocol in a small-scale project to implement an exoskeleton in OR's. We conclude that implementation activities and implementation instructions included in this protocol are useful, complete, and ready for use in more complex projects. Refinement of this protocol can be achieved by clarifying instructions and

removing scientific references. Although this study was carefully prepared and executed, several limitations can be identified. The intention was to evaluate this protocol with a case to digitize pathology inquiries at the hospital laboratory. This project was discontinued due to Covid-19 measures and priorities. We argue, that included activities in our protocol for implementation are relevant and similar for the digitizing activities in hospitals. In previous studies, we identified and relevant implementation activities and instructions. We based these activities and instructions on a literature research and questionnaire, in which we included implementations of information systems, electronic healthcare records and digitized processes in hospitals (Rivkin, 2009; Ehrenfeld and Rehman, 2011). Although results and findings to this case study are based on a small case and cross case analysis was not possible, we assured data quality and rigor by using various sources of data as triangulation measures. Data collection was only conducted and analyzed after the device was implemented and after the protocol had been used according to the study procedure. In future research, this implementation protocol needs to be evaluated in other projects with increased complexity. Other future research should include refinement of this protocol based on the first finding, in particular, the development of a tool to select and sort implementation activities and instructions based on user preference and tailored to context.

References

Beiser, M. et al. (2021) 'Electronic Health Record Usage Patterns: Assessing Telemedicine's Impact on the Provider Experience During the COVID-19 Pandemic', Telemedicine and e-Health, 00(00), p. tmj.2020.0490. doi: 10.1089/tmj.2020.0490.

Dutch Hospital Association (2016) Convenant Veilige toepassing van medische technologie in de medisch specialistische zorg. Bilthoven: Barnyard Creative Powerhouse. Available at: http://www.nfu.nl/img/pdf/16.11163_NVZ-NFU_ea_Convent_Medische_Technologie_-_definitieve_versie_2016.pdf.

Edmondson, A. C., Bohmer, R. M. and Pisano, G. P. (2001) 'Disrupted Routines: Team Learning and New Technology Implementation in Hospitals', Administrative Science Quarterly, 46(4), p. 685. doi: 10.2307/3094828.

Ehrenfeld, J. M. and Rehman, M. A. (2011) 'Anesthesia information management systems: a review of functionality and installation considerations.', Journal of clinical monitoring and computing. Netherlands, 25(1), pp. 71–79. doi: 10.1007/s10877-010-9256-y.

Fennelly, O. et al. (2020) 'Successfully implementing a national electronic health record: a rapid umbrella review', International Journal of Medical Informatics. Elsevier B.V., 144(September), p. 104281. doi: 10.1016/j.ijmedinf.2020.104281.

Gagnon, M. P. et al. (2012) 'Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system', Telemedicine and e-Health, 18(1), pp. 54–59. doi: 10.1089/tmj.2011.0066.

N. Sewberath Misser, J. Jaspers, B. van Zaane, H. Gooszen & J. Versendaal:
Evaluating an Implementation Protocol for Digitization and Devices in Operating Rooms: a Case Study

363

Heijden, H. van der (2004) 'User Acceptance of Hedonic Systems', MIS Quarterly, 28(4), pp. 695–704.

Koopman, A. S. et al. (2019) 'Effects of a passive exoskeleton on the mechanical loading of the low back in static holding tasks', Journal of Biomechanics, 83(3), pp. 97–103. doi: 10.1016/j.jbiomech.2018.11.033.

Looze de, M. P. et al. (2016) 'Exoskeletons for industrial application and their potential effects on physical work load', Ergonomics, 59(5), pp. 671–681. doi: 10.1080/00140139.2015.1081988.

Maimbo, H. and Pervan, G. (2005) 'Designing a Case Study Protocol for application in IS research', 9th Pacific Asia Conference on Information Systems: I.T. and Value Creation, PACIS 2005, pp. 1281–1292.

Meyer, B. C. et al. (2020) 'Medical Undistancing Through Telemedicine: A Model Enabling Rapid Telemedicine Deployment in an Academic Health Center During the COVID-19 Pandemic', Telemedicine and e-Health, 00(00), pp. 1–10. doi: 10.1089/tmj.2020.0327.

Rivkin, G. (2009) 'Challenges of Technology Integration and Computer-Assisted Surgery', The Journal of Bone and Joint Surgery (American). United States, 91(Supplement_1), p. 13. doi: 10.2106/JBJS.H.01410.

Rodriguez Socarrás, M. et al. (2020) 'Telemedicine and Smart Working: Recommendations of the European Association of Urology', European Urology, 78(6), pp. 812–819. doi: 10.1016/j.eururo.2020.06.031.

Scott, B. K. et al. (2020) 'Advanced digital health technologies for COVID-19 and future emergencies', Telemedicine and e-Health, 26(10), pp. 1226–1233. doi: 10.1089/tmj.2020.0140.

Sewberath Misser, N., Zaane, B. Van, et al. (2018) 'Implementing Medical Technological Equipment in the OR : Factors for Successful Implementations', Journal of healthcare engineering, 2018. doi: https://doi.org/10.1155/2018/8502187.

Sewberath Misser, N., Jaspers, J., et al. (2018) 'Transforming operating rooms: factors for successful implementations of new medical equipment', Digital Transformation – Meeting the challenges, (June), pp. 279–289. doi: 10.18690/978-961-286-170-4.18.

Sewberath Misser, N. et al. (2020) 'A protocol for the implementation of new technology in a highly complex hospital environment: the operating room', International Journal of Networking and Virtual Organisations, 22(2), p. 199. doi: 10.1504/IJNVO.2020.105543.

Tantiponganant, P. and Laksitamas, P. (2014) 'An analysis of the technology acceptance model in understanding students' behavioral intention to use university's social media', Proceedings - 2014 IIAI 3rd International Conference on Advanced Applied Informatics, IIAI-AAI 2014, 12, pp. 8–12. doi: 10.1109/IIAI-AAI.2014.14.

Wu, J. H. and Wang, S. C. (2005) 'What drives mobile commerce? An empirical evaluation of the revised technology acceptance model', Information and Management, 42(5), pp. 719–729. doi: 10.1016/j.im.2004.07.001.

Yin, R. (2018) Case Study Research and Applications Design and Methods. 6th edn. Sage Publications Limited

WHO ELSE DO YOU NEED FOR A DATA-DRIVEN BUSINESS MODEL? EXPLORING ROLES AND EXCHANGED VALUES

FLORIAN LESKI,[1] MICHAEL FRUHWIRTH[2] & VIKTORIA PAMMER-SCHINDLER[1]

[1] Graz University of Technology, Graz, Austria, e-mail: florian.leski@student.tugraz.at, viktoria.pammer-schindler@tugraz.at
[2] Know-Center GmbH, Graz, Austria, e-mail: mfruhwirth@know-center.at

Abstract The increasing volume of available data and the advances in analytics and artificial intelligence hold the potential for new business models also in offline-established organizations. To successfully implement a data-driven business model, it is crucial to understand the environment and the roles that need to be fulfilled by actors in the business model. This partner perspective is overlooked by current research on data-driven business models. In this paper, we present a structured literature review in which we identified 33 relevant publications. Based on this literature, we developed a framework consisting of eight roles and two attributes that can be assigned to actors as well as three classes of exchanged values between actors. Finally, we evaluated our framework through three cases from one automotive company collected via interviews in which we applied the framework to analyze data-driven business models for which our interviewees are responsible.

1   Introduction

The increasing volume of available data and the advances in analytics and artificial intelligence hold the potential for competitive advantage, business growth, and new business models. Thus, also offline-established organizations are seeking new so-called data-driven business models (DDBMs). This innovation and transformation process is often challenging, as it requires new skills and capabilities (e.g., data science or IT infrastructure), deep relationships, and partner information ecosystems (Schüritz et al., 2017). These interdependencies and the complexity of the ecosystem increase the risk in business model innovation (Dellermann et al., 2017).

In general, tools and methods support the process of business model innovation. Still, there is a limited amount of tool support that explicitly focuses on data utilization and mainly supports idea generation (Fruhwirth, Ropposch, Pammer-Schindler, 2020). Furthermore, to successfully implement DDBMs, it is crucial to understand the environment and involved stakeholders, as companies will collaborate more and increase their dependencies (Hunke et al., 2017). Typically, traditional firms rely on new external partners for their DDBM, such as a cloud- or data provider. Thus, it is important to know the roles that need to be fulfilled by actors in the own business model. Business model representations with a transactional focus are useful to understand, develop, and model business models (Fruhwirth et al., 2021; Täuscher and Abdelkafi, 2017). There, types of actors and exchanged values support modeling a business model (Terrenghi et al., 2018). Despite this, existing research on DDBMs mostly overlooks the partner and ecosystem perspective. Accordingly, we ask the following research question: *What roles do exist in a data-driven business model, and how can the exchanged values be categorized?*

To answer this research question, we conducted a structured literature review and derived a framework with a set of eight roles and two attributes that can be assigned to actors as well as three classes of exchanged values. The framework was evaluated by applying it to three use cases from one company in the automotive industry.

F. Leski, M. Fruhwirth & V. Pammer-Schindler:
Who Else do You Need for a Data-Driven Business Model? Exploring Roles and Exchanged Values

367

2   Background

Data-driven business models are digital business models with a conceptual focus on value creation with data (Guggenberger et al., 2020). DDBMs rely on data as a key resource and apply data analytics techniques as key activities to discover insights from data and that are transformed into a data-based value proposition that supports customers in their decision-making process (Hartmann et al., 2016; Kühne and Böhmann, 2019; Schüritz et al., 2019). Other researchers denote such models as »data-infused business models« (Schüritz and Satzger, 2016) or »data-driven services« (Azkan et al., 2020). Data-driven services could be either offered as stand-alone or as an add-on to existing products or services (Breitfuß et al., 2019; Wixom and Ross, 2017). Existing conceptualization and classification approaches of DDBMs have a company-centric focus, studying the value creation process with types of data sources and key activities related to data and analytics (e.g., Hartmann et al., 2016), the value proposition (e.g., Fruhwirth, Breitfuß, Pammer-Schindler, 2020) or the value delivery with service flows or platform types (e.g., Azkan et al., 2020). Studying and developing DDBMs also involves the ecosystem perspective with involved roles and actors (Hunke et al., 2017). Little attention has been paid to that in contemporary research on DDBMs.

A business model can be understood as »*an architecture of the product, service and information flows, including a description of the various business actors and their roles; a description of the potential benefits for the various business actors; a description of the sources of revenues*« (Timmers, 1998, p. 4). These **actors** are economic independent entities and »*exchange value objects, which are services, products, money, or even consumer experiences. A value object is valuable to one or more actors.*« (Gordijn and Akkermans, 2001, p. 13). A business model also can be seen as a set of activities, performed by the focal organization itself, by its customers, suppliers, and/or partners (Zott and Amit, 2010). Thus, every actor has one or more **roles**, that describe an actor's activities, functions, or contributions in the business model (Terrenghi et al., 2018). This understanding of business models takes a network-centric and transactional view, focusing on value exchange among actors. Similar concepts in that regard have been established, such as the »value network« (Allee, 2008), or the »business ecosystem« (Jacobides et al., 2018).

3   Methodology

Aiming to identify types of actors and classes of exchanged values in DDBMs, we conducted a structured literature review and adopted an inductive category formation approach to analyze and synthesize the literature. Finally, we evaluated our framework in three cases from one automotive company.

We based our **search and selection process** on the guidelines and recommendations of Vom Brocke et al. (2009) and Webster and Watson (2002). We started with a database search as summarized in Table 1. Our search strings were informed by previous literature denoted in the background section. We searched separately for papers dealing with actors and exchanged values in ecosystems. By applying the stated logical search terms in the respective databases[1], 2513 articles were found; 917 for actors, 1496 regarding exchanged values.

Table 1: Summary of the Database Search Results

			AISel	IEEE Xplore	Science Direct	Scopus	Web of Science	ACM
("digital" OR "data-driven" OR "data-infused" OR "data-based") AND ("business model" OR "service")	AND	(role" OR "actor" OR "partner")	300	22	230	162	118	85
		("network" OR "ecosystem" OR "value chain")	281	214	258	306	193	244

Further, we applied a three-step selection process: First, we selected 119 relevant articles based on their titles. Second, we scanned the abstracts of these selected papers for relevance, limiting them to 62 articles. Third, we read the full text of the remaining papers and made a final selection of 26 articles that are relevant for our

---

[1] Search strings were applied to title, abstract, keywords and/or full text depending on the database to retrieve a manageable number of articles per query. A detailed description can be provided by the authors upon request.

F. Leski, M. Fruhwirth & V. Pammer-Schindler:
Who Else do You Need for a Data-Driven Business Model? Exploring Roles and Exchanged Values

369

research. We also conducted a forward and backward search (Webster and Watson, 2002), leading to an additional set of 11 publications. Therefore, we arrived at a final sample of 38 articles (17 for actors, and 21 for exchanged values), whereas five articles were present in both categories, resulting in 33 articles[2] without overlaps.

After the search and selection process, we **analyzed and synthesized** the selected literature following an inductive category formation approach (Mayring, 2015). We examined the papers' content aiming to define distinct roles for actors and classes of exchanged values present in DDBMs. Initially, we specified that the level of abstraction of the resulting classes must be generic so that they can be applied to a broad spectrum of industries. We analyzed the material focusing on the results, findings, conclusions, figures, and tables and summarized the essential parts of the material for both actors and exchanged values. Subsequently, we synthesized this interim outcome to a generic set of categories, consisting of ten roles and two attributes that can be assigned to actors as well as three classes of exchanged values. After the evaluation, two roles were dropped or merged with other roles.

We **evaluated** and refined our framework in three use cases from the automotive industry. Therefore, we conducted three semi-structured interviews with managers from one automotive company (as shown in Table 2), each responsible for developing a business model for a data-driven innovation. We selected only cases where a data-driven service was provided to external B2B customers. To ensure the confidentiality of the company, interviewees, and use cases, all names and specific information were anonymized. In the beginning, we introduced the framework of roles and classes of exchanged values. The interviewees were then asked to apply the framework in the context of their use cases, in terms of involved actors as well as exchanged values. The outcome of each interview was a visual network-based representation of the business model. Further, we asked how understandable, useful and comprehensive the framework is and if some roles were missing or unnecessary.

---

[2] The full list of identified articles was omitted in this paper due to space restrictions and can be provided by the authors upon request.

Table 2: Overview of interviews

	Position	Use Case	Duration
A	Product Manager	Service in the field of autonomous driving	45 min
B	Product Manager	Providing end-customer insights as a service	30 min
C	Project Manager	Fleet monitoring service for electric vehicles	45 min

4   Results

In the following, we present our framework, as shown in Figure 1, by introducing eight roles and two attributes that can be assigned to an actor and follow with three classes of values that are exchanged between actors. Note that we describe here only the final framework after performing the initial evaluation described later in Section 5 due to space limitations.
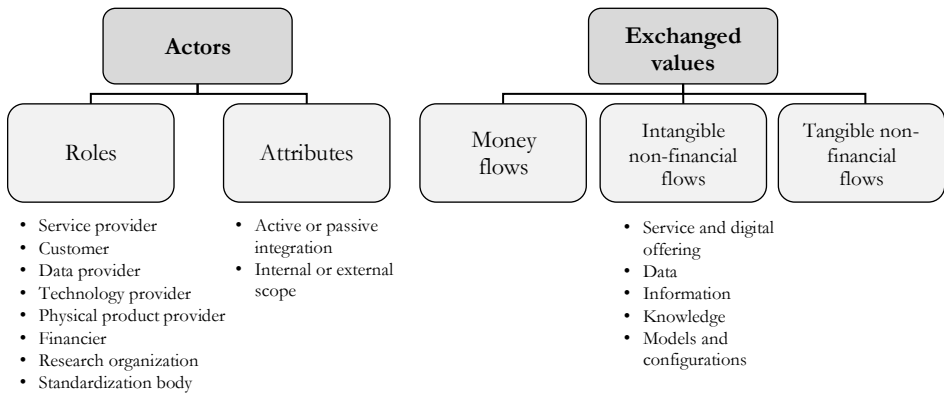
Figure 1: Classification of actors and exchanged values in data-driven business models.

4.1   Roles of Actors in Data-Driven Business Models

We have identified eight different roles that can be assigned to an actor in a DDBM:

- A **service provider** is an actor who utilizes data as a resource to create or co-create value for other actors (Immonen et al., 2014; Kaiser et al., 2019; Schüritz et al., 2019), for instance, by adding data-driven services to a physical product (Terrenghi et al., 2018).

F. Leski, M. Fruhwirth & V. Pammer-Schindler:
Who Else do You Need for a Data-Driven Business Model? Exploring Roles and Exchanged Values

371

- A **customer** is a recipient of the offering who also has a need. The customer initiates the value generation process triggered by his needs (Terrenghi et al., 2018), can actively participate in the value (co-)creation (Cova and Salle, 2008), or act just as a passive receiver. A customer creates real value via utilizing the potential value of the data-driven service offered by the service provider (Schüritz et al., 2019).

- A **data provider** is an actor who collects and aggregates data from public or private sources, performs the necessary preprocessing steps, and provides it to other actors who request data (Curry, 2016). This includes for instance collecting data about the conditions of a physical object in a cyber-physical system (Terrenghi et al., 2018).

- A **technology provider** is an actor who provides the necessary technical infrastructure, platforms, and tools to the business model owner, such as data management solutions or cloud technology (e.g., Curry, 2016; Immonen et al., 2014)

- A **physical product provider** is an actor who manufactures and sells a physical core product equipped with data-collecting devices, such as sensors, that should be enriched with a data-driven service (Kaiser et al., 2019; Papert and Pflaum, 2017; Terrenghi et al., 2018).

- A **financier** is a provider of financial resources for the business model innovation, such as a pre-financing investor, an incubator, or a venture capitalist (Curry, 2016; Papert and Pflaum, 2017).

- A **research organization**, such as a university, a research partner, or an internal research and development department, is an actor that engages with other actors in the business model to support the value generation process (Kindström et al., 2015; Schymanietz and Jonas, 2020).

- A **standardization body** is responsible for introducing common standards and controlling the economy legally addressing topics such as transparency or data privacy in the ecosystem (Curry, 2016; Terrenghi et al., 2018).

4.2 Actor Attributes

We further found two attributes that can be assigned to each actor describing their interaction in the business model: the level of integration into the business model (i.e., active or passive) and the scope (i.e., internal or external).

The **integration** of an actor into the business model can be either active or passive. An actively integrated actor can benefit from working with other actors in the business model (Zolnowski et al., 2016). Turetken et al. (2019) distinguish between core partners that are actively engaged in the value creation process; and enriching partners. The focal organization is overseeing the business model and takes an active role. Also, a customer can play an active role in the value co-creation process of a data-driven service (Schüritz et al., 2019) or stay passive by just receiving an offering.

The **scope** describes the internal or external relation relative to the focal business model that is currently analyzed. A service ecosystem consists of internal and external roles (Sklyar et al., 2019). All actors that are clustered within the same organizational unit of the focal business model are considered as internal. The business model owner cannot run the business by himself alone and therefore is usually supported by external and internal actors (Schymanietz and Jonas, 2020). For instance, a DDBM can rely on internal data sources, external data sources, or both (Hartmann et al., 2016), thus involves also external data providers.

4.3   Exchanged Values between Actors

Actors are exchanging values in a DDBM, that we clustered into three classes: money flows, intangible non-financial flows, and tangible non-financial flows.

Money flow summarizes all exchanged values of financial nature. Money serves as an enabler for negotiation and trading activities between economic actors (Allee, 2008). Flows of money can occur in different forms, denoted as revenue models, such as subscription fees or a pay-per-use model (Terrenghi et al., 2018). The choice of one model is influenced by several factors, such as capabilities and the characteristic of the service (Enders et al., 2019).

The class of **intangible non-financial flow** summarizes all exchanged values between actors that are non-monetary and intangible meaning that it »*cannot be seen, felt, tasted or touched*« (Chowdhury and Åkesson, 2011, p. 4). Such values are denoted in general as **services** (e.g., Immonen et al., 2014; Täuscher and Laudien, 2017) and **digital offerings** (Sklyar et al., 2019; Täuscher and Laudien, 2017). On a more granular level, flows can be divided into **data** (e.g., Engelbrecht et al., 2016; Terrenghi et al., 2018), **information** (e.g., Curry, 2016; Schüritz et al., 2019),

F. Leski, M. Fruhwirth & V. Pammer-Schindler:
Who Else do You Need for a Data-Driven Business Model? Exploring Roles and Exchanged Values

373

knowledge (e.g., Brownlow et al., 2015; Schüritz et al., 2019), and **models** or configuration of models (Hirt and Kühl, 2018).

The category **tangible non-financial flow** summarizes all exchanged values between actors that are non-monetary and tangible, such as physical products, raw materials, or other physical resources (e.g., Allee, 2008; Täuscher and Laudien, 2017). DDBMs can also rely on hardware as a key resource, such as measurement instruments, data transmission devices, or data-generating products.

5   Evaluation and Discussion of Results in three Cases

We evaluated the initial framework from the literature synthesis in three cases, collected via interviews with managers from one automotive company. We structured our insights along with the findings from the cases and the interviewees' feedback regarding the completeness, discriminability, and understandability of the individual elements as well as the usefulness of the overall framework.

The framework was sufficient to describe and analyze the cases: all roles, attributes, and values of the framework appeared in the cases and no additional ones that were not included in the framework emerged through the data collection process. When analyzing the content of our cases, we found that in two of three cases, one actor in the business model fulfilled both roles of the customer as well as a data provider, because the business model was relying on data provided by the customer. Further, the business models relied on partners with the role of a technology provider (e.g., providing a platform for data analysis), as well as a physical product provider (e.g., providing measurement hardware for data acquisition as well as devices for storage and transmission). Also, the attribute scope has been helpful in distinguishing between roles that are fulfilled internally and by external partners in the business model. Regarding the completeness of the framework, one interviewee mentioned that it is »*quite comprehensive, I can't think of an actor that is missing*« (Interviewee A). Interviewee C mentioned that he missed the role of the user, as in B2B context, the user and buying customer are often separate actors within one organization.

Regarding the definition of roles, we found that the initial roles of »data supplier« and »data collector« were difficult to distinguish, as one interviewee mentioned: »*the differentiation between the collector and supplier is diffuse for me, it might make more sense for other use cases*« (Interviewee A). Therefore, we decided to merge both under a new role

»data provider«. Similarly, distinguishing the initial roles of »resource integrator«, »service provider« and »technology provider« was challenging. We decided to drop the role of »resource integrator« and refined the definition of a technology provider, in the sense that a technology provider only offers the resource without supporting the application of it compared to a service provider. Further, the interviewees found it beneficial to have a more nuanced level of integration. We found that it is necessary to elaborate a more granular differentiation of exchanged intangible values, for instance what distinct type of exchanged data in a DDBM is essential.

Regarding the usefulness of the framework, it could help to understand complex environments by illustrating all involved actors as a whole. Interviewee B for instance, mentioned that a visual representation with this framework could be beneficial in complex and multi-layered business models and »*where there are no flows on both sides, it is critical because only one actor can benefit from the setting*« (Interviewee B).

6 Conclusion

In this research, we developed a framework with roles and attributes that can be assigned to actors and classes of values exchanged between actors in DDBMs. We contribute to the body of knowledge by shedding first lights on the partner perspective in DDBMs, which was overlooked by previous research. We extend the existing research of Hunke et al. (2020) on the orchestration of key activities in data-driven services, by a nuanced reflection of the partner perspective. Further, our framework extends a network-based representation for DDBMs from Fruhwirth et al. (2021) by providing a template for identifying and then visualizing relevant actors and value flows. Additionally, we extend existing research on flow-based business model representations (e.g., Gordijn and Akkermans, 2001) by introducing new attributes for actors as well as data as a central type of value flows.

Note as a limitation, that we found few publications that had a specific focus on data due to the topic's novelty. Further, we might have missed some literature due to the difference in the denomination of the concept of DDBMs that our search strings might not cover. Another limitation is the sparse evaluation of the results only via three experts of one company in one particular industry. To further improve our framework, it should be applied to an additional number of cases from multiple industries to make our results more generalizable. Also, a detailed investigation of

F. Leski, M. Fruhwirth & V. Pammer-Schindler:
Who Else do You Need for a Data-Driven Business Model? Exploring Roles and Exchanged Values

375

data-related value exchanges and the roles of service and technology providers seems like a fruitful direction for further research. Further, the framework could be instantiated in a tool with visual representations of the actors and values to design and communicate DDBMs.

References

Allee, V. (2008) 'Value network analysis and value conversion of tangible and intangible assets', Journal of Intellectual Capital, vol. 9, no. 1, pp. 5–24.

Azkan, C., Iggena, L., Gür, I., Möller, F. and Otto, B. (2020) 'A Taxonomy for Data-Driven Services in Manufacturing Industries', PACIS 2020 Proceedings.

Breitfuß, G., Fruhwirth, M., Pammer-Schindler, V., Stern, H. and Dennerlein, S. (2019) 'The Data-Driven Business Value Matrix: A Classification Scheme for Data-Driven Business Models', 32nd Bled eConference Humanizing Technology for a Sustainable Society, pp. 803–820.

Brownlow, J., Zaki, M., Neely, A. and Urmetzer, F. (2015) Data and Analytics - Data-Driven Business Models: A Blueprint for Innovation [Online], Cambridge Service Alliance. Available at http://cambridgeservicealliance.eng.cam.ac.uk/resources/Downloads/Monthly%20Papers/2015FebruaryPaperTheDDBMInnovationBlueprint.pdf (Accessed 19 July 2017).

Chowdhury, S. and Åkesson, M. (2011) 'A proposed conceptual framework for identifying the logic of digital services', PACIS 2011 Proceedings.

Cova, B. and Salle, R. (2008) 'Marketing solutions in accordance with the S-D logic: Co-creating value with customer network actors', Industrial Marketing Management, vol. 37, no. 3, pp. 270–277.

Curry, E. (2016) 'The Big Data Value Chain: Definitions, Concepts and Theoretical Approaches', in Cavanillas, J. M., Curry, E. and Wahlster, W. (eds) New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe, Switzerland, SpringerOpen, pp. 29–38.

Dellermann, D., Fliaster, A. and Kolloch, M. (2017) 'Innovation risk in digital business models: the German energy sector', Journal of Business Strategy, vol. 38, no. 5, pp. 35–43.

Enders, T., Schüritz, R. and Frey, W. (2019) 'Capturing Value from Data: Exploring Factors Influencing Revenue Model Design for Data-Driven Services', WI 2019 Proceedings, pp. 1738–1752.

Engelbrecht, A., Gerlach, J. and Widjaja, T. (2016) 'Understanding the anatomy of data driven business models - towards an empirical taxonomy', ECIS 2016 Proceedings.

Fruhwirth, M., Breitfuß, G. and Pammer-Schindler, V. (2020) 'The Data Product Canvas: A Visual Collaborative Tool for Designing Data-Driven Business Models', 33rd Bled eConference Enabling Technology for a Sustainable Society. Online, June 28-29 2020, pp. 515–528.

Fruhwirth, M., Pammer-Schindler, V. and Thalmann, S. (2021) 'A Network-based Tool for Identifying Knowledge Risks in Data-Driven Business Models', Proceedings of the 54th Annual Hawaii International Conference on System Sciences, pp. 5218–5227.

Fruhwirth, M., Ropposch, C. and Pammer-Schindler, V. (2020) 'Supporting Data-Driven Business Model Innovations: A structured literature review on tools and methods', Journal of Business Models, vol. 8, no. 1, 7-25.

Gordijn, J. and Akkermans, H. (2001) 'Designing and evaluating e-business models', IEEE Intelligent Systems, vol. 16, no. 4, pp. 11–17.

Guggenberger, M. T., Möller, F., Boualouch, K. and Otto, B. (2020) 'Towards a Unifying Understanding of Digital Business Models', PACIS 2020 Proceedings.

Hartmann, P. M., Zaki, M., Feldmann, N. and Neely, A. (2016) 'Capturing value from big data – a taxonomy of data-driven business models used by start-up firms', International Journal of Operations & Production Management, vol. 36, no. 10, pp. 1382–1406.

Hirt, R. and Kühl, N. (2018) 'Cognition in the Era of Smart Service Systems: Inter-organizational Analytics through Meta and Transfer Learning', ICIS 2018 Proceedings.

Hunke, F., Seebacher, S., Schüritz, R. and Illi, A. (2017) 'Towards a Process Model for Data-Driven Business Model Innovation', 19th IEEE Conference on Business Informatics, pp. 150–157.

Hunke, F., Seebacher, S. and Thomsen, H. (2020) 'Please Tell Me What to Do – Towards a Guided Orchestration of Key Activities in Data-Rich Service Systems', DESRIST 2020 Proceedings.

Immonen, A., Palviainen, M. and Ovaska, E. (2014) 'Requirements of an Open Data Based Business Ecosystem', IEEE Access, vol. 2, pp. 88–103.

Jacobides, M. G., Cennamo, C. and Gawer, A. (2018) 'Towards a theory of ecosystems', Strategic Management Journal, vol. 39, no. 8, pp. 2255–2276.

Kaiser, C., Stocker, A. and Fellmann, M. (2019) 'Understanding Data-driven Service Ecosystems in the Automotive Domain', AMCIS 2019 Proceedings.

Kindström, D., Kowalkowski, C. and Alejandro, T. B. (2015) 'Adding services to product-based portfolios', Journal of Service Management, vol. 26, no. 3, pp. 372–393.

Kühne, B. and Böhmann, T. (2019) 'Data-Driven Business Models: Building the Bridge Between Data and Value', ECIS 2019 Proceedings.

Mayring, P. (2015) Qualitative Inhaltsanalyse: Grundlagen und Techniken, 12th edn, Weinheim, Beltz.

Papert, M. and Pflaum, A. (2017) 'Development of an Ecosystem Model for the Realization of Internet of Things (IoT) Services in Supply Chain Management', Electronic Markets, vol. 27, no. 2, pp. 175–189.

Schüritz, R., Farrell, K., Wixom, B. H. and Satzger, G. (2019) 'Value Co-Creation in Data-Driven Services: Towards a Deeper Understanding of the Joint Sphere', ICIS 2019 Proceedings.

Schüritz, R. and Satzger, G. (2016) 'Patterns of Data-Infused Business Model Innovation', IEEE 18th Conference on Business Informatics, pp. 133–142.

Schüritz, R., Seebacher, S., Satzger, G. and Schwarz, L. (2017) 'Datatization as the Next Frontier of Servitization: Understanding the Challenges for Transforming Organizations', ICIS 2017 Proceedings.

Schymanietz, M. and Jonas, J. M. (2020) 'The Roles of Individual Actors in Data-Driven Service Innovation – A Dynamic Capabilities Perspective to Explore its Microfoundations', Proceedings of the 53rd Annual Hawaii International Conference on System Sciences (HICSS 2020), pp. 1135–1144.

Sklyar, A., Kowalkowski, C., Tronvoll, B. and Sörhammar, D. (2019) 'Organizing for digital servitization: A service ecosystem perspective', Journal of Business Research, vol. 104, pp. 450–460.

Täuscher, K. and Abdelkafi, N. (2017) 'Visual tools for business model innovation: Recommendations from a cognitive perspective', Creativity and Innovation Management, vol. 26, no. 2, pp. 160–174.

Täuscher, K. and Laudien, S. M. (2017) 'Uncovering the Nature of Platform-based Business Models: An Empirical Taxonomy', Proceedings of the 50th Hawaii International Conference on System Sciences.

Terrenghi, N., Schwarz, J. and Legner, C. (2018) 'Towards Design Elements to Represent Business Models for Cyber Physical Systems', ECIS 2018 Proceedings.

Timmers, P. (1998) 'Business Models for Electronic Markets', Electronic Markets, vol. 8, no. 2, pp. 3–8.

Turetken, O., Grefen, P., Gilsing, R. and Adali, O. E. (2019) 'Service-Dominant Business Model Design for Digital Innovation in Smart Mobility', Business & Information Systems Engineering, vol. 61, no. 1, pp. 9–29.

Vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R. and Cleven, A. (2009) 'Reconstructing the Giant: On the Importance of Rigour in documenting the Literature Search Process', ECIS 2009 Proceedings.

Webster, J. and Watson, R. T. (2002) 'Analyzing the past to prepare for the future: Writing a literature review', Management Information Systems Quarterly, vol. 26, no. 2, pp. 13–23.

F. Leski, M. Fruhwirth & V. Pammer-Schindler:
Who Else do You Need for a Data-Driven Business Model? Exploring Roles and Exchanged Values

377

Wixom, B. H. and Ross, J. W. (2017) 'How to Monetize Your Data', MIT Sloan Management Review, vol. 58, no. 3, pp. 10–13.

Zolnowski, A., Towe, C. and Gudat, J. (2016) 'Business Model Transformation Patterns of Data-Driven Innovations', ECIS 2016 Proceedings.

Zott, C. and Amit, R. (2010) 'Business Model Design: An Activity System Perspective', Long Range Planning, vol. 43, 2-3, pp. 216–226

ICT-Driven Business Model Innovation in SMEs: The Role of Organizational Capabilities, Firm Size and Age

Mohammad-Ali Latifi, Harry Bouwman &
Shahrokh Nikou

Delft University of Technology, Faculty of Technology, Policy, and Management, South Holland, The Netherlands, e-mail: s.m.a.latifirostami@tudelft.nl

Abstract Research has shown that business model innovation (BMI) can create competitive advantages and enhance firm performance. However, many small and medium-sized enterprises (SMEs) fail to supreme their performance. BMI can create unexpected consequences for businesses and their ecosystem. Therefore, knowing how and under what circumstances BMI affects a firm's performance is a primary concern for managers/owners of SMEs. Using data from 460 European SMEs, this paper aims to examine three paths through which ICT-driven BMI can impact firm's performance. Introducing organisational capabilities as a mediator, this study has extended prior literature on BMI by showing that organisational capabilities are as strong as other existing mediators of revenue and efficiency growth regarding improving the firm's performance. The findings provide guidelines for practitioners to execute informed-decisions about the implementation of BMI based on their firm's strategies and the available capabilities while considering contingent factors of firm size and age.

1   Introduction

Performance improvement is at the heart of any firm, and according to scholars (Hartmann et al., 2013; Karimi and Walter, 2016) BMI positively contributes to firm performance. Some classic examples of innovation in BMs and their association with the firm's performance are Dell, Wal-Mart, Uber, and Southwest. All these companies developed a novel BM by either introducing or reorganising key components of their BM. The changes to core components or to the architecture of a firm's BM (Nair et al., 2013), in comparison to innovations in product, service, and process, have been associated with high risk and uncertainty (Chesbrough, 2010; Waldner et al., 2015). So, if not handled properly, a well-formulated BM may fail to lead to improved performance (Chesbrough, 2010; Knab and Rohrbeck, 2014). Christensen et al. (2016) revealed that more than 60% of BMI efforts did not deliver the expected performance. So BMI can have both positive and negative outcomes. Hence, knowing how and when to innovate a BM is a serious challenge for managers/owners of firms (Hartmann et al., 2013).

In this paper, we focus on '*how*' firms exploit or modify their BMs to improve their overall performance. Therefore, we have two objectives: (a) to develop and examine a conceptual framework that illustrates the complex mechanisms through which implementation of strategic BMI decisions related to a focus on efficiency or growth, as well as organizational capabilities, influences a firm's overall performance, and (b) to explore whether specific characteristics of the firm (i.e., size, age) have a different impact on performance.

We contribute to the BM literature in three ways. First, by examining the proposed model using empirical data, we contribute to both the practical knowledge and theoretical enrichment. Second, by considering the mediating effect of organizational capabilities, this research attempts to understand how managers can ensure that BMI provides more benefits to their firms in terms of performance. Third, by focusing specifically on SMEs, we contribute to the body of knowledge on BMI in relation to SMEs.

This paper is structured as follows. First, a systematic literature review on mediating factors between BMI and performance is presented. The research method is then described. Next, the research results are presented. The paper ends with some conclusions, a discussion of limitations and suggestions for future research.

2   Literature Review and Hypotheses Development

A business model, which uses as a tool to communicate and implement strategic choices, is seen as a realized expression of strategy and articulates how available resources can be used more effectively, how costs can be managed and reduced, and how new sources of revenues can be leveraged (Chesbrough, 2007). Although a significant number of companies have gained advantages from BMI, there are many more that have performed extremely poorly, failed to meet their objectives or even exited business.

To explore the causal mechanism, our in-depth analysis of 37 articles resulted in 12 distinct mediating factors through which the BMI indirectly influences a firm's overall performance. Analysing the factors, we found that some mediators mainly were related to generating revenue by increasing the firm's sales, by a combination of exploring new markets, new customers and new value propositions, and by new ways of service, product and price bundling; we, therefore, called them '*Revenue growth*.' Other mediators focus on efficiency – that is, by introducing new ways to minimise cost, increasing productivity, or reducing time to market – are referred to as '*Efficiency growth*.' The last two foci, were related but not identical to the design of efficiency- and novelty-oriented BMs as highlighted in BMI literature (Heikkila et al., 2018; Zott and Amit, 2008).

However, we identified some more generic mediators that do not belong to mediators related to revenue or efficiency growth groups, e.g., organisational learning, opportunity recognition, organizational culture; these mediators enable companies to increase their revenue and efficiency. We called this group of concepts '*Organisational capabilities*', which is vital to long-term performance of business, since a culture of openness and knowledge sharing reinforce a high level of cooperation within the firm and its associated network and contribute to a firm's readiness to change, and in particular to its ability to survive in the longer term, rather than merely achieving short-term growth (Latifi et al., 2021).

Scholars and practitioners agree that the BM is vital to, for example, the success of organisations, and especially to those that wish to grow (Teece, 2010), to gain competitive advantage, to enhance long-term performance (Bock et al., 2012) and as a new source of innovation (Zott, 2011). However, recent studies have produced inconclusive results when testing the strength of the relationship between BMI and firms' performance in different regions and industries. We, therefore, to examine whether BMI impacts the performance of SMEs in the European context, propose the following hypothesis:

H1: *If an SME engages in BMI, its overall performance will improve.*

Heikkilä et al. (2018) stress that BMI influences firm performance occurs when there is a strategic focus on efficiency. Their findings confirm the research by Zott and Amit (2007) on the impact of efficiency-centred BM design on a firm's overall performance. BMI can take ICT ventures to complete their transactions efficiently, by reducing transaction costs within the firm and with its outsiders (Ladib and Lakhal, 2015). According to Chesbrough (2007), BMI leverages performance not only by reducing production costs but also by utilising available resources more effectively. Gronum et al. (2016) and Wei et al. (2017) also found that BM designs that focus on efficiency enhance a firm's performance by reducing inventory costs – thus benefitting both customers and suppliers – and decreasing marketing, sales, and other communication expenditures. In light of this, we propose the next hypothesis:

H2: *Efficiency growth mediates the relation between BMI and an SME overall performance*

However, the focus can also be on implementing a growth strategy by attracting new customers and expanding the firm's markets (Heikkilä et al., 2018). Some scholars argue that BMI, through the creation of new value propositions (Teece, 2010; Wei et al., 2017) or opportunity recognition (Guo et al., 2017) can attract new customers by exploring a market niche not addressed by competitors (Zott and Amit, 2007). These could occur via new ways of market penetration or new ways of market development. Moreover, BMI by combining existing and new channels in a smart way can create new value (Ladib and Lakhal, 2015). Based on this review, we propose the following hypothesis:

H3: *Revenue growth mediates the relation between BMI and an SME's overall performance*

The capacity to innovate is seen as one of the key factors that improve business performance (Burns and Stalker, 1966). An organisation's culture –norms, values, and beliefs as expressed within the organisation – can boost behaviour that is ultimately related to business performance (Hult et al., 2004). A culture that supports the implementation of a strategic attempt and encouraged by the enthusiastic support of all employees is not easy to imitate and can lead to a sustainable competitive advantage (Anning-Dorson, 2017). A large number of studies found a significant relationship between firm innovativeness and performance in different types of organisations (Rubera and Kirca, 2012). Hult et al. (2004) concluded that innovativeness appeared to be a key mediator in their empirical research. The role of opportunity-sensing and seeking behaviour in BMI has also been emphasised in several studies (Chesbrough, 2010). Several studies investigated the direct effects of corporate entrepreneurship on performance (George and Bock, 2011). However, BMI as a mediator of the relationship between corporate entrepreneurship and a firm's performance was also considered by Karimi and Walter (2016). So the concepts might have a direct or a mediating effect, here we consider entrepreneurship as belonging to this group of meditators. We, therefore, propose to consider organisational capability, as discussed for innovativeness, opportunity recognition, and culture, as an alternative group of mediating factor:

H4: *Organisational capabilities mediate the relation between BMI and an SME's overall performance.*

A summary of the identified mediating factors that indirectly affect the relationship between BMI and a firm's overall performance is presented in the Appendix. Figure 1 illustrates the proposed conceptual research model, which is based on the literature review.
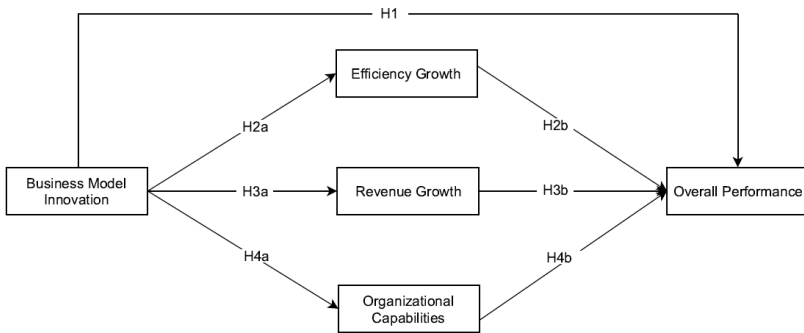
Figure 1: Research conceptual model: BMI mechanism to boost firms' performance

3   Research Method

An extensive literature review in the domains of entrepreneurship, innovation management, and BMs was conducted to assemble a comprehensive list of reliable measures. The research constructs were measured by using existing items in the literature. Since firm size and firm age could impact the relationship between BMI and firm's overall performance (Hartmann et al., 2013; Zott and Amit, 2007), it is appropriate to account for their impact on the path relationships identified in Fig.1.

3.1   Survey administration, sample and data collection

This study's population are European SMEs in any industry engaged in business model innovation in the previous 24 months and used ICT to enable their product and service offering. The sample was based on Dun and Bradstreet's database. Firms were randomly selected from this sample frame an established quotas for micro-enterprises, small and medium-sized enterprises resulting in a final distribution of 36%, 32%, and 31%, respectively. Data was collected by a professional research agency that uses native speakers and computer-aided telephone interviewing. The final dataset contains 460 SMEs in 17 different industries in 13 European countries engaged in BMI. The questionnaire was based on the previously mentioned scales and was pretested in each of the 13 countries by reading aloud to managers and academics to improve the clarity of the questions and to prevent any potential ambiguous expressions.

4   Data analysis

All constructs fulfill the requirement for Cronbach's alpha, i.e., 0.70 or higher (Hair et al.,2014) and for composite reliability (CR) (0.70 or higher) (Table 1). Convergent validity is represented by average variance extracted (AVE), which is recommended to be at least 0.50 (Hair et al., 2011). (Table 2). Discriminant validity guarantees the uniqueness of a measuring construct and indicates that the phenomenon of interest is not captured in other measures (latent variables) within the research model (Hair et al., 2010). For assessing discriminant validity, an alternative criterion is HTMT. An HTMT value close to 1 indicates a lack of discriminant validity. The HTMT values are lower than 0.85. We, therefore, conclude that discriminant validity is not an issue.

5   Results

To test the hypotheses, we employed structural equation modelling (SEM) using SmartPLS v.3 software. The mediation analyses with regard to three variables, namely efficiency growth, revenue growth, and the organisational capability, were also computed. Additionally, a multi-group analysis was conducted to evaluate the role of control variables (i.e., firm size and age).

In the path model analysis, the firms' overall performance is explained by a variance of 30%, and the three mediators – namely efficiency growth, organisational capabilities, and revenue growth – are explained by a variance of 24%, 27%, and 37%, respectively, in the model. Consistent with our expectations, the direct path

Table 1: Descriptive statistics, convergent validity, consistency and reliability of items

Constructs	Items	Factor Loadings	t-statistics	Cronbach's alpha	CR	AVE
Overall performance	The sales growth of the enterprise	0.82	48.40	0.91	0.93	0.65
	The profit growth of the enterprise	0.83	53.41			
	Market share	0.75	26.16			
	Speed to market	0.71	26.37			
	Penetration rate	0.77	33.73			
	Market value	0.79	42.02			
	Net income	0.81	37.51			
	Return on Investment (ROI)	0.79	34.67			
Business model innovation	Introduced new products as a new value proposition	0.72	16.52	0.73	0.81	0.54
	Introduced new services as a new value proposition	0.74	21.80			
	Started to collaborate with new business partners	0.66	15.25			
	Shared new responsibilities with business partners	0.67	13.82			
	Created new revenue streams	0.71	12.61			
	Introduced a new pricing mechanism	0.69	12.61			
	Focused on a completely new market segment	0.76	21.23			
Revenue growth	Introduced new ways to transact with customers	0.70	25.47	0.75	0.83	0.60
	Introduced new ways of organising relations with customers	0.73	30.14			
	Advertising products and services in a new way	0.71	26.26			
	Scale up your business	0.71	27.98			
	Focus your product offering	0.70	18.40			
Efficiency growth	Introduced new ways to reduce fixed costs	0.71	13.71	0.71	0.82	0.54
	Introduced new ways to reduce variable costs	0.72	14.85			
	Business processes standardisation	0.81	29.99			
	Business processes integration	0.82	34.42			
Organisational capabilities	Managers encourage employees to think outside the box	0.76	21.41	0.87	0.90	0.53
	Our corporate culture is focused on constant innovation	0.78	33.96			
	Our enterprise shows perseverance in turning ideas into reality	0.74	28.08			
	Our enterprise ability to identify new opportunities	0.76	34.13			
	Our enterprise aims to create multiple innovations annually	0.78	37.25			
	Our enterprise introduces innovations that are completely new to the market	0.75	36.02			
	Creating more than one innovation at the same time is common practice in our enterprise	0.73	31.11			

Note: CR = Composite reliability; AVE = Average variance extracted

between BMI and the firm's overall performance is significant (in the absence of mediators); thus, H1 is supported by the model ($\beta = 0.41$, $t = 10.52$, $p < 0.001$). However, this direct path between BMI and the firm's overall performance is not significant when the three mediators are included in the analysis. As shown in Figure 2, all six paths from BMI to efficiency growth, revenue growth, and organisational capabilities, and from these mediators to the firm's overall performance, are statistically significant ($p < 0.001$).
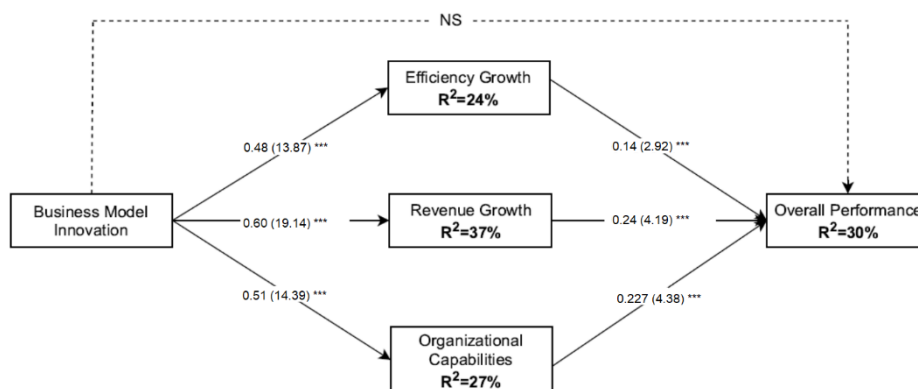
Figure 2: Structural model results (Significance levels: *** p < 0.001, and NS means not significant)

Mediation analysis showed a significant direct relationship between BMI and the firm's overall performance. Therefore H1 was supported confirming that the independent variable (BMI) is a significant predictor of the dependent variable (firm's overall performance). Satisfying this condition provides the ground to test the mediation relationship between BMI and the firm's overall performance. Based on the SEM results, when the mediators are included in the analysis, the direct path between BMI and overall firm performance is not significant (see Figure 2). Moreover, as we hypothesized, the mediation test results show that the path between BMI and the firm's overall performance is fully mediated by three variables (i.e., efficiency growth, revenue growth, and organisational capabilities) in our proposed model. The individual effects of each mediator can be seen in Table 2; thus hypotheses H2, H3, and H4 are supported by the model.

When considering the size of firms, in our multi-group analysis, we found a significant direct relationship between BMI and overall performance for different sizes. Table 3 shows that the mediation effect of efficiency growth is significant only for micro-sized firms, and revenue growth mediates the relationship between BMI and the overall performance of both small and medium-sized firms. However, firm

Table 2: The mediation results between BMI and a firm's overall performance

Mediation paths	β	t-Statistics	p-values	Mediation
BMI → Efficiency growth → Overall performance	0.07	2.91	0.005	Full
BMI → Revenue growth → Overall performance	0.14	4.15	0.001	Full
BMI → Organisational capabilities → Overall performance	0.12	4.29	0.001	Full

size has no significant effect on the mediation of organisational capabilities; therefore, organisational capabilities mediate between BMI and performance in firms of all sizes.

Table 3: The effect of firm size on mediation relationships

Paths	All Firms	Contingency factor: Firm Size		
		Micro	Small	Medium
BMI → Efficiency growth → Overall performance	0.07 (2.91) ***	0.09 (2.47) **	NS	NS
BMI → Revenue growth → Overall performance	0.14 (4.15) ***	0.18 (2.71) **	0.14 (2.70) **	0.17 (2.67) ***
BMI → Organisational capabilities → Overall performance	0.12 (4.29) ***	NS	0.23 (4.77) ***	0.16 (2.61) ***

NS = Not Significant

Considering firm age as a moderator, none of the three factors mediate the relationship between BMI and the firm's overall performance in newly-established firms. While efficiency and revenue growth mediate the path between BMI and firm's overall performance for young and well-established firms, the organizational capabilities mediate this relationship solely in well-established firms (Table 4).

Table 4: The effect of firm age on mediation relationship

Paths	All Firms	Contingency factor: Firm Age		
		Newly-established	Young	Well-established
BMI → Efficiency growth → performance	0.07 (2.91) ***	NS	.12 (2.26) **	.07 (2.28) **
BMI → Revenue growth → Performance	0.14 (4.15) ***	NS	NS	.15(3.76) ***
BMI → Organisational capabilities → Performance	0.12 (4.29) ***	NS	0.12 (2.24) **	.13(3.79) ***

NS: Not Significant

6   Conclusion

This study proposes a model that would enable researchers and practitioners to understand causal mechanisms through which business model innovation (BMI) influences firm performance, specifically when SMEs which use digital technologies are concerned. Our findings contribute to the literature and confirm that 'efficiency growth,' 'revenue growth,' and 'organisational capabilities' are relevant mediators for SMEs that engage in BMI to increase firm's overall performance. We contribute to the literature by considering organisational capabilities as a mediator between BMI and firms' overall performance. The research findings also enhance our understanding by demonstrating the importance of firm size and age as moderators between BMI and SME performance. However, we did not find any significant difference in the influence of the contingency factors of firm size and firm age on the direct relationship between BMI and firms' performance; the results showed that well-established firms achieve better performance through developing organisational capabilities by doing BMI. Moreover, efficiency improvement is not the primary goal of newly-established firms (start-ups) that implement BMI to improve performance, although it is for young (scale-up) and well-established firms.

Although various organisational capabilities exist, we investigated only the entrepreneurial orientation, innovativeness, and organisational culture. However, other capabilities that might mediate the relation between BMI and performance – for example, employees' training and leadership style – are worth further investigation. Furthermore, the focus of the present study was on exploring mediation factors, and only a limited number of contingency factors (i.e., firm size and firm age) were taken into consideration. In particular, we did not take into account the industry characteristics or the BMI implementation skills within firms, for instance, employees' knowledge and skills, management support and the use of

BMI tooling (Latifi and Bouwman, 2018). This could be an interesting avenue for further research.

Our results have implications both for the academic literature on BMI and for practitioner discussions. First, by considering mediating effects, the model elucidates how managers can ensure that BMI brings more benefits to their firms in terms of performance. Second, by examining the proposed model using empirical data from 17 different industries in 13 European countries, we could add practical knowledge along with theoretical enrichment. Third, by focusing specifically on SMEs, we contribute to SMEs' knowledge of BMI. Fourth, by taking into account the influence of firm size and age on the relationship between BMI and the performance of firms, we provide insight that the owners/managers of firms need to carefully assess their specific situation in order to take appropriate measures to improve the effect of BMI on performance and to choose a focus on growth or efficiency to exploit all benefits of BMI efforts fully. Moreover, managers need to be aware of the organizational capabilities related to BM Innovation. This also clarifies that generic advice of consultants or training programs on BMI needs to take differences in foci as well as capabilities, age, and size into account. More tailored programs are advisable.

This study also has some limitations that should be considered when interpreting the findings. First of all, although cross-sectional data are used extensively in business and management research, such data represent a single point in time and hardly allows the cause and effect or the impact of changes over time to be determined. Second, although the respondents – mainly the firms' top managers/owners – possessed a high degree of relevant knowledge, all of the measures were self-reported using a self-assessed scale, which may represent a potential source of common method bias. Future research should collect objective measurements to eliminate common method bias.

While the business world is constantly changing in terms of technology, regulations and customer's needs, we believe that these results advance BMI research by opening the black box of the causal relationship between BMI and a firm's overall performance to better understand the BMI phenomenon, how it works and how we can gain the greatest benefit from it. We hope that our work leads to improved managerial practices and helps future research to probe more deeply into these constructs in small and medium-sized firms.

References

Anning-Dorson, T., 2017. Moderation-mediation effect of market demand and organisation culture on innovation and performance relationship, Marketing Intelligence & Planning, 35, 222–242.

Bock, A.J., Opsahl, T., George, G., Gann, D.M., 2012. The Effects of Culture and Structure on Strategic Flexibility during Business Model Innovation. Journal of Management Studies, 49, 279–305.

Burns, T. and Stalker, G.M. (1966) The Management of Innovation. London: Tavistock.

Chesbrough, H., 2010. Business model innovation: Opportunities and barriers. Long Range Planning, 43, 354–

Chesbrough, H., 2007. Business model innovation: it's not just about technology anymore. Strategy & Leadership, 35, 12–17. doi:10.1108/10878570710833714

Christensen, C.M., Bartman, T., Van Bever, D., 2016. The hard truth about business model innovation. MIT Sloan Management Review, 58.

Foss, N.J., Saebi, T., 2017. Fifteen Years of Research on Business Model Innovation: How Far Have We Come, and Where Should We Go? Journal of Management, 43 (1), 220-227.

Gronum, S., Steen, J., Verreynne, M. L., 2016. Business model design and innovation: Unlocking the performance benefits of innovation. Australian Journal of Management, 41, 585–605.

Guo, H., Tang, J., Su, Z., Katz, J.A., 2017. Opportunity recognition and SME performance: the mediating effect of business model innovation. R&D Management 47, 431–442. doi:10.1111/radm.12219

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2010. Multivariate Data Analysis, Prentice-Hall.

Hair, J.F., Ringle, C.M., Sarstedt, M., 2011. PLS-SEM: indeed, a silver bullet. Journal of Marketing Theory and Practice, 19 (2), 139-151.

Hair, J.F., Sarstedt, M., Kuppelwieser, V., Hopkins, L., 2014. Partial Least Squares Structural Equation Modeling (PLS-SEM): An Emerging Tool for Business Research. European Business Review 26, 106-121.

Hartmann, M., Oriani, R., Bateman, H., 2013. The Performance Effect of Business Model Innovation: An Empirical Analysis of Pension Funds. 35th DRUID Celebration Conference, Barcelona, Spain.

Heikkilä, M., Bouwman, H., Heikkilä, J., 2018. From strategic goals to business model innovation paths: an exploratory study. Journal of Small Business and Enterprise Development, 25, 107–128.

Hu, B. (2014). Linking business models with technological innovation performance through organizational learning. European Management Journal, 32, 587–595. doi:10.1016/j.emj.2013.10.009

Hult, G.T.M., Hurley, R., Knight, G.A., 2004. Innovativeness: Its antecedents and impact on business performance. Industrial Marketing Management, 33, 429–438. doi:10.1016/j.indmarman.2003.08.015

Karimi, J., Walter, Z., 2016. Corporate Entrepreneurship, Disruptive Business Model Innovation Adoption, and Its Performance: The Case of the Newspaper Industry. Long Range Planning 49, 342–360.

Knab, S., Rohrbeck, R., 2014. Why intended business model innovation fails to deliver, Proceedings of the R&D Management Conference, Stuttgart, Germany.

Kumar Pati, R., Nandakumar, M.K., Ghobadian, A., Ireland, R.D., Regan, N.O., 2018. Business model design performance relationship under external and internal contingencies: Evidence from SMEs in an emerging economy. Long Range Planning, 51 (5). doi:10.1016/j.lrp.2018.01.001

Ladib, N., Lakhal, L., 2015. Alignment between business model and business strategy and contribution to the performance. Journal of High Technology Management Research, 26, 168–176.

Latifi, M-A, Nikou, S., Bouwman, H., 2021. Business model innovation and firm performance: Exploring causal mechanisms in SMEs, Technovation, Volume 107. 10.1016/j.technovation.2021.102274.

Latifi, M-A., Bouwman, H., 2018. Business Model Innovation and Firm Performance: The Role of Mediation and Moderation Factors. 31st Bled eConference, June 17 – 20, Bled, Slovenia

Liu, Y., Han, Z., 2013. The Empirical Study of Relationship between Business Model and Enterprise Performance. Advanced Materials Research, 850–851, 1078–1081.

Majumdar, S.K., 1997. The Impact of Size and Age on Firm-Level Performance. Review of Industrial Organization, 12, 231–241

Marshall, A., 1920. Principles of Economics, 8th Edition. London: Macmillan.

Nair, S., Paulose, H., Palacios, M., Tafur, J., 2013. Service orientation: effectuating business model innovation. The Service Industries Journal 33, 958–975. doi:10.1080/02642069.2013.746670

Pucci, T., Nosi, C., Zanni, L., 2017. Firm capabilities, business model design and performance of SMEs. Journal of Small Business and Enterprise Development, 24, 222–241. doi:10.1108/JSBED-09-2016-0138

Rubera, G., Kirca, A.H., 2012. Firm Innovativeness and Its Performance Outcomes: A Meta-Analytic Review and Theoretical Integration. Journal of Marketing 76, 130–147. doi:10.1509/jm.10.0494

Teece, D.J., 2010. Business models, business strategy and innovation. Long Range Planning 43, 172–194.

Velu, C., 2015. Business model innovation and third-party alliance on the survival of new firms. Technovation 35, 1–11. doi:10.1016/j.technovation.2014.09.007

Waldner, F., Poetz, M.K., Grimpe, C., Eurich, M., 2015. Industry Structure and Business Model Innovation Antecedents and consequences of business model innovation: at DRUID15 conference, Rome, Italy.

Wei, Z., Song, X., Wang, D., 2017. Manufacturing flexibility, business model design, and firm performance. International Journal of Production Economics 193, 87–97. doi:10.1016/j.ijpe.2017.07.004

Zott, C., Amit, R., 2008. The fit between product market strategy and business model: implications for firm performance. Strategic Management Journal 29(1), 1–26. doi:10.1002/smj.642

Zott, C., Amit, R., 2007. Business model design and the performance of entrepreneurial firms. Organization Science 18(2), 181-199. doi:10.1287/orsc.1060.0232

DECEPTIVE DESIGN: COOKIE CONSENT AND MANIPULATIVE PATTERNS

THOMAS MEJTOFT, ERIK FRÄNGSMYR,
ULRIK SÖDERSTRÖM & OLE NORBERG

Umeå University, Digital Media Lab, Sweden; e-mail: thomas.mejtoft@umu.se,
erik.frangsmyr@gmail.com, ulrik.soderstrom@umu.se, ole.norberg@umu.se

Abstract As a larger proportion of our lives moves onto the web, so does important and valuable information. This has led to an increase in different kinds of manipulative patterns (dark patterns) in web design with the sole purpose of being deceptive and tricking users. This paper discusses the comprehensive suite of deceptive design patterns on Internet services where the users are expected to comply with the use of cookies. This was done by analyzing 50 different home cooking recipe websites, regarding their appliance to GDPR and how they use different dark patterns in their design. Even though legislation tries to move the choices from the website to the user, it is clear that by using deceptive design patterns it is possible to "bypass" the legislation and trick the user into making a favorable choice for the owners behind the website. The results show that out of the websites that were GDPR approved, a majority still use two types of deceptive design patterns - misdirection and sneak into basket.

1 Introduction

Since the very first web page, there has been constant changes to the design to create usable designs that increases the user experience. New guidelines have gradually evolved to enable great designs for various digital devices (Nielsen, 2018; Shneiderman, 2009). This constant perfection has made the vast majority of web sites intuitive and easy to use for most users. However, as the field of user experience on the web has gradually evolved during the last 25 years, so have underhanded tactics colloquially known as dark patterns or deceptive designs (Brignull, 2011). These features of design are just as carefully crafted but with another purpose than to lead the user in the right direction. Hence, a dark pattern is "a user interface carefully crafted to trick users into doing things they might not otherwise do, such as buying insurance with their purchase or signing up for recurring bills" (Brignull, 2013). Unlike concepts like e.g. digital nudging, which is about creating solutions that help the user to make the choices in their best interest by altering the choice environment (Thaler & Sunstein, 2008; Mejtoft et al., 2019), deceptive design is about manipulating a user into doing something that is not in the user's best interest but in the interest of the owner of the website.

As people spend an increasing proportion of their lives on the web and the self-disclosure increase with more user generated content (Blackshaw & Nazzaro, 2006) and the use of e.g. social media (Kaplan & Haenlein, 2010), the need for digital integrity and privacy becomes important issues. This has become even more important during the Covid-19 pandemic, when an increasing part of our leisure and work time is spent online. Integrity is a "personal choice" (Killinger, 2010) and privacy is the "right to be let alone" (Warren & Brandeis, 1890). All this collection of data can e.g. be done by a company with, or without, consent or totally voluntarily by the users, so called personal informatics (Wilson, 2012). One important difference between the old analogue systems with notes and the current digital systems is the much higher traceability and foreverness of digital information and the opportunities to manipulate users using this information to alter systems (Kramer, Guillory & Hancock, 2014). This has led to an increase in different types of manipulative patterns in web design with the sole purpose to be deceptive and trick the users. A common way of creating deceptive design patterns and purposely tricking the user into making non-favorable decisions when collecting and using user data is to use cookies. Recently, there have been legal initiatives to try to strengthen the consumers

rights within this area e.g. the General Data Protection Regulation (GDPR) (European Commission, n.d.).

The objective of this paper is to discuss and analyze different deceptive web design patterns where users are expected to comply with the use of cookies. To achieve the objective the following research questions will be discussed: (1) What are the most commonly used deceptive design patterns and the effect of those? (2) How can users avoid involuntary sharing of personal data?

2   Background and Theory

The main idea behind web design and the guidelines (Nielsen, 2018; Shneiderman, 2009) for creating accessible and usable websites is to create an honest design. This is also one of the ten principles of good design stated by Dieter Rams in the 1970s - "Good Design is honest: It does not make a product more innovative, powerful or valuable than it really is. It does not attempt to manipulate the consumer with promises that cannot be kept" (Rams, n.d.). However, deception is pretty much the opposite. Deception can be described as the act of "hiding the truth" and in the realm of business this means "dishonest or illegal methods that are used to get something, or to make people believe that something is true when it is not" (Cambridge dictionary, n.d.). This can be used to describe design choices that have the users unconsciously share information which they normally would not do (D'Onfro, 2015).

The term Dark Patterns was coined in the aftermath of the boom of e-commerce websites that in order to generate sales and traffic were designed using deceiving user interfaces to manipulate users in different ways (Jaiswal, 2018). In a broader aspect, Dark Patterns use developers' and designers' knowledge of human psychology (Gray et al., 2018) and UX design to theoretically flip "honest" design into "evil" (Valjak, 2018).

The design patterns in Table 1, describes some of the psychological effects, inflicted among the users, that the designers want to build on. Persuasive design is a practice where the idea is to purposely influence the users' behaviors through the characteristics of a service. This can be done by designing for a behavior as a product of motivation, ability and triggers (Fogg, 2009). Consequently, many game

mechanics used to enhance activities through gamification (Robson, 2015; Zichermann & Cunningham, 2011) touch upon the problem and similar patterns as deceptive design. There are usually two sides - the same mechanics that can be used to enhance motivation to make users do good (Papworth & Mejtoft, 2015), can also be used to cause harm to the users in different ways.

Table 1: Description of common Dark Patterns (Brignull & Darlo, 2019)

Dark pattern	Description
Misdirection	The design purposefully focuses the attention on one thing to distract the attention away from another.
Sneak into Basket	The user attempt to purchase something, but somewhere in the purchasing journey, the site sneaks an additional item into your basket, often using an opt-out radio button or checkbox on a prior page.
Trick Questions	While filling in a form the users respond to a question that tricks the user into giving a non-intended answer. When glanced upon quickly, the question appears to ask one thing, but when reading carefully it actually asks another thing.
Bait and Switch	The users set out to do one thing, but a different, undesirable thing happens instead.
Confirmshaming	Guilting users to "opt in" by making them feel bad for saying no.

2.1 GDPR approved

The General Data Protection Regulation (GDPR) (European Commission, n.d.) is a regulation on privacy and data protection within the European Union (EU). The regulation applies to all companies that collect or use data of citizens that reside within the EU and the EEA, regardless of the location of the company. Since May 2018, when GDPR was put into action, Internet services who utilize cookies (Koch, n.d.) to collect user data are obligated to inform the user of e.g. what type of data is collected and how the service is using the data. To comply with the GDPR, Internet services appear to do the following (Dabrowski et al, 2019; Koch, n.d.): (1) services refrain from using persistent cookies at all, (2) EU users are banned from using the specific service, and (3) a service asks for explicit user consent and only then sets the cookies, leaving the site usable without consent. If consent is asked for, there is frequently a banner spanning over a service's pages asking for consent. The latter

alternative sites tend to use deceptive design patterns to have the user consent to cookies by applying different types of Dark Patterns.

To be able to define if the usage of cookies is according to the regulations, GDPR.EU (European Commission, n.d.) has a general explanation of what cookies are and how they should be implemented according to the GDPR (Koch, n.d.). There are four different types of cookies purposed - Strictly necessary cookies, Preferences cookies, Statistics cookies, and Marketing cookies. In this paper the focus is on the Strictly necessary cookies, as Strictly necessary cookies are the cookies that are essential for the user to be able to use the website and its features in an intended fashion. Other types of cookies are those which must be confirmed by the user or those that the user need to be informed about according to the GDPR. According to EU requirements, cookies on a website must comply with the following: (1) Have the users' consent before any cookies are in use, except strictly necessary cookies, (2) Provide the necessary information about the cookie and its collection of data before the consent, (3) Save the consent information from the user, (4) Provide the user of accessing the service even if they do not allow the use of certain cookies, and (5) Provide an easy way for the user to change their consent or cookie settings.

3 Method

A comprehensive analysis of 50 different home cooking recipe websites was done by first distinguishing how many websites were GDPR approved. Then out of the approved websites, an extended evaluation of the design was made to discern any types of deceptive designs. From these designs, an A/B test was created consisting of two websites (Test A and B) combined with a survey, to then be evaluated.

3.1 Website analysis

Using Alexa (2019) Top 500 Ranking, the top 50 websites on the list were analyzed by comparing each website's approach to using cookies to the definition of a correct usage according to GDPR.EU (European Commission, n.d.; Koch, n.d.). To see if the websites are using cookies correct each website was launched into Google Chrome Incognito mode, where the website was inspected with Developer tools >

Application > Cookies. In this mode all active cookies appeared, which could be categorized between Strictly necessary cookies or others.

After distinguishing which websites were GDPR approved, these websites were evaluated to determine if, and what type of, deceptive designs were being used. Brignull and Darlo's (2019) Types of Dark Patterns were used as a reference to find the most commonly used deceptive design patterns.

3.2   Testing and survey

An A/B-test was conducted with two websites created, to be able to compare two different cookie prompt approaches. Test A included a website with a cookie prompt made by using the most common Dark Patterns found from the Website analysis, and Test B's cookie prompt was made without Dark Patterns. The two different websites had three possible surveys, where all surveys had the same content and questions. The only thing differentiating the tests was the fact that if a participant Accepted or Declined the cookie prompt they would come to the corresponding survey. If the participant remained undecided, they stayed on the current survey (Figure 1).
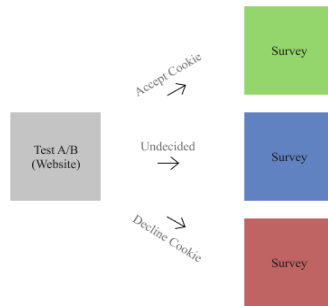
Figure 1: Flowchart showing how participants decision on *accepting*, *declining* or remaining *undecided* affected which survey was shown

Test A (small banner cookie prompt): The Test A cookie prompt (Figure 2a) was inspired by two defining Dark Pattern Methods - Sneak in the Basket and Misdirection. The prompt was small to not distract the user from the website content too much. The text on the prompt stated: "This site uses cookies to provide you with a great user

experience. By using this site, you accept our use of cookies.", which corresponded to the Sneak in the Basket method, as it informed the user that if a decision is not actively made, the user automatically accepted the cookies when using the site. Note that there were no other cookies used on the website than the strictly necessary ones, to be able to know if the participants Accepted or Declined the cookie prompt. The Accept button was green and bright, while the Decline button was gray an unattractive, which corresponded to Misdirection. The text even had a hidden link under "Use of Cookies", where it would explain what types of cookies were being used. This was purposely hidden to not draw attention away from a big green button.

| (a) Screenshot of Test A. The user is shown a small banner cookie prompt. | (b) Screenshot of Test B. The user is shown a full screen cookie prompt. |
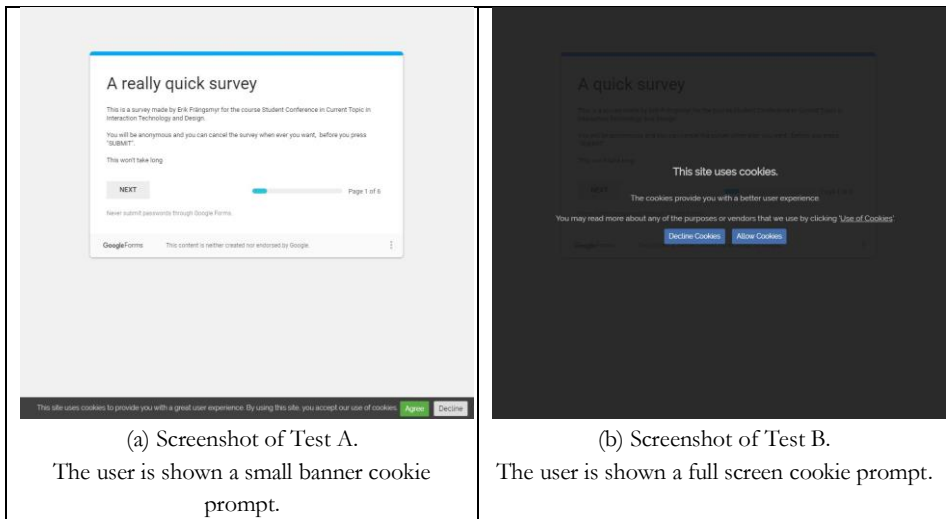
Figure 2: Screenshot of (a) Test A and (b) Test B

Test B (full screen cookie prompt): The cookie prompt created for Test B was much clearer and more informative than Test A (Figure 2b). The text was more informative and showed the option to go to "Use of Cookies" more distinctly. The option to Decline or Accept cookies was made less hierarchically than in Test A by having the buttons the same color and they explained what each button was meant to do, either to decline cookies or to allow cookies.

3.3   Participants

In total 40 respondents took part of the study, 20 respondents did Test A and 20 respondents did Test B. All participating respondents were from Sweden and was between the age of 20 and 65 years, where a majority were students between the age of 20-25 years. The research study was carried out during 2019. The tests were conducted online and a link to the corresponding test website was sent to each participant, where they could participate in the test and fill out the survey as they perceived it.

4   Results and discussion

Out of the 50 home cooking websites analyzed, less than half, 22 websites, turned out to comply to the GDPR. The other 28 websites either did not give a choice of complying with the cookies, or they asked the question for user compliance after the fact that they already ran all the cookies. Even if the usage of cookies was declined, the cookies were already in use and, consequently, the compliance did not matter. Hence, these websites did not meet the requirements of the GDPR legislation. Since the websites were chosen from the top of the Alexa Top 500 ranking, the websites in this test are very popular websites. Noteworthy is that still over half of 50 most common websites did not comply to the legislation.

Out of the 22 GDPR approved home cooking websites, 7 websites had no dark patterns and 15 websites had some kind of dark pattern in their design. From the websites analyzed there were two clear Dark Patterns used – *misdirection* and *sneak into basket*.

Misdirection methods were used in such way that the design purposefully focused the attention to accept all cookies by having the user focus on the biggest green button that says "Accept recommended setting" to distract the attention from reading more about the other cookies used. Out of the 22 GDPR approved websites analyzed, 4 websites had "misdirection" patterns.

Sneak into Basket methods were used in such a way that the user was prompted to comply with the website only uses the Necessary Cookies (Koch, n.d.) but it also adds third-party cookies without full consent from the user, which then manually

have to be removed through the use of an opt-out radio button or checkbox on a prior page. Out of the 22 GDPR approved websites analyzed, 11 websites had a "sneak into basket" design.

The task to present a how-to guide on how to avoid involuntary sharing of data is a seemingly hard task. One could think it would be as simple as being careful of not accepting the Privacy Policies or specifically the Use of Cookies. Maybe to go to the settings of every site and uncheck any unwanted cookie, or do so the first time launching a site. But to give this as convincing advice would give the false impression of having the control of what is being shared and not. Sadly, this is not the case, which can be seen from the results of the website analyses.

4.1 Cookie consent test

The cookie consent test was done on the two mock-up websites created to simulate the two dark patterns identified in the website analysis described above. The results from the survey show that all participants answered that they did *not* read the Use of Cookies page. One of the participants even expressed their concern about this issue by commenting this in the survey: "I usually get super annoyed when these prompts appear... I always click 'accept' because I have this weird fear that I won't be able to enter if I don't accept the cookies. I want to click 'decline', but for some reason I always accept." In Test A, where the users were shown the small banner cookie prompt a majority (60%) made no decision on either accepting or declining the cookies.

In Test B, with full screen cookie prompt, all participants made a decision, and a majority (80%) accepted the use of cookies. The full screen cookie prompt, however, made it more or less necessary to make a decision to remove the prompt and either continue to the website or go back. Sure, as a user, it is a good thing to feel like you are in control. But if you do not trust the website, does it really matter?

Is it really necessary for websites to use Dark Patterns? The results show that in Test A, a clear majority are not making an active choice of whether they want to accept or decline the cookie. However, when asked "Did you notice the cookie prompt on this site?" both tests resulted in similar answers, where the majority did notice the cookie prompt. However, almost non (in both cases) did read the read the "Use of

Cookies" page. Hence, even though many of the participants made a active choice to accept the cookies, how free is the choice when focus is not on the cookie consent but to consume actually information on the website. It is hard to believe that the users would not care about the privacy, but still no not take time to read through how personal data is used. The symbiosis of the commercial companies' collection of data for customization as well as the privacy and choice of the users have been discussed by e.g. Appelgren, Leckner and Mejtoft (2014).

5   Concluding discussion

Legislation has become more important for users to avoid involuntarily sharing of personal data. EUs GDPR is one of the more common legislations in recent years. However, looking at the Alexa Top 500 popular websites, only about half of the top websites complied with the legislation. Other ways of getting a user to share data is by designing to deceive the users to, in different ways, give away data. However, to avoid involuntary sharing of personal data it does not matter if deceptive design patterns are avoided or not if the website is not abiding by the regulations. While deceptive design has a purpose to create deceiving design based on the general design principles to make e.g. intuitive choices, it is important that honest design makes us think. Even though legislation tries to move the choices from the website to the user, it is clear that by using deceptive design patterns it is possible to somehow "bypass" the legislation. This is done by purposely designing for moving the users from the reflective situation that the cookie consent should be to an automatic behavior (Hansen & Jespersen, 2013). The most common ways to deceive the users to give away more information than needed, is the patterns *misdirection* and *sneak into basket*. Both of these either trick you into accepting all cookies or show you options but add cookies that are not necessary for function.

One way of dealing with the automatic behavior that cookie consent has become is to purposely introduce more intentional friction into the design that encourage a reflective behavior. Consequently, it is possible to focus on the important elements at hand and make users do reflective choices (Mejtoft, Hale, & Söderström, 2019; Hansen & Jespersen, 2015; Kahneman, 2008).

Even if websites ask for user compliance there is no way of knowing how they will be using the data without reading the Privacy Policy or Use of cookie page. If the website in question seems to follow the GDPR, the key to limiting sharing of involuntary personal data seems to be (in incremental order); 1) not trusting any websites, 2) become familiar with Developer Tools or Cookie managements for the preferred web browser and 3) make sure to look out for Dark Patterns in the design.

References

Appelgren, E., Leckner, S., & Mejtoft, T. (2014). The media consumers' conscious and unconscious choices – a key to understanding the news media consumption of tomorrow. In S. Zlitni, F. Liénard, D. Dula & C. Crumiére (Eds.), Communication électronique, cultures et identités (pp. 521-528). Editions Klog.

Blackshaw, P., & Nazzaro, M. (2006). Consumer-Generated Media (CGM) 101 (2nd ed.). Technical Report. Nielsen BuzzMetrics. Retrieved December 2, 2020, from http://www.nielsen-online.com/downloads/us/buzz/nbzm_wp_CGM101.pdf

Brignull, H. (2011, November 1). Dark Patterns: Deception vs. Honesty in UI Design. A List Apart 338. Retrieved January 2, 2021, from https://alistapart.com/article/dark-patterns-deception-vs-honesty-in-ui-design/

Brignull, H. (2013, August 29). Dark Patterns: inside the interfaces designed to trick you. The Verge. Retrieved June 2, 2020, from https://www.theverge.com/2013/8/29/4640308/dark-patterns-inside-the-interfaces-designed-to-trick-you

Brignull, H., & Darlo, A. (2019). Types of dark patterns. Retrieved April 2, 2021, from https://www.darkpatterns.org/types-of-dark-pattern

European Commission. (n.d.). Data protection in the EU. Retrieved April 1, 2021, from https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

Dabrowski, A., Merzdovnik, G., Ullrich, J., Sendera, G., & Weippl, E. (2019). Measuring Cookies and Web Privacy in a Post-GDPR World. In D. Choffnes and M. Barcellos (Eds.), Passive and Active Measurement (pp. 258–270). Springer.

Cambridge Dictionary. (n.d.). Deception. Retrieved February 7, 2020, from https://dictionary.cambridge.org/dictionary/english/deception

D'Onfro, J. (2015, October 3). LinkedIn might have to pay you money for spamming your email contacts. Business Insider. Retrieved June 1, 2020, from https://www.businessinsider.com/linkedin-settles-class-action-lawsuit-2015-10

Fogg, B. J. (2009). A Behavior Model for Persuasive Design. In Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09), Article 40. ACM.

Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The Dark (Patterns) Side of UX Design. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 534. ACM.

Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the Manipulation of Choice. European Journal of Risk Regulation 4(1), 3–28.

Alexa Internet. (2019). The top 500 sites on the web. Retrieved March 23, 2019, from https://www.alexa.com/topsites/category/Top/Home/Cooking.

Jaiswal, A. (2018, April 16). Dark patterns in UX. UX Collective. Retrieved June 4, 2020, from https://uxdesign.cc/dark-patterns-in-ux-design-7009a83b233c

Kahneman, D. (2008). Thinking, fast and slow. Farrar, Straus and Giroux.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite!. Business Horizons 53(1), 59–68.

Killinger, B. (2010). Integrity: Doing the Right Thing for the Right Reason (2 ed.). McGill-Queen's University Press.

Koch, R. (n.d.). Cookies, the GDPR, and the ePrivacy Directive. Retrieved June 25, 2020, from https://gdpr.eu/cookies/

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National Academy of Sciences 111(24), 8788–8790.

Mejtoft, T., Hale, S., & Söderström, U. (2019). Design Friction: How intentionally added friction affect users level of satisfaction. In Proceedings of the 31st European Conference on Cognitive Ergonomics (pp. 41-44). New York, NY: ACM.

Mejtoft, T., Ristiniemi, C., Söderström, U., & Mårell-Olsson, E. (2019). User experience design and digital nudging in a decision making process. In 32nd Bled eConference Proceedings (pp. 427-442). University of Maribor Press.

Nielsen, J. (2018). 10 Usability Heuristics for User Interface Design. Nielsen Norman Group. Retrieved June 6, 2020, from https://www.nngroup.com/articles/ten-usability-heuristics/

Papworth, S., & Mejtoft, T. (2015). Using game mechanics for motivational design in products and services. In 2015 ANZMAC Conference proceedings, 1047 – 1054.

Rams, D. (n.d.). The power of good design. Retrieved June 2, 2020, from https://www.vitsoe.com/us/about/good-design

Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2015). Is it all a game? Understanding the principles of gamification. Business Horizons 58(4), 411-420.

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs S. (2009). Designing the User Interface: Strategies for Effective Human-Computer Interaction (5th ed.). Addison-Wesley Publishing Company.

Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving decisions about health, wealth and happiness. Penguin Putnam Inc.

Valjak, A. (2018, April 16). Dark patterns in UX. Retrieved June 1, 2020, from https://infinum.com/the-capsizedeight/dark-patterns-designs-that-pull-evil-tricks-on-our-brains

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. Harvard Law Review, 4(5), 193-220.

Wilson, H. J. (2012). You, By the Numbers. Harvard Business Review, 90(9), 119 – 122.

Zichermann, G., & Cunningham, C. (2011). Gamification by Design. O'Reilly Media

DESIGNING CALL TO ACTION: USERS' PERCEPTION OF DIFFERENT CHARACTERISTICS

THOMAS MEJTOFT,[1] JONATHAN HEDLUND,[1]
HELEN CRIPPS,[2] ULRIK SÖDERSTRÖM[1] & OLE NORBERG[1]

[1] Umeå University, Digital Media Lab, Sweden; e-mail: thomas.mejtoft@umu.se, jonathan.hedlund.371@gmail.com, ulrik.soderstrom@umu.se, ole.norberg@umu.se
[2] Edith Cowan University, School of Business and Law, Perth, Australia; e-mail: h.cripps@ecu.edu.au

Abstract This paper aims to provide guidance when designing a call to action in a digital system with the purpose to create an intended feeling and user engagement. The paper is based on a test of four different simple call to action constructions. The users clearly prefer constructions that have a high explainability and feels intuitive. Hence, the design should have a high level of transparency and show the user straight away what it demands from the user and what the result of the action is. Furthermore, the design should have a high usability to make it clear how to use the call to action.

1   Introduction

Measuring and maximizing customer profitability is becoming more and more important in a constantly evolving landscape of customer engagement (Jorge, Pons, Rius, Vintro , Mateo, & Vilaplana, 2020; Gao & Javier, 2020; Pereira, de Fa tima Salgueiro & Rita, 2017). Companies are constantly looking to improve their customer relationship management (CRM) strategies (Kumar, 2013; North & Ficorilli, 2017). Measuring the click-through rate (CTR) for example gives a fair representation of how engaged the customers are with their website (Rashid, 2017; Liu, Chen, Chou, & Lee, 2018).

A call to action (CTA) is a marketing tool to prompt an immediate response or encouragement, is a crucial part for many websites (Horner, 2012; Steinberg, 2005). They can be the deciding factor between whether a user takes the action the company wants to induce (Chen, Yeh & Chang, 2020). Companies can use a call to action to create customer profitability that is easy to measure and analyze. Technically, a call to action is divided into several parts – getting attention, making it possible for the user to take action and make the user take the appropriate action from the user perspective. Call to actions can take many forms and might be designed in many different ways, some are buttons to direct the user to another page while some are small input forms to gather data from the user directly in the CTA component. Some examples of CTAs commonly used in marketing are components to encourage sign ups, subscription, and an option to learn more about a specific product or service (Chen, Yeh & Chang, 2020). There are many things that determine the effectiveness of a call to action on a website. The placement has an impact since the call to action must be found (Hernandez & Resnik, 2013), affordance-based cues (Norman, 2013) and skeuomorphs (Basalla, 1988) enabling the user to know how to engage with a call to action as well at the actual design of the call to action to induce a certain sense of feeling among the audience.

This paper will focus on different strategies of designing the call to action to create the intended feeling and user engagement. The objective of this short paper is to analyze how users interpret different designs of a call to action. The aim is to provide guidance to designers and managers that intend to design call to action on websites.

T. Mejtoft, J. Hedlund, H. Cripps, U. Söderström & O. Norberg:
Designing Call to Action: Users' Perception of Different Characteristics

407

2 Theory and background

Customer engagement advocates developing a portfolio of customers and nurturing this relationship (Gao & Javier, 2020; Imhof & Klaus, 2020). With this approach companies focus on how many products they can sell to a customer. How they can highlight the product benefit that aligns with customer need (Markey, 2020). Which customer segment they should focus on and what strategies should they use to develop their relationship (Schwartz, Bradlow & Fader, 2017; Jorge, Pons, Rius, Vintro☐ Carla, Mateo & Vilaplana, 2020). An important part of customer engagement is the management of profitable and unprofitable customers (Kumar, 2013; Imhof & Klaus, 2020). Click-through rate (CTR) is used to estimate the probability of users clicking on an advertisement or product displayed to them. CTR prediction is a key technique in internet advertising. Online advertising will play an increasingly important role in the future economy. Therefore, advertising CTRs is going to be the most important factor in developing the future of advertising. Accurately estimating the CTR of an advertisement has a crucial impact on the revenue for businesses (Jiang, Xu, Xu & Xie, 2021).

2.1 Call to actions

From a marketing perspective, a call to action has the aim to get the audience to perform a certain task immediately (Eisenberg & Eisenberg, 2006). It is a "call" to take an "action" that the sender of the message wants to induce. Strongly connected to the definition within marketing, in terms of web design, call to action is a term for "elements in a web page that solicit an action from the user" (Gube, 2009). Hence, a CTA is typically a part of the website or application that drives the reader to click through to engage further with a brand. A call to action is often created to drive the user to produce some type of immediate, measurable result (Chen, Yeh & Chang, 2020). Some examples of a CTA can involve a request to receive more information about a product or service.

According to Chen, Yeh and Chang (2020) some of the more popular manifestations of a CTA on websites are- (1) a link to a web page with additional and further information (e.g. 'Learn more'), (2) a request for the user to take action after browsing the web (e.g. 'Contact us'), and (3) the use of buttons that, when clicked, perform an action (e.g. 'Show now'). A CTA can vary in the amount of user

engagement, some CTA contains a button while some also contain an input field for the user to fill.
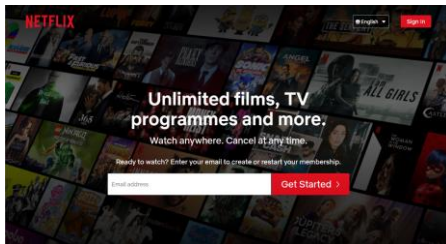
(a) Signup CTA.
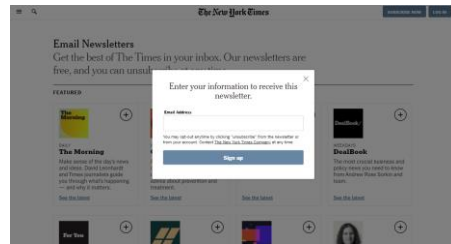
(b) Newsletter CTA.

(c) CTA to direct users to more info.
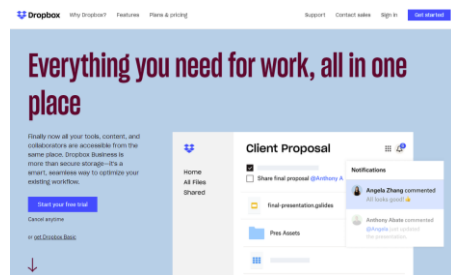
(d) Sign up to a premium service CTA.

Figure 1: Examples of different manifestations of how a CTA can be designed and used (screenshots by authors)

In Figure 1, some examples of different manifestations of how a CTA can be used to prompt the user to take an action are illustrated. The first example is from a note-taking application with a prompt to get the user to sign up for their service (Figure 1a). The CTA contains a short description of the service they are offering, two input fields, one for email and one for the password, and a button that creates the account. The second example is an example of a newsletter CTA (Figure 1b). The CTA contains a description of what the newsletter will do for the user. An input field for your email and a button to subscribe the user to the newsletter. The third example is from a drink producer with a CTA to prompt the user to find out more about their products (Figure 1c). The CTA contains a title and two buttons that direct the user to different pages, one to "Find out more" and one to "View products". The fourth example is a CTA from a music streaming service (Figure 1d). The purpose of the CTA is to prompt the user to sign up for their premium service. The CTA contains a description of their offer and two buttons to direct the user.

T. Mejtoft, J. Hedlund, H. Cripps, U. Söderström & O. Norberg:
Designing Call to Action: Users' Perception of Different Characteristics

409

2.2 Designing for effective call to action

There are several aspects that affect the effectiveness of a call to action within web design. The most prominent ones are based on the different steps that the user goes through when encounter a call to action – (1) noticing the call to action, (2) taking action, and (3) believing that the right action was taken. Attraction of attention is related to several issues both regarding the actual design and wording of the call to action (Bashinsky, 2016) but also the placement of the call to action. Previous research has identified how users scan through a website to find placement for important information (e.g. Hernandez & Resnik, 2013). Regarding the possibility to make the user take action it is needed that that the user understands that an action is possible from a technical perspective. This is related to two issues - the affordance-based cues in the actual design (e.g. understanding that a button can be clicked) and the usability of the system (e.g. using the call to action is intuitive). Affordances (Gibson, 1977; Norman, 2013) define what actions are possible when facing an object. Hence, in digital systems affordances provide strong clues to how something is going to be used, often based on connection to the non-digital world, so-called skeuomorph (Basala, 1988). A skeuomorph is a design element that has no necessary meaning in the new setting but was essential in the old setting (e.g. the raised button with, consequently, a shadow). These affordance-based ques are important from an interaction design perspective to provide the users with information about how to use certain features (Kaptelinin, 2014) and, consequently, pushing for a call to action. Creating high usability (Nielsen, 1994) and frictionless design is often in focus and make it intuitive for the user to navigate through a design and make choices. However, the frictionless design might decrease the possibility to make reflective choices due to the higher speed of navigation. It is therefore important that the design a high explainability. An explainable design is easy for the user to understand the effect when a certain task is taken, and it create a design that is in line with the ideas of Dieter Rams (n.d.) – Good design is honest and "does not attempt to manipulate the consumer with promises that cannot be kept". Design friction has been discussed during latter years due to its ability to make the user reflect on choices and, hence make the user take action the user intended to take (Mejtoft, Hale & Söderström, 2019).

3   Method

The aim of this study is to determine how users interpret different designs of a call to action. The study presents different construction of a CTA to see which construction the users prefer and how the design is interpreted by the users. Four different constructions of a CTA with the purpose to get the participant to sign up for a newsletter were presented to the participant. Followed by a short survey where the participant could choose which CTA they preferred and why. The tests were conducted during fall 2020. The test was conducted on 32 participants with an even distribution between male and female participants.

3.1   The different call to actions designs

As previously mentioned, four calls to action were constructed to be used in the study. The purpose of the CTA is to get the participant to sign up for a newsletter. The CTA contains a header, a description, a text field, and or a button for the participant to interact with. The CTAs differed in the level of user engagement presented in the initial component, as well as swapping the headers to be either informative or appealing.

Two constructions gave the participant a text field to fill out their email as well as a button to complete the action. The other two constructions gave the participant a button to press to start the process of signing up to the newsletter. The header and description were also swapped to determine if the participant wanted the CTA to be more appealing or informative. The header was either "Sign up to our newsletter" to be informative or "With our tips you will become a better manager" to be appealing for the participant. The description of the CTA was dependent on the header, if the header was "Sign up to our newsletter" then the description was "With our tips you will become a better manager" and vice versa.

The four CTAs used in the study (Figure 2) had the following characteristics – (a) the header is informative and the description is appealing while also having a higher initial user engagement with a text field and button, (b) the header is informative and the description is appealing while also having a lower initial user engagement with only a button, (c) the header is appealing and the description is informative while also having a higher initial user engagement with a text field and button, and (d) the

T. Mejtoft, J. Hedlund, H. Cripps, U. Söderström & O. Norberg:
Designing Call to Action: Users' Perception of Different Characteristics

411

header is informative and the description is appealing while also having a lower initial user engagement with only a button. To remove the effect of e.g. skeuomorphs (Kaptelinin, 2014), typography and font size were consistent (Koníček & Světlík, 2020). These things that could affect the CTA was omitted in the study.
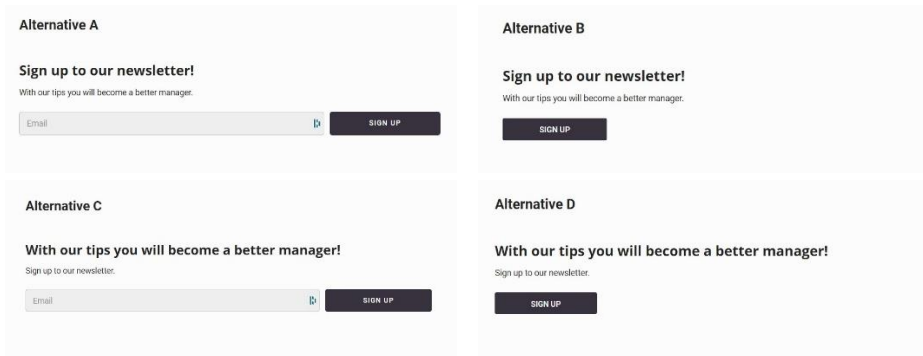
Figure 2: The four CTAs used in the study

3.2 Setup

Due to restrictions, the test was performed online. However, it is the authors strong belief that the validity or reliability of the results are not affected by the online test environment. An online test environment creates a realistic environment for this test. The participant was sent a link to a webpage where the participant would first read a short description of what the test was going to be used for, as well as an assurance that their personal information would not and could not be connected to their result. The participant would then be presented with the CTA used in this study in the same order they are presented in this paper. After the participant decided which CTA they preferred they would fill out a questionnaire regarding their preferred choice. Some basic demographics were gathered and analyzed without any significant differences being found.

4 Results and discussion

The effectiveness of a CTA is an important field to study, both from a marketing perspective and a design perspective. Focusing on developing strategies to enhance companies' customer relationship management (CRM) is crucial for success in an everchanging business landscape. Measuring and maximizing profitability of online

assets is becoming increasingly important. One way to increase profitability is by improving one's use of CTA. For user experience designers monitoring the click-through rate can be a great way to evaluate how well a website is fulfilling its purpose.

The results from the questions about which CTA the participant prefers are shown in Figure 3. Most participant preferred design A and design C and there were no major differences between male and female participants. Both Design A and Design C wanted the participant to fill in their email right away and had an upfront higher user engagement than Design B and Design D. For a CTA with few inputs, e.g. signing up for a newsletter, the user seems to prefer being able to see directly what they need to fill out to complete the action.
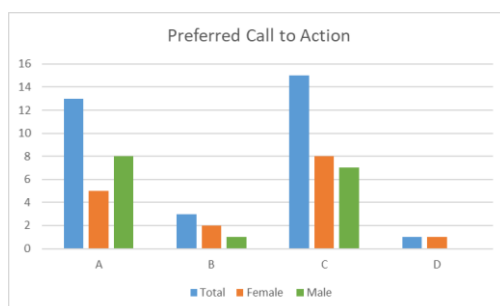
Figure 3: Number of votes each CTA received in the study

Of the 34 participants 24 out of the 32 were cited motivations for their choice of CTA. The motivations were analyzed and categorized based on the reasons for making a certain choice. Patterns were found and five different reasons based on design, usability and marketing theories and reasoning, emerged from the answers. The reasons were classified as *intuitive and frictionless*, *High explainablity and honest*, *Aesthetic and appealing*, *attention-grabbing* and *design friction*. In Figure 4 the total number of mentions for each of these reasons are plotted for the different designs.

T. Mejtoft, J. Hedlund, H. Cripps, U. Söderström & O. Norberg:
Designing Call to Action: Users' Perception of Different Characteristics

413
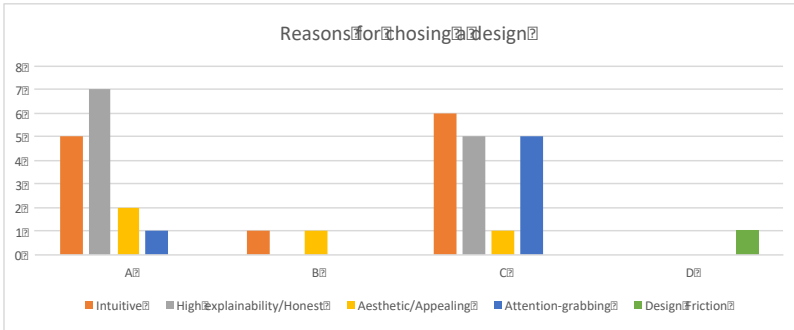
Figure 4: Reasons mentioned for choosing a specific design
(one participant could mention several reasons)

Since very few respondents preferred Design B and Design D there were few comments mentioned. However, there were comments mentioned for *not* choosing Design B and Design D e.g.: "Design B and D require me to click on and then I do not know what will happen after that" and "unclear what happens if I click 'SIGN UP' on B and D, a box will appear where I enter an email or have my email via a cookie, etc.". Thematically, these comments were in line with the user getting a feeling of the website wanting to trick the user into doing something that is not in the users' best interest, so-called deceptive design or dark patterns (Brignull, 2011).

Design A was the second most preferred by the participants and was high regarding the factors of intuitiveness and high explainability/honest. Regarding intuitive, respondents stated e.g. that "a clear title that describes the purpose and functionality, and when I see that I only have to enter in one field makes it feels frictionless. Nothing is hidden or feels indistinct that would otherwise make me hesitate", "it felt like the easiest way to sign up", "Easy to Sign Up without having to click on to another page to enter your email" and "Simplest and clearest". Regarding having a high explainability and a feeling of being honest in the design the respondents stated that "if the email field is above the button, you also know what it will mean if you press the button", "shows that they only need my email and nothing more", "prefer text fields for emails as I do not know how the others will then get hold of my email, a new page opens?" and "just need to fill in the email so it's clear, nothing uncertain. With only a button and no input field, you would be taken to a new page, and you do not know what more you may need to do".

Design C was the most preferred design being high regarding the factors intuitive, high explainability/honest and attention-grabbing. Regarding intuitive, the respondents stated that it is "nice to avoid through to another page", "easy to get the newsletters with as few clicks as possible" and "nice to fill in the email immediately so you avoid another step". The reasons for high explainability were mainly due to reasons like: "I think it's nice that it clearly says why I should sign up first", "like the fact that I get everything at once. That I have to enter everything and then click 'sign up'" and "that it is email contact and I get a direct answer to what it can give me". The respondents also believed that this design attracted attention since e.g. "attractive title, not with the main goal of getting people to sign up, "it attracts more attention" and "you have to attract people to get them to sign up for newsletters, by offering something they do not get somewhere else".

One recurring opinion on whether or not to have the input field was that the participants preferred to be provided with everything at once, making the interaction more frictionless. With only a button and no input field, some of the participants wrote that they felt unease with not knowing where the button would take them or if a modal would pop up or they would get directed to a new page. Some participants preferred the informative title and being able to see right away that it was a newsletter they were asked to sign up for while others felt that a more appealing title was attracting more of their attention. On the other hand, there are no clear results on whether or not the participant had a preference for the title of the CTA to be more appealing (Design C) or informative (Design A) since they both was equally preferred. A few possible errors could have affected the outcome of this study. One of them being the choice of the appealing title used in this study. All of the participants were presented with the CTA in the same order presented in this paper, which could have an impact on their choice. A study where the different CTA are presented in random order would remove this error.

5   Conclusions

There is a wide range of call to actions found online. This study investigates two different characteristics and the impact of these regarding the reactions of the users' and how they want to engage with the call to action. The results indicate positive attitudes towards the design features such as intuitive and high explainability/honest are important for the users' willingness to engage with a call to action. Clearly

T. Mejtoft, J. Hedlund, H. Cripps, U. Söderström & O. Norberg:
Designing Call to Action: Users' Perception of Different Characteristics

415

showing the user what to expect is important when designing for a high click through rate. Hence, the study shows that moving the user engagement upfront is better for enabling users to trust the design. This makes the call to action more transparent by showing the user right away what they need to fill out to complete the action.

Acknowledgements

References

Basalla, G. (1988). The evolution of technology. Cambridge, UK: Cambridge University Press.

Bashinsky, A. (2016, January 13). 6 Easy Ways to Incorporate Emotional Appeals Into Your Website's Call to Actions. Entrepreneur. Retrieved December 13, 2020 from https://www.entrepreneur.com/article/253284.

Brignull, H. (2011, November 1). Dark Patterns: Deception vs. Honesty in UI Design. A List Apart 338. Retrieved March 2, 2021, from https://alistapart.com/article/dark-patterns-deceptionvs-honesty-in-ui-design/.

Chen, T.-Y., Yeh, T.-L., & Chang, C.-I. (2020). How different advertising formats and calls to action on videos affect advertising recognition and consequent behaviours. The Service Industries Journal, 40(5-6), 358–379.

Eisenberg, B, & Eisenberg, J. (2006). Call to Action: Secret Formulas to Improve Online Results. Nashville, TN: Nelson Business.

Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), Perceiving, Acting, and Knowing: Towards an Ecological Psychology (pp. 67-82). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gao, L., & Javier, S. F. (2020). Customer equity drivers, customer experience quality, and customer profitability in banking services: the moderating role of social influence. Journal of Service Research, 23(2), 174–193.

Gube, J. (2009, October 13). Call to Action Buttons: Examples and Best Practices. Smashing Magazine. Retrieved January 30, 2021, from https://www.smashingmagazine.com/2009/10/call-to-action-buttons-examples-and-best-practices/.

Hernandez, A., & Resnik, M. L. (2013). Placement of Call to Action Buttons for Higher Website Conversion and Acquisition: An Eye Tracking Study. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57(1), 1042-1046.

Horner, T. (2012, April 30). Writing a Better Call to Action. Marketing Professors. Retrieved December 13, 2020 from http://www.marketingprofs.com/articles/2012/7772/writing-a-better-call-to-action.

Imhof, G., & Klaus, P. (2020). The dawn of traditional cx metrics? examining satisfaction, exq, and war. International Journal of Market Research, 62(6), 673–68.

Jiang, D., Xu, R., Xu, X., & Xie, Y. (2021). Multi-view feature transfer for click-through rate prediction. Information Sciences, 546, 961–976.

Jorge, O., Pons, A., Rius, J., Vintro C., Mateo, J., & Vilaplana, J. (2020). Increasing online shop revenues with web scraping: a case study for the wine sector. British Food Journal, 122(11), 3383-3401.

Kaptelinin, V. (2014). Affordances. In The Encyclopedia of Human-Computer Interaction (2nd Ed.)
        (Chapter 44). Interaction Design Foundation. Retrieved May 18, 2021 from interaction-
        design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-
        ed/affordances

Koníček, I., & Světlík, J. (2020). Some principles for the design of successful and appealing websites.
        European Journal of Media, Art and Photography, 8(1), 106-116.

Kumar, V. (2013). Profitable customer engagement: Concept, metrics and strategies. SAGE.

Liu, D.-R., Chen, K.-Y., Chou, Y.-C., & Lee, J.-H. (2018). Online recommendations based on dynamic
        adjustment of recommendation lists. Knowledge-Based Systems, 161, 375–389.

Markey, R. (2020). Are you undervaluing your customers? Harvard Business Review, 98(1), 42–42.

Mejtoft, T., Hale, S., & Söderström, U. (2019). Design Friction: How intentionally added friction affect
        users level of satisfaction. In M. Mulvenna & R. Bond (Eds.), Proceedings of the 31st European
        Conference on Cognitive Ergonomics (pp. 41-44). New York, NY: ACM.

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. CHI '94: Proceedings of
        the SIGCHI Conference on Human Factors in Computing Systems, 152-158.

Nielsen, J. (2012). Usability 101: Introduction to Usability. Nielsen Norman Group. Retrieved January
        12, 2021, from https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Norman, D. (2013). The design of everyday things: Revised and expanded edition. New York, NY:
        Basic Books.

North, M., & Ficorilli, M. (2017). Click me: an examination of the impact size, color, and design has
        on banner advertisements generating clicks. Journal of Financial Services Marketing, 22(3), 99–
        108.

Pereira, H. G., de Faᵀtima Salgueiro, M., & Rita, P. (2017). Online determinants of e-customer
        satisfaction: application to website purchases in tourism. Service Business, 11(2), 375–403.

Rams, D. (n.d.). The power of good design. Retrieved January 23, 2021, from
        https://www.vitsoe.com/gb/about/good-design

Rashid, U. (2017). A framework to explore results in multiple media information aggregated search.
        Multimedia Tools and Applications, 76(24), 25787–25826.

Schwartz, E. M., Bradlow, E. T., & Fader, P. S. (2017). Customer acquisition via display advertising
        using multi-armed bandit experiments. Marketing Science, 36(4), 500–522.

Steinberg, B. (2005, March 22), 'Call to Action' Ads Give Clients Results They Can Measure. Wall Street
        Journal. Retrieved November 13, 2020 from
        https://www.wsj.com/articles/SB111145597859585890

FORMING SUSTAINABLE PHYSICAL ACTIVITY PROGRAMS AMONG YOUNG ELDERLY - A COMBINED ELM & UTAUT APPROACH

CHRISTER CARLSSON,[1] PIRKKO WALDEN,[1]
TUOMAS KARI,[1,2] MARKUS MAKKONEN[1,2] &
LAURI FRANK[2]

[1] Institute for Advanced Management Systems Research, Turku, Finland; e-mail: christer.carlsson@abo.fi, pirkko.walden@abo.fi, tuomas.t.kari@jyu.fi, markus.v.makkonen@jyu.fi
[2] University of Jyvaskyla, Faculty of Information Technology, Jyvaskyla, Finland; e-mail: lauri.frank@jyu.fi

Abstract There is consensus in health studies that regular physical activities of sufficient intensity and duration contribute to better health both in the short and long term. In an ongoing research program, we focus on getting the young elderly, the 60-75 years age group, to adopt and include physical activities as part of their daily routines. One reason for addressing young elderly is large numbers – in Finland health care costs for the elderly was 3.7 B€ in 2019 and will increase dramatically if the young elderly group is in bad shape when they reach the 75+ age group. We are finding out that systematic physical activities can serve as preventive health care for the young elderly. We are also learning that digital services can be instrumental for building sustainable physical activity programs.

1   Introduction

There is a consensus (Bangsbo et al (2019)) that regular and systematic physical activities (PA) can serve as preventive health care (Jonasson (2017)). In the following we will refer to a set of systematic PA as a *PA program*, if the physical activities are carried out over a period of time and to reach some goal,   The Copenhagen Consensus statement of 2019 (Bangsbo et al (2019)): (i) "being physically active is a key factor in maintaining health …" (ii) "physically active older adults, compared with older inactive adults, show benefits in terms of physical and cognitive function …" (iii) "physical inactivity in older adults is associated with a trajectory towards disease and increased risk of premature all-cause mortality …". Health recommendations agree on that PA at moderate intensity for at least 150 minutes per week will have health effects (Wallen et al (2014)). This standard applies to healthy adults; in practice there will be differences in the actual PA programs; the differences are in terms of female/male, age groups, level and history of PA capability, physical requirements for everyday activities, and to meet individual long- and short-term objectives. The Copenhagen Consensus also finds that less than 150 minutes could be sufficient for older adults (Bangsbo et al (2019)). Further study will, no doubt, show that a good variety of PA programs will give individually sufficient health effects.

Our focus is the young elderly – the 60-75 years age group – and we want to work out sustainable PA programs that could give preventive health effects for this group. We immediately found common ground with the pensioners' associations that have developed PA programs for their members. There was, however, a feeling among active pensioners that these activities (i) are not intensive enough, (ii) are not of sufficient duration, and (iii) are not regular enough to be sustainable habits; in short, they did not meet the perceived standards for preventive health care (Reyes-Mercado (2018), Hukkanen et al (2018)). The associations found that our proposed program – called *DigitalWells* (DW) – could be useful for them and agreed to work with us; they stated as a requirement that we also include the 75+ age group, who regard themselves as modern time young elderly.

C. Carlsson, P. Walden, T. Kari, M. Makkonen &L. Frank:
Forming Sustainable Physical Activity Programs Among Young Elderly - A Combined ELM &
UTAUT Approach

419

DW is designed as an interactive research and development program aimed at building sustainable PA programs for young elderly.

There are a few central elements that define the program (cf. Carlsson and Walden (2018, 2019)); <u>first</u>, we use a synthesis of existing research results to develop PA programs for groups of young elderly; <u>second,</u> experience collected from field studies is used to validate and verify the theory constructs; <u>third,</u> we aim for PA programs which as such will give the participants short- and long-term health effects; <u>fourth,</u> the PA programs should be sustainable, i.e. the participants should stay with them for extended periods of time (months and years, more than weeks); <u>fifth,</u> we design and implement digital services to guide and support users to stay with and be active with their PA programs.

The participants are all volunteers, at the moment (May 2021) about 750 participants in 30 groups, they come from the local associations of three pensioners' federations which have more than 230 000 members.

DW works out contrasts between the users' own perceptions of their physical activities, intentions to spend (more) time with their PA programs and the time they have actually spent. The users are anonymous and identified only with individual 8-digit pseudonyms in both the surveys that cover perceptions and intentions, in the DW-app 2.5 (cf. section 2) and in the secure data cloud. This makes it possible to trace their activities and to test and validate theory frameworks for the adoption of PA programs.

With the user groups and the digital service tools developed in DW (cf. Kari et al (2020a, 2020b); Makkonen et al (2020a, 2020b)), we aim to tackle the following research questions:

- What drivers could get young elderly users to adopt and use PA programs?
- What factors or drivers could help support and sustain the adoption of PA programs?

The DW research program is running for several years and is ongoing; in this paper we formulate some initial, partial answers to the research questions, - i.e. in short, how to get young elderly to become users of PA programs and to stay with them.

In section 2, we will introduce the context and present the support tools we use for the young elderly users; in section 3 we work through the conceptual frameworks of ELM and UTAUT (introduced in detail in the section); section 4 offers answers to the research questions and adds details from previous research to the framework to offer roadmaps for further research.

2 Context and Tools for Young Elderly

Physical activity among Finnish adults is way too low. The 2010-2017 studies in the health and wellness of adults in Finland (THL, 2019a) shows that in the age group of 30-54 years, only 30% spent several hours per week at regular physical activities; in the 55-74 age group, it decreases to 15%; in the 75+ age group only 7% are regularly physically active. Moreover, the recent FinHealth2017 study (THL, 2019b) found that only 39% of men and 34% of women reach recommended levels of PA to get health effects. The FinHealth2017 study builds on a representative sample of the Finnish population.

Physical wellness comes from physical exercise to build stamina, muscle strength, and balance, and to ward off age-related serious illness. Sustained physical exercise helps to meet everyday requirements of life. Studies (Jonasson, 2017; Wallén et al., 2014) show that systematic PA contributes to good quality of life in senior years.

The DW research program aims to meet the recommendations of Bangsbo et al (2019), Hukkanen et al (2018), Jonasson (2017) and Wallén et al (2014). In the following, we will work through some details.

DW has found (cf. Kari et al (2020a), (2020b), Makkonen et al (2020a, 2020b)) that typical PA forms for young elderly include walking, yard work, Nordic walking, cycling, cross-country skiing, golf, gym training, swimming and home gymnastics. The FinHealth2017 (THL, 2019b) study found that popular forms of PA mostly were the same; the study looks at physical activity during leisure time and registers

activities that are carried out several times per week. What is interesting, there is no dramatic decrease in PA when moving to the older age group.

The FinHealth2017 study (THL, 2019b) offers a couple of challenges for the young elderly age group: (i) the proportion of young elderly that carries out PA according to health recommendations should be much higher; (ii) PA carried out several times per week should be of sufficient intensity and duration.

Over time, when PA programs developed in the DW research program have been adopted and taken into use, young elderly will start meeting the FinHelath2017 recommendations. The DW actually works out a list of proposed PA forms that will give sufficient PA intensity and duration ; some details on how this was worked out in the next few steps.

The 2011 Compendium of Physical Activity (CPA) (Ainsworth et al., 2011) offers advice and guidelines for the design of PA programs that will start giving health effects. CPA quantifies the energy cost of 821 specific activities in terms of the metabolic equivalent of task (MET) (cf. fig.1). This offers support for dealing with both challenges (i) and (ii). METs show the energy cost (effort) of a PA relative to sitting (Ainsworth et al., 2011). It is measured with groups in actual exercises and in lab experiments, which makes it verifiable and useful for goal setting, for registration of activities and for follow-up of PAs.

In the DW we adopted the CPA guidelines as a baseline for the PA programs (cf. fig.1). We work with 35 different activities that can be included in a PA program; we replaced the CPA MET levels with more "activity-describing" labels: low = light; medium = moderate; high = vigorous.

English	Finnish	Swedish	CPA MET		
			Low	Medium	High
Basketball	Koripallo	Basketboll	4.5	6.0	8.0
Soccer	Jalkapallo	Fotboll	4.9	7.0	10.0
Volleyball	Lentopallo	Volleyboll	3.0	4.0	6.0
Ice hockey	Jääkiekko	Ishockey	6.4	8.0	10.0
Baseball	Pesäpallo	Boboll	4.0	5.0	6.0
Floorball	Salibandy	Innebandy	6.2	7.8	9.8
Rowing or paddling	Soutu tai melonta	Rodd eller paddling	2.8	3.5	5.8
Rowing indoors	Soutu soutulaitteella	Rodd med roddmaskin	3.8	4.8	6.0
Swimming	Uinti	Simning	3.5	6.0	9.8
Aquajogging	Vesijuoksu	Vattenlöpning	2.5	4.5	6.8
Water aerobics	Vesivoimistelu	Vattengymnastik	3.2	5.5	9.0
Ice skating	Luistelu	Skridskoåkning	5.5	7.0	9.0
Cross-country skiing	Maastohiihto	Längdskidåkning	6.8	9.0	12.5
Aerobics	Aerobic	Aerobics	5.0	7.3	10.0
Jogging	Hölkkä	Joggning	6.0	7.0	8.0
Walking	Kävely	Gång	2.8	3.5	4.3
Yoga	Jooga	Yoga	2.0	2.5	4.0
Running	Juoksu	Löpning	6.0	8.0	9.8
Stretching	Venyttely	Töjning	1.8	2.3	3.0
Nordic walking	Sauvakävely	Stavgång	3.4	4.8	6.8
Home gymnastics	Kotivoimistelu	Hemgymnastik	2.8	3.8	8.0
Partner dance	Paritanssi	Pardans	3.0	4.5	5.5
Badminton	Sulkapallo	Badminton	4.3	5.5	7.0
Squash	Squash	Squash	4.4	7.3	12.0
Tennis	Tennis	Tennis	5.0	7.3	8.0
Boccia	Boccia	Boccia	2.6	3.3	4.2
Golf	Golf	Golf	3.5	4.8	5.3
Bowling	Keilailu	Bowling	3.0	3.8	4.8
Gym training	Kuntosaliharjoittelu	Gymträning	3.5	5.0	6.0
Circuit training	Kuntopiiri	Cirkelträning	2.3	4.3	8.0
Orienteering	Suunnistus	Orientering	6.8	9.0	11.0
Cycling	Pyöräily	Cykling	6.8	8.0	10.0
Cycling indoors	Pyöräily kuntopyörällä	Cykling med motionscykel	4.8	7.0	8.8
Household work	Kotityöt	Hushållsarbete	2.8	3.5	4.3
Yard work	Pihatyöt	Gårdsarbete	3.0	4.0	6.0
Other activity	Muu aktiviteetti	Annan aktivitet	4.1	5.6	7.5

Figure 1: List (subset) of physical activities, including CPA MET values

The names of the activities are standardized in English, Finnish and Swedish. The list of 35 activities is now a tool and a guide to find systematic and practical ways to build health-effective PA programs for young elderly. The health recommendations (THL, 2019a; Bangsbo et al., 2019) correspond to roughly 650 MET-minutes per week, which is a minimum to give short- and long-term health effects. A DW participant will simply work down the list and compose his/her weekly PA program for sufficient health effects.

In the DW we opted for composing and running weekly PA programs and the registration of the actual activities with digital support. For this purpose, we developed an application (cf.fig.2) for smart mobile phones (Android, iOS), now called the DW-app (2.5), which is still being developed with new and more advanced features.

C. Carlsson, P. Walden, T. Kari, M. Makkonen &L. Frank:
Forming Sustainable Physical Activity Programs Among Young Elderly - A Combined ELM &
UTAUT Approach

423
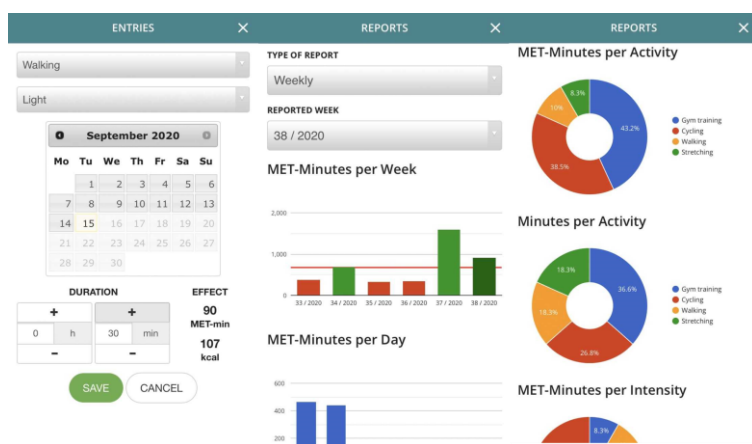
Figure 2: Logging and reporting of activities

The logging of activities on the smart phone is done in the left part of the screen (cf. fig.2): (i) the user selects the activity (*walking*), (ii) the intensity (*light*), (iii) the *date* from the calendar, (iv) the duration (*hours, minutes*) after which the app (v) calculates and shows the effect of the activity (*MET-min, kcal*).

The results of the activity are used to update a secure database where the user is registered with an individual 8-digit pseudonym. The data is used to produce reports on the smart phone for the user (the middle part): (i) the type of report is specified (*weekly*), then (ii) the reported week (*38/2020*), and (iii) the wanted report is shown (*MET-minutes per week*) and (iv) further specified (*MET-minutes per day*). Further graphical reports are shown in the right side that specify (for instance) MET-minutes per activity and Minutes per intensity.

A growing number of users have opted to get even more automated and use sports trackers and smart watches to collect and register their activities; the DW-app 2.5 now collects data from these devices, up-dates the secure database and produces the reports for the user on his/her smartphone.

There are quite a few research reports published on the use of the DW program and the DW-app 2.5 (cf. Carlsson et al (2020), Kari et al (2020b), Makkonen et al (2020b), Kettunen et al (2020)); here it suffices to give a snapshot of user comments selected

from 265 respondents that have participated in the program for about 12 months (cf. Table 1). The comments have been selected with an aim to find both positive and negative experiences, but they also give a glimpse of the attitudes to exercise that can be found among the participants. The key to the quotes (translated from Finnish and Swedish) is: [participant #] M or F/age/BMI.

Table 1: Participant reactions to the WFR after about 12 months

The distance should be measured, in km or number of steps … there should be more measures for physical work [24] M/71/26.40

Summertime changes habits, we spend time at the summer cottage with no regular walking tours; fall is a better time for exercise [49] F/69/22.04

I regularly spend time with exercise; thus, the WFR has not changed my exercise routines [64] F/71/26.40

I spend as much time with exercise as feel good (quite much) [65] M/74/25.86

This program does not motivate to exercise as such; it is too simple and cumbersome; I am in favour of technology that automatically registers exercise and other activities; MET points are not calibrated to active exercise nor to health exercise [72] F/70/23.73

I left the WFR in mid-summer because my motivation was not sufficient to go on by myself [covid-19 restrictions closed group activities]; technology by itself is not enough as a motivation [75] F/68/33.25

Our WFR group was closed because of the covid-19 restrictions; all the indoor group events that I was attending every week were closed; I had to switch to walking and Nordic walking, which I increased quite a lot (to several times per day); now fall is coming but it seems as if the restrictions will continue [78] F/65/30.48

I follow up on my own exercise more than before and I have also been checking my results both from the tracker and from the [app on the] phone [83] M/71/26.01

I am not yet familiar with the WFR; I need to get guidance in person; I does not want to spend time to find out the functions for myself from the Internet [112] F/81/21.09

I increased exercise after I retired, because now I have more time for it; WFR and the need to reduce weight have increased exercise; yet I have not decided on any goals [120] M/63/34.48

[WFR] has improved the follow up of different forms of exercise [128] M/75/35.98

This [WFR] is good for people that have not been active on exercise; for me personally there is not much effect [134] F/66/23.59

The application has made me surer that my exercise activities are quite sufficient without any programs; I easily get 10 000 steps every day in my daily routine tasks [141] F/67/23.38

Before the covid-19 restrictions I worked out in gym programs 3 times/week and spent 3 times/week in water aerobics or aqua-jogging; now everything is closed which reduced my exercises to yoga and qigong once a week [164] F/71/27.83
A very good program that makes me follow my own activity program much better [174] M/66/28.78
For myself, joining the program did not "force" me to increase exercise [180] M/70/25.51

The DW-app 2.5 is a digital service, and we know something about digital services. Drivers for the adoption of digital services have been identified and studied with a basis in the UTAUT2 model. They include (Yuan et al., 2015):

- performance expectancy ("degree to which the use of a technology will help users to perform chosen activities")
- effort expectancy ("degree of ease in the use of a technology")
- social influence ("perception that important others support the use of technology")
- facilitating conditions ("factors that facilitate or impede adoption of technology")
- hedonic motivation ("fun or pleasure with using a technology")
- price value ("trade-off between perceived benefits of and monetary cost for using a technology")
- habit ("perception of automatically engaging in a certain behaviour").

Thus, we can find drivers that could make participants adopt and use the DW-app 2.5, but will that be sufficient to make them adopt and use a PA program? Then sustain the use of the adopted PA program. We will try out the proposal that UTAUT2 *will not be sufficient* as a conceptual framework for the adoption and use of PA programs; cf. Table 1 ["*This program does not motivate to exercise as such; it is too simple and cumbersome*" or "*I spend as much time with exercise as feel good (quite much)*"]. We propose to use the Elaboration Likelihood Model (ELM) to find drivers for the adoption and use of PA programs.

3   ELM and UTAUT and Variations

The ELM was introduced by Petty et al (1995), (i) to work out influence processes and their impacts on human perceptions and behaviour, and (ii) to explain why a given influence process may have varying outcomes across different users in each context. This is promising for our aim to find drivers that could get the DW program adopted for sustained use among young elderly.

Bhattacherjee and Sandford (2006) use the ELM as a theory framework to describe and explain IT usage intention. They state, following the Diffusion of Innovation theory (Rogers (1995)), that the acceptance of IT is fundamentally a problem of social influence. The ELM offers instruments for systematic studies of alternative influence processes and their effects, and the impact of moderating factors.

Bhattacherjee and Sandford (2006) make a distinction between argument quality as the key influence process and peripheral cues as moderating factors. They work out perceived usefulness as driven by argument quality to influence the intention to use IT (the perceived usefulness is one of the drivers in the TAM framework (Davis (1989)). They add source credibility and attitude as additional influences on perceptions of perceived usefulness. Finally, job relevance and user experience, are peripheral cues and moderating factors.

The key influence processes are surprisingly similar to the social influence processes we assume to describe and explain intentions for a sustained use of PA programs (cf. fig.3). The main influences come from perceived usefulness ("to get more good years"), which in turn is influenced by argument quality and source credibility (cf. the Copenhagen Consensus and verifiable medical research results); argument quality is at least partly influenced by source credibility; in some cases, argument quality can help decide source credibility.
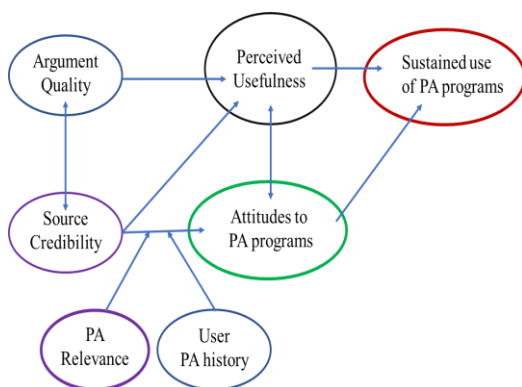
Figure 3: The ELM applied to the sustained use of PA programs

In the ELM framework, perceived usefulness influences attitudes, and we can find support for a proposed influence on attitudes to PA programs (typical comments from young elderly, cf. Table 1: "*A very good program that makes me follow my own activity program much better*" versus "*I left the DW in mid-summer because my motivation was not sufficient to go on by myself; technology by itself is not enough as a motivation*"). Attitudes to PA programs influence the perceptions of usefulness.

We also propose a couple of moderating effects (peripheral clues) that influence attitudes to PA programs: (i) user PA history – (cf. Kettunen et al (2019)) typically identifies (i.1) regular PA users who have been active for multiple years (even decades); (i.2) sporadic PA users in on-off mode for multiple years; (i.3) inactive PA users with on-off intentions to get active; (ii) PA relevance –typically lists different goals for being active with PA programs (Linke et al (2011)), (ii.1) to lose weight and get in (much) better shape; (ii.2) to stay in sufficient shape to be independent and able to carry out all daily tasks; (ii.3) to enjoy life and (social) pleasures of pensioners (with no stress for spending time on PA); (ii.4) a multitude of other goal variations.

The ELM framework will show the drivers that get users to adopt and use PA programs (and then – always/most of the time/sometimes/seldom - to stay with the programs). There are challenges to get the users to stay with their PA programs. Linke et al (2011) cite statistics that up to 50% of people who start an exercise program drop out within 6 months; Stiggelbout et al (2005) show a dropout incidence of 0.15 per 6 months in a large 12 month program for seniors with more

than 2000 participants; Rossi et al (2018) studied the participants in a 10-year supervised physical exercise program for older adults and found that the main reasons for dropout were lack of time, disease and need to care for family members.

In section 2 we worked out the DW-app 2.5 as a digital service and noted that conceptual drivers for the adoption and use of digital services build on the Unified Theory of Acceptance and Use of Technology (UTAUT2) framework (cf. Venkatesh et al (2016)) that identifies seven drivers. We will use {DS} to denote a set of digital services that (following Venkatesh et al (2016)) will be specified by the context in which they are used.

The UTAUT2 framework will show the drivers that get users to adopt and use {DS} – which will then assist and support the sustained use of PA programs ({DS} includes DW-app 2.5 as an artefact).

We propose to combine the UTAUT2 with the ELM to work out how {DS} could contribute to the sustained use of PA programs when users are already using a program or have decided to start to use one.

If we assume that {DS} will contribute to a sustained use of PA programs we will only need three of the UTAUT2 drivers to trace the impact (ELM will cover the rest of the drivers, cf. fig.3): (i) perceived usefulness (≈ performance expectancy (PE)); (ii) perceived ease of use (≈ effort expectancy (EE)); and (iii) behavioural intention (≈ behavioural intention (BI)). The drivers are collected and described as perceived by the users (Davis (1989), Venkatesh et al (2012)). If we apply this to the {DS} we get the following constructs (cf. Table 2).

Table 2: Some selected UTAUT2 drivers for sustainable PA programs

- (PE1) I find {DS} useful in achieving my daily PA goals
- (PE2) Using {DS} helps me achieve my PA goals more quickly
- (PE3) Using {DS} increases my efficiency in achieving my PA goals
- (EE1) Learning how to use {DS} to achieve my PA goals is easy for me
- (EE2) I find using {DS} to achieve my PA goals easy
- (EE3) It is easy for me to become skilful at using {DS} to achieve my PA goals

C. Carlsson, P. Walden, T. Kari, M. Makkonen &L. Frank:
Forming Sustainable Physical Activity Programs Among Young Elderly - A Combined ELM &
UTAUT Approach

429

- (BI1) I intend to continue using {DS} to achieve my PA goals
- (BI2) I will always try to use {DS} to achieve my PA goals
- (BI3) I plan to use {DS} regularly to achieve my PA goals

Perceived usefulness now appears in both the ELM and the UTAUT2 frameworks. In the ELM framework we have the perceived usefulness of PA programs and work this out from attitudes to systematic PA, and the relevance and history of PA (for a user), which then shows what {DS} features and functions would be helpful to gain sustained use of PA programs. In the UTAUT2 framework we work out what {DS} will be useful and easy to use for sustained use of PA programs. There may be facilitating conditions for the adoption and use of the {DS}; in the UTAUT2 framework we typically find digital experience, self-efficacy, and trust (cf. Venkatesh et al (2016)).

If we combine the ELM and UTAUT2 we get the following framework (cf. fig.4) that will show the conceptual drivers that describe and (partly) explain the sustained use of PA programs. This type of combination of conceptual frameworks was used in (e.g.) Brown et al (2010) and Lallmahomed et al (2013).
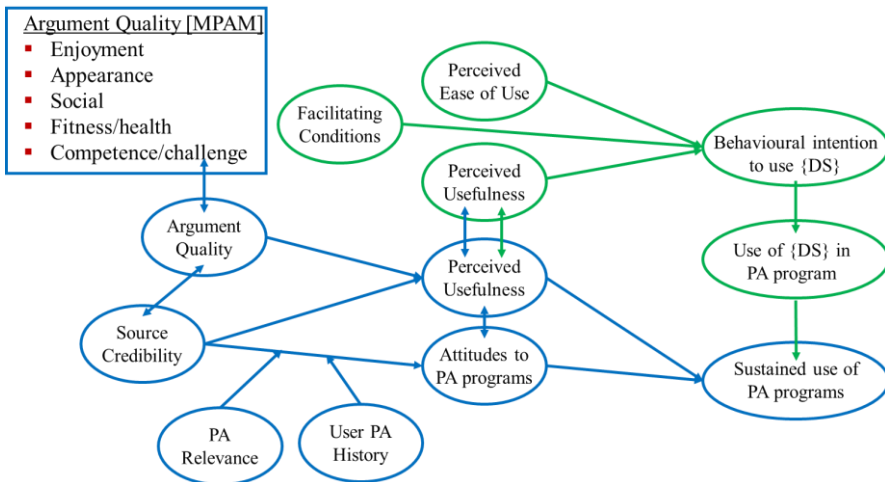
Figure 4: Combined ELM & UTAUT2 model

The framework has later been augmented with the *Motives for Physical Activity Measures* to add to the Argument Quality.

4 Preliminary Conclusion and Discussion

In section 1 we set out to tackle two research questions:

- What drivers could get young elderly users to adopt and use PA programs? Perceived usefulness & attitudes to PA programs combined with argument quality & source credibility (on the health effects of systematic PA) supported with user PA relevance and PA history will be drivers for the use of PA programs.
- What factors or drivers could help support and sustain the adoption of PA programs? Behavioural intention to use {DS} & use of {DS} in PA programs, which are supported by perceived usefulness & perceived ease of use with further support by facilitating conditions will be drivers to sustain the adoption and use of PA programs.

We realize that these answers are preliminary. Further studies are needed with the combined ELM & UTAUT2 model to find out what drivers are necessary for the sustained use of PA programs, and then what subset of drivers are sufficient to get sustained use of PA programs. We have seen in the DW research program (cf. Carlsson et al (2020)) that the {DS} will have an important (or even key) role for the sustained use of PA programs among the young elderly, but we have also seen that they will probably not be sufficient. Systematic empirical experiments and studies will give the answers.

We can use the UTAUT2 framework to enhance the proposed combined ELM & UTAUT2 model (cf. fig.5). There are *endogenous* motivations (cf. Venkatesh et al (2016)) that fit with argument quality and source credibility for sustained adoption of PA programs. There is the objective "*to get more good years*". There are facts, news, and media discussions about the effects of PA on health and the chance to avoid serious illness in senior years (Linke et al (2011)). Young elderly has short-term goals to continue with their everyday routines and longer-term goals for plans on activities that require good or better physical shape (Stiggelbout et al (2005)).

There are *exogenous* motivations (cf. Venkatesh et al (2016)) that support PA relevance in ELM; these include physicians' health recommendations for PA programs (Hukkanen et al (2018)). Potential public policy decisions on reduction in elderly care are dark drivers - "*better stay in shape, care may not be available when we need it*".

Social influence (cf. Venkatesh et al (2016)) is an exogenous factor for PA programs that supports attitudes to PA programs in ELM. The strongest influence, reasonably enough, comes from family and loved ones that want to contribute to the "*more good years*".

Aims for wellness & quality of life (cf. Venkatesh et al (2016)) also support attitudes to PA programs in ELM.

We found first validations of the constructs through empirical studies with young elderly in the DW research program (cf. Carlsson et al (2020b); Kari et al (2020a), (2020b); Makkonen et al (2020a), (2020b)) We have applied both cross-sectional (cf. Kari et al (2020a)) and longitudinal (cf. Makkonen et al (2020b)) studies of the actual use of PA programs (which fit the ELM mode), and the expected continued use of PA programs (in the UTAUT mode). The early results are interesting and growing sets of empirical data are collected in longitudinal test programs as the DW program progresses.

References

Ainsworth, B.E. et al. (2011). 2011 Compendium of Physical Activities, 0195-9131/11/4308-1575/0, Medicine & Science in Sports & Exercise, DOI: 10.1249/MSS.0b013e31821ece12

Bangsbo, J. et al. (2019). Copenhagen Consensus statement 2019: physical activity and ageing, Br J Sports Med 2019;0:1–3. doi:10.1136/bjsports-2018-100451The 2011 Compendium of Physical Activity

Bhattacherjee, A., Sanford, C. (2006). Influence Processes for Information Technology Acceptance: An Elaboration Likelihood Model, MIS Quarterly, Vol.30, No.4, 805-825

Brown, S. A., Dennis, A. R., & Venkatesh, V. (2010). Predicting collaboration technology use: Integrating technology adoption and collaboration research. Journal of Management Information Systems, 27(2), 9-53.

Carlsson, C., Walden P. (2018). Digital Wellness Services: Key to Better Quality of Life for Young Elderly, Andreja Pucihar et al. (editors), Proceedings of the 31st Bled eConference, Bled 2018, pp 248-261.

Carlsson, C., Walden, P. (2019). Digital Support to Guide Physical Activity – Augmented Daily
        Routines for Young Elderly, Andreja Pucihar et al (editors), Proceedings of the 32nd Bled
        eConference, Maribor 2019, pp 783-802.
Carlsson, C., Walden, P., Kari, T., Makkonen, M., and Frank, L. (2020a): Sustainable Physical Activity
        Programs for Young Elderly – A Fuzzy AHP Approach. A. Pucihar et al (eds.) Proceedings of
        the 33rd Bled eConference, 29.6.2020.
Carlsson, C., Kari, T., Makkonen, M., Frank, L. and Walden, P. (2020b): Sustained Adoption of
        Systematic Physical Activity Programs for Young Elderly – A Developed UTAUT Approach.
        A. Pucihar et al (eds.) Proceedings of the 33rd Bled eConference, 29.6.2020.
Davis, F.D. (1989): Perceived usefulness, perceived ease of use and user acceptance of information
        technology, MIS Quarterly 13(3), 319-339
Hukkanen, H., Husu, P., Sievänen, H., Tokola, K., Vähä-Ypyä, H., Valkeinen, H., Mäki-Opas, T., Suni,
        J.H., Vasankari, T. (2018). Aerobic physical activity assessed with accelerometer, diary,
        questionnaire, and interview in a Finnish -population sample, Scandinavian Journal of Medical
        Science in Sports, 2018;28:2196–2206, https://doi.org/10.1111/sms.13244
Jonasson, L. (2017). Aerobic Fitness and Healthy Brain Aging. Cognition, Brain Structure, and
        Dopamine, Doctoral Dissertation, Umeå University
Järveläinen J., Sell A. and Walden P., (2020). The Impact of IT Mindfulness on Complex Task
        Performance, Proceedings of the HICSS-54 Conference, pp 1232-1241
Kari, T., Rinne, P. (2018). Influence of digital coaching on physical activity: motivation and behaviour
        of physically inactive individuals. In Proceedings of the 31st Bled eConference "Digital
        Transformation – Meeting the Challenges", pp. 127-145. University of Maribor Press, Bled,
        Slovenia.
Kari, T., Makkonen, M., Frank, L., Carlsson, J. and Sell, A. (2020a): The Effects of Using a Mobile
        Wellness Application on Physical Activity Levels: A Four-Month Follow-Up Study Among
        Aged People. A. Pucihar et al (eds.) Proceedings of the 33rd Bled eConference, 29.6.2020.
Kari, T., Sell, A., Makkonen, M., Wallin, S., Walden, P., Carlsson, C., Frank, L., Carlsson, J. (2020):
        Implementing a Digital Wellness Application into Use – Challenges and Solutions among Aged
        People. Proceedings of the ITAP 2020, 6th International Conference on Human Aspects of
        IT for the Aged Population. 19-24 July 2020, Copenhagen, Denmark
Kari T., Makkonen M., Carlsson J. and Frank L. (2020b). Using a Physical Activity Application to
        Promote Physical Activity Levels Among Aged People: A Follow-Up Study. Proceedings of
        the HICSS-54 Conference, pp 1242-1251
Kettunen, E., Kari, T. Makkonen, M., Frank, L. and Critchley, W. (2020): Young Elderly and Digital
        Coaching: A Quantitative Intervention Study on Exercise Self-efficacy. A. Pucihar et al (eds.)
        Proceedings of the 33rd Bled eConference, 29.6.2020.
Kettunen E.K., (2020). Using a Digital Coach to Promote Physical Activity to University Students with
        Low Levels of Physical Activity: A Qualitative Intervention Study. Proceedings of the HICSS-
        54 Conference, pp 1252-1261
Kettunen, E., Makkonen, M., Kari, T., & Critchley, W. (2019). Using Sport and Wellness Technology
        to Promote Physical Activity: An Intervention Study among Teenagers. Proceedings of the
        52nd Hawaii International Conference on System Sciences (HICSS). 8.-11.1.2019, Grand
        Wailea, Maui. University of Hawai'i at Manoa. (pp. 1341-1350).
Kettunen, E., Critchley, W., & Kari, T. (2019). Can Digital Coaching Boost Your Performance? A
        Qualitative Study among Physically Active People. Proceedings of the 52nd Hawaii
        International Conference on System Sciences (HICSS). 8.-11.1.2019, Grand Wailea, Maui.
        University of Hawai'i at Manoa. (pp. 1331-1340).
Lallmahomed, M. Z. I., Ab Rahim, N. Z., Ibrahim, R., & Rahman, A. A. (2013). Predicting different
        conceptualizations of system use: Acceptance in hedonic volitional context (Facebook).
        Computers in Human Behavior, 29(6), 2776-2787
Linke, S.E., Gallo, L.C. and Norman, G.J. (2011), Attrition and Adherence Rates of Sustained vs.
        Intermittent Exercise Interventions. Ann Behav Med, 42(2), 197-209

Makkonen, M., Kari, T. and Frank, L. (2020a): Applying UTAUT2 to Explain the Use of Physical
    Activity Logger Applications Among Young Elderly. A. Pucihar et al (eds.) Proceedings of the
    33rd Bled eConference, 29.6.2020.

Makkonen M., Kari T. and Frank L., (2020b). Changes in the Use Intention of Digital Wellness
    Technologies and Its Antecedents Over Time: The Use of Physical Activity Logger
    Applications Among Young Elderly in Finland. Proceedings of the HICSS-54 Conference, pp
    1262-1271

Petty, R.E., Cacioppo, J.T. (1986). Communication and Persuasion: Central and Peripheral Routes to
    Attitude Change. Springer-Verlag, New York

Reyes-Mercado P. Adoption of fitness wearables. Insights from partial least squares and qualitative
    comparative analysis, Journal of Systems and Information Technology, Vol. 20 No. 1, 2018,
    pp. 103-127

Rogers, E.M. (1995). Diffusion of Innovations, Free Press, New York

Rossi, P.G., Carnaz, L., Bertollo, W.L., de Medeiros Takahasi, A.C. (2018). Causes of drop out from a
    physical exercise super-vised program specific to older adults, Fisioter Mov., 31, 1-11

Sell, A., Mezei, J. & Walden, P. (2020). Digital Coaching: A study on potential motivators. Proceedings
    of the 53rd Hawaii Conference on System Sciences (HICSS-53).

Stiggelbout, M., Hopman-Rock, M., Tak, E., Lechner, L. and van Mechelen. W. (2005). Droput from
    exercise programs for seniors: a prospective cohort study, Journal of Aging and Physical
    Activity, 13, 409-421

THL – Finnish Institute for Health and Welfare. 2019a. Aikuisten terveys-, hyvinvointi- ja
    palvelututkimus, ATH 2010-2017.

THL – Finnish Institute for Health and Welfare. 2019b. National FinHealth Study.

Wallén, M. B., Ståhle, A., Hagströmer, M., Franzén, E. & Roaldsen, K. S. (2014), Motionsvanor och
    erfarenheter av motion hos äldre vuxna, Karolinska Institutet, Stockholm, March 2014

Venkatesh, V., Brown, S. A., Maruping, L. M., Bala, H. (2008). Predicting different conceptualizations
    of system use: The competing roles of behavioral intention, facilitating conditions, and
    behavioral expectation. MIS Quarterly, 32(3), 483-502.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information
    technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly,
    36(1), 157-178.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2016). Unified Theory of Acceptance and Use of Technology:
    A synthesis and the Road Ahead. JAIS, Vol 17, Issue 5, 328-376

Yuan, S., Ma, W., Kanthawala, S., Peng, W. (2015). Keep Using My Health Apps: Discover Users'
    Perception of health and Fitness Apps with the UTAUT2 Model, Telemedicine and e-Health,
    Vol.21, No.9, 7335-741

SEGMENTATION OF THE YOUNG ELDERLY BASED ON TECHNOLOGY READINESS

ANNA SELL[1] & PIRKKO WALDEN[2]

[1] Åbo Akademi University, Faculty of Social Sciences, Business and Economics, Åbo, Finland; e-mail: anna.sell@abo.fi
[2] Institute for Advanced Management Systems Research and Åbo Akademi University, Åbo, Finland; e-mail: pirkko.walden@abo.fi

Abstract We examine the young elderly's technology readiness in order to understand the propensity to adopt and use technology for personal use. We use the Technology Readiness Index 2.0 as segmentation basis to segment a sample of mainly young elderly individuals. Our aim is to find meaningful segments within this demographic group regarding their technology readiness, and to contrast the segments with previous research. Our findings based on 538 retirees revealed a similar segmentation profile as found within working-age populations, and a surprisingly different profile than previous research with a mature target group. We identified five distinct segments portraying the young elderly as diverse technology users, ranging from 'pioneers' to 'hesitators'. The findings give arise to discussion regarding the impact of age on the technology readiness of individuals and the importance of age as a predictor of technology use. We propose that commonly held views on age as an inhibitor of technology use are becoming outdated as the diffusion of technology reaches a certain level of maturity in a market.

1   Introduction

It is often claimed that elderly people are technology averse and that a project is doomed if it introduces digital services for this group of consumers. It is not unusual that elderly people think that they are too old to use digital technology. Studies have shown (Czaja and Lee, 2007, Villarejo et al. 2019) that there are indeed those elderly who can be portrayed as technophobic i.e., who fear, dislike, or avoid technology. Studies have also shown the opposite. For example, Neves (2012) found that the elderly do not think that they are too old for digital technology nor do they see themselves as technophobic. Several studies have also shown that the elderly are willing as well as competent to use digital technology (Czaja and Lee, 2006, 2007, Sell et al. 2011, 2017).

Research on elderly people's use of technology has focused on adoption behavior, e.g., Deng et al. (2014), attitudes towards technology, e.g., Mitzner et al. (2010), evaluations of wellness-supporting applications or devices e.g., Scandurra and Sjölinder (2013), and Mercer et al. (2016). Several studies have investigated how mobile technologies can be employed to support the care of individuals with dementia-related diseases. The technology user in such studies is either the caregiver (Maiden et al. 2013; Zachos 2013) or the patient herself (Upton et al. 2011; Yamagata et al. 2013). The technology is utilized to e.g., aid communication, support reminiscence/recall and provide stimulation. These studies and mobile technologies are not as such applicable to the cognitively healthy and independent young elderly. As technology users, a digital divide seems to exist between younger and older adults, visible in lower usage of and experience with technology in general, as well as computers and the Internet (Czaja et al. 2006, sample aged 60-91 years, and König et al. 2018, sample aged 50+). The divide widens when combined with a lower education level, i.e., older adults with a lower education level are less likely to use different technologies (Czaja et al. 2006; Vroman et al. 2015; König et al. 2018). On the other hand, it has been suggested that digital technology can protect older adults from the digital divide (Hill et al. 2015).

When examining initial adoption decisions, it has been found that subjective norm and perceived behavioral control are more important for older adults than younger (Morris and Venkatesh 2000); subjective norm loses its significance with prolonged usage as domestication occurs. Heart and Kalderon (2013) followed up on the study

by Morris and Venkatesh (2000) thirteen years later (sample aged 60-90+), finding that the health status of the elderly person is a strong moderating factor regarding technology use; healthy older people are much more likely to be technology users. Nevertheless, age was still found to be a barrier to ICT use by Heart and Kalderon (2013) and based on their findings the authors concluded that older adults are not yet ready to adopt health-related ICTs. Other studies have identified facilitating conditions to be of special importance for older adults' decisions to adopt technology, such as Barnard et al. (2013) (samples aged 58-78 and over 65, mean age 68) and Nägle and Schmidt (2012) (sample aged 50-90). Marital status has also been found to be of importance; those living alone (and/or single, widowed, divorced etc.) are less likely to adopt technology. Finally, a positive and optimistic disposition is related to a higher likeliness to adopt technology (Vroman et al. 2015). In summary, research indicates that younger age, better health, higher education and not living alone are features associated with a higher likeliness to adopt and use technology. Some of these factors, i.e., age and education are also related to technology readiness (Blut and Wang 2020).

DigitalWells is an ongoing interactive research and development program with the aim to build sustainable systematic technology-supported physical activity (PA) programs for young elderly. The focus on the young elderly is a new approach for digital services, a market for which there has been little interest to develop digital value services (Bouwman et al. 2014, Carlsson and Walden 2012). The PA programs are based on self-tracking, meaning that the participants use an application on their mobile phone to log and keep track on their physical activities. It is also possible to use wearables that automatically collect the physical activities which can be synchronized to the mobile app, i.e., no manual keying is needed.

In this study we concentrate on how well prepared the young elderly are to accept and use technology in the DigitalWells program. This is carried out by using the technology readiness index (TRI) 2.0 by Parasuraman and Colby (2015) with the aim to find out technology-related beliefs. These beliefs are not easy to change within a short time frame which make them especially suitable to study.

Thus, our aim is to find meaningful segments within this demographic group with regard to their technology readiness, and to contrast the segments found in this age group with the segments found by Parasuraman and Colby (2015) in a sample

representing ages 18-65+. We assume that the young elderly differ in their technology readiness, and, if so, then it is most essential to understand the needed support of the participants in the PA-program in order for them to feel comfortable with the digital technology.

2   Technology Readiness Index 2.0

Technology readiness is defined based on Parasuraman's seminal work (2000 p. 308) as *"people's propensity to embrace and use new technologies for accomplishing goals in home life and work"*. Following Parasuraman's study we recognize that there are both positive and negative feelings that describe the domain of technology readiness. The positive feelings push people towards technologies whereas the negative may have the opposite direction, i.e., holding them back (Parasuraman 2000). These feelings Parasuraman (2000) sees as mental motivators and inhibitors that collectively determine a person's predisposition to use new technologies. The mental motivators can be categorized into two dimensions, optimism and innovativeness and the mental inhibitors as well into two dimensions, discomfort and insecurity. The dimensions are defined in the following way: *Optimism* is defined as *"a positive view of technology and a belief that it [technology] offers people increased control, flexibility, and efficiency in their lives"* (Parasuraman & Colby, 2001, p. 34). It generally portrays positive feelings about technology. *Innovativeness* is defined as *"a tendency to be a technology pioneer and thought leader"* (Parasuraman & Colby 2001, p. 36). This dimension generally measures to what degree individuals perceive themselves as being at the forefront of technology adoption. *Discomfort* is defined as *"a perceived lack of control over technology and a feeling of being overwhelmed by it"* (Parasuraman & Colby 2001, p. 41). This dimension generally measures the fear and concerns people experience when confronted with technology. *Insecurity* is defined as a *"distrust of technology and skepticism about its ability to work properly"* (Parasuraman & Colby, 2001, p. 44). This dimension focuses on concerns people may have in face of technology-based transactions.

Optimism and innovativeness are drivers of technology readiness. A high score on these dimensions will increase overall technology readiness. Discomfort and insecurity, on the other hand, are inhibitors of technology readiness. Therefore, a high score on discomfort and insecurity dimensions will reduce overall technology readiness (Parasuraman, 2000).

Parasuraman and Colby (2001) emphasize that the four dimensions are fairly independent, indicating that each of them make a unique contribution to an individual's technology readiness. An individual can for example score high on motivations simultaneously mitigated by strong inhibitions which would be considered as a *"paradoxical state"* (Parasuraman and Colby, 2001 p. 61).

3 Segmentation studies based on technology readiness

It is suggested that using attitudes and lifestyles as segmentation bases can be useful when segmenting technology users (Sell and Walden 2012, Sell et al. 2014). Technology readiness represents an individual's technology-related beliefs, and it has been used as a basis for segmentation in several research projects. Parasuraman and Colby (2001) originally found five segments which they named explorers (high motivation, low inhibition), pioneers (high motivation, high inhibition), skeptics (low motivation, low inhibition), paranoids (moderate motivation, high inhibition) and laggards (low motivation, high inhibition). Using another method (latent class analysis) and a streamlined 16-item scale for segmenting the consumers Parasuraman and Colby (2015) again found a five-cluster solution being the best and similar to the prior TR-based segmentation (2001). The segments were labeled and described as follows: *(i) skeptics, a detached view of technology, with less extreme positive and negative beliefs; (ii) explorers, a high degree of motivation and low degree of resistance; (iii) avoiders, a high degree of resistance and a low degree of motivation; (iv) pioneers, both strong positive and negative views about technology; and (v) hesitators, a low degree of innovativeness* (Parasuraman and Colby 2015, p.71).

Tsikriktsis (2004) replicated and extended the taxonomy proposed in 2001 by Parasuraman and Colby. He found that there are both similarities and differences between the two segmentation studies. Four segments had a good match but the fifth segment – paranoids – could not be identified in his study. Victorino et al. (2009) explored the use of a ten-item abbreviated version of TRI (Parasuraman and Colby 2001) for hotel customer segmentation and found that it was a reliable method for segmenting customers. A three-cluster solution came out as the best. The segments they found were similar to the explorers/pioneers, paranoids, and laggards.

One of the few studies researching the TRI-profile of mature consumers is by Rose and Fogarty (2010). They use the abbreviated technology readiness scale by Parasuraman and Colby (2001) in order to segment Australian consumers over 50 years of age. They found support for the original five segments by Parasuraman and Colby (2001). Mature consumers in their study were more likely to belong to the skeptic and laggard segments (34% and 23% respectively), and less likely to belong to the explorer and pioneer segments (16% and 13%) than the general population sample their results were contrasted to (22% skeptics, 17% laggards, 19% explorers and 26% pioneers). Demographic analysis of segments revealed that older, female, and less educated individuals were more likely to be in the laggard segment, mirroring previous research (Czaja et al 2006, König et al 2018).

Kim et al. (2018) used the TRI 2.0 to segment users of sports wearables and found three distinct groups of users, explorers, laggards and pioneers. The method they used was a two-stage cluster analysis in contrast to latent class analysis which was used by Parasuraman and Colby (2015). In a recent article Wiese and Humbani (2020) applied TRI 2.0 for segmenting the mobile payment market in South Africa. They found four segments of which three shared similarities with pioneers, paranoids and explorers and the fourth - hesitant-skeptics - had no similarities with the original ones. Ramirez-Correa et al. (2020) validated the TRI 2.0 in Chile and used LCA for segmentation. They found that four segments are similar compared to the study by Parasuraman and Colby (2015) whereas the segment of skeptics is different. There are also differences in the size of the segments as hesitators and pioneers together count for more than 80% of the Chilean sample.

All the above-mentioned studies have both similarities and differences. They are targeting different countries with different technology adoption stages and different contexts that most probably explain some of the variation in the outcomes.

4   The study

Our study is carried out in Finland which repeatedly has been named as one of the most technologically advanced countries in the world. Further, according to the European Commission (2020), Finland is leading the EU in digital competence and use of internet services amongst its population. We are targeting a group of people which we call the young elderly, 60 to 75 years of age. There are few studies that

focus on technology readiness amongst elderly or young elderly and technology readiness-based segmentation, thus our study is to our knowledge one of the first of its kind.

Our study is divided into several waves, depending on when the participants joined the DigitalWells research program. In the first wave which we report here are 575 participants who joined the program in 2019-2020. All participants are retired, they are members of a local retiree's association, they have volunteered to be in the program and a prerequisite is that they have a smart phone, i.e., they need to have a mobile internet connection in order to be able to participate in the program. When we first start working with the participants, we do not require that they have deep knowledge in digital technology as we give them hands-on guidance. However, it is vital for the program to measure their readiness for technology and for that purpose we used the Technology Readiness Index 2.0 by Parasuraman and Colby (2015). The TRI 2.0 questions were translated to Finnish and Swedish by a team of researchers and then translated back to English by another team of researchers. We aim at having the participant feel confident about and comfortable with the digital technology used in the project. As pointed out by Parasuraman and Colby (2015, p. 61) technology readiness is *"…an individual-level characteristic that does not vary in the short term…"* and another notable aspect is that it does not change suddenly in response to a stimulus. Several researchers have also found that higher technology readiness levels are correlated with higher adoption rates of cutting-edge technologies, more intense use of technology, and greater perceived ease in doing so (Kuo 2011, Fisk et al. 2011, Massey et al. 2007) which would all contribute to the success of the DigitalWells program.

The demographics of the participants are as follows. 62% of the participants are women and the mean age of the participants is 68,2 years (range 48-83[1]), median value being 68 years (n=528). The great majority of the participants (68,5%) are married, 10,1% are divorced, 8,2% are widowed, 9,7% are common-law married and the remaining 3% are single. Many of the participants (67,2%) live in a small city (less than 20.000 inhabitants) or in a rural area. The mother tongue of the participants is Finnish 82% or Swedish 18%. Roughly one-fourth of the respondents

---

[1] In our sample, a few of the respondents were younger or older than the young elderly age range of 60-75. However, all respondents are retirees and belong to a retired persons' association.

(25,3%) have a university degree or a degree from university of applied sciences, but there are a lot of missing values concerning education (n=226).

At the onset, we had 575 observations. After cleaning out respondents who (i) had no answers on one of the four TRI dimensions, (ii) had only answered eight or less of the TRI statements, or (iii) only had one answer on one or more of the TRI dimensions, we ended up with 538 usable observations. Remaining missing values were replaced with mean value for the sample.

5   Results

We calculated the Technology Readiness Index and its components following Parasuraman and Colby (2015), table 1 below. For comparison, we include the mean values for the TRI components in the Parasuraman and Colby study (*P&C* in table 1). The scale ranges from strongly agree 5.0 to strongly disagree 1.0. The overall technology readiness index mean score is 3.01, close to the scale's midpoint 3.0. The participants are generally optimistic (OPT) about technology, mean score being 3.42. Innovativeness (INN) score is below midpoint. The participants are on discomfort (DISC) below the midpoint and insecurity (INSEC) above the midpoint. Distribution of the scores is near-normal, as skewness values are all between -0,5 and 0,5 and kurtosis values between -1 and 1. As expected, correlation between motivating factors (optimism and innovativeness) is positive and correlation between inhibiting factors (discomfort and insecurity) is also positive. Correlations for motivator-inhibitor combinations are negative.

Table 1: Summary statistics for Technology Readiness Index 2.0 and its components

	Mean values				Correlation coefficients			
Components	*Current study*	*P&C*	*SD*	*Skewness*	*Kurtosis*	*OPT*	*INN*	*DISC*	*INS*
OPT	3,42	3,75	0,80	-0,41	-0,06	1,00			
INN	2,89	3,02	0,94	0,05	-0,76	0,53	1,00		
DISC	2,88	3,09	0,85	-0,03	-0,64	-0,39	-0,45	1,00	
INSEC	3,41	3,58	0,91	-0,37	-0,25	-0,36	-0,31	0,53	1,00
Overall TR score	3,01	3,02	0,66	0,10	-0,42	0,74	0,77	-0,78	-0,73

We employed exploratory factor analysis with Varimax rotation to examine whether the factor structure of the original four TRI dimensions can be found in this dataset. We utilized principal axis factoring with Varimax rotation requesting a four-factor solution. After examining initial factor loadings, we decided to remove one of the variables on the Insecurity scale (INSEC4) from further analysis, as it did not load correctly on the Insecurity dimension and its communality was below 0.5. Our subsequent factor analysis (without the INSEC4 item) explains 59.2% of the variance. The eigenvalues of the first three factors surpass 1, the eigenvalue of the fourth one being 0.93. Examining a scree plot confirmed the suitability of a four-factor solution. Also, each of the four factors explains a significant part of the variance: 16.7%, 16.1%, 14.1% and 12.2% respectively. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.891 and Bartlett's test was significant ($x^2_{(105)}$ = 2324.0, $p < 0.000$).

The four-factor solution aligned well with the four technology readiness dimensions, with only one item (DISC2) with a standardized factor loading just below 0.5 at 0.475. Remaining items loaded from 0.584 to 0.771. There is one significant cross-loading, with INN4 cross-loading on the Optimism dimension. The same cross-loading is reported by Parasuraman & Colby (2015) in their validation of the TRI 2.0 scale. Factor loadings and Cronbach's α for each of the sub-scales can be seen in table 2.

Table 2: Factor loadings and reliability analysis

	Factor 1	Factor 2	Factor 3	Factor 4	Cronbach's α[a]
OPT 1	.771				
OPT 2	.598				.756
OPT 3	.715				
OPT 4	.741				
INN 1		.763			
INN 2		.699			.770
INN 3		.749			
INN 4	*.426*	.596			
DISC 1			.727		
DISC 2			.475		.708
DISC 3			.584		
DISC 4			.702		
INSEC 1				.767	
INSEC 2				.710	.725
INSEC 3				.769	

a.   For reliability analysis, discomfort and insecurity statements were reverse-coded.

At the following stage, K-means cluster analysis in SPSS 24 was used to create a five-cluster solution. The five resulting segments were analyzed and found to be similar to the five segments found by Parasuraman and Colby (2015) in a sample of 878 respondents with a median age of 51 years and ranging from 18 to 65+ (exact upper range not reported). Hence, we have named the clusters with the same monikers as Parasuraman and Colby. A description of the segments can be found in chapter 3.

In the following we present the segments and highlight similarities and differences to those presented by Parasuraman and Colby (2015). In table 3, the segments are described through the mean values of each segment for the four technology readiness dimensions, as well as the TR score for each segment, the size of the segment and the rank of the segment based on the TR score (TR scores were calculated according to Parasuraman and Colby 2015). TR score, percentage per cluster and rank is provided also for the Parasuraman and Colby (2015) sample for the sake of comparison.

Table 3: Mean values, TR scores, size and rank

									Parasuraman and Colby (2015)		
Clusters	*OPT*	*INN*	*DISC*	*INSEC*	*N*	*TR SCORE*	*%*	*RANK*	*TR SCORE*	*%*	*Rank*
Pioneers	3,75	3,37	3,12	3,63	137	3,09	25 %	3	3,05	16 %	3
Skeptics	3,16	2,49	2,35	2,77	83	3,13	15 %	2	3,06	38 %	2
Hesitators	3,04	1,78	3,53	3,88	117	2,35	22 %	4	2,74	13 %	4
Avoiders	2,59	2,54	3,49	4,27	80	2,34	15 %	5	2,13	16 %	5
Explorers	4,16	3,94	1,92	2,57	121	3,90	22 %	1	3,92	18 %	1

In our sample, the *Pioneers* is the largest segment, comprising 25% of the total respondents, compared to 16% in the Parasuraman and Colby (P&C) study. In P&C, *Skeptics* are the largest segment with 38% of total respondents, whereas in our study, only 15% of respondents fall into the *Skeptics* segment. Both *Avoiders* and *Explorers* are similarly sized segments in the two studies, but a clearly larger proportion of our respondents fall into the *Hesitators* segment (22%) than in the P&C study (13%). When looking at the ranking of the segments, based on the mean TR scores, the two

studies exhibit the same ranking, with *Explorers* showing the highest mean TR score and the *Avoiders* the lowest. Interestingly, the mean TR scores are higher for the *Pioneers, Skeptics* and *Avoiders* in our study, higher for *Hesitators* in P&C, and roughly the same in both studies for *Explorers.*

When looking at the segments through the lens of the four technology readiness dimensions, some interesting observations can be made. We visualize the segments from our study in figure 1. The Pioneer, Skeptic and Explorer segments present similar profiles in our study and in that by P&C. The profiles for Pioneers and Skeptics are very similar on all dimensions, but the scores are lower for the Skeptics in our study. The Hesitators, characterized by their low level of innovativeness, present a distinctly higher level of optimism in the P&C study than in our study. The Avoiders show a markedly lower level of innovativeness in the P&C study than in our study.
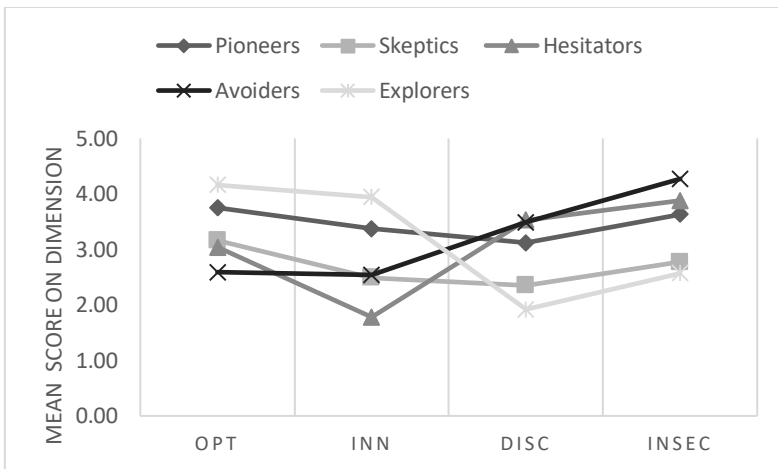
Figure 1: Mean values for technology readiness dimensions per segment in current study (n=538)

In table 4 we outline the demographic characteristics of the five TRI-based segments derived from our data. Age is excluded, as there were only negligible differences in age between the segments. Due to missing data, it was not possible to analyze the segments according to their level of education. We notice that the biggest segments are Pioneers, Explorers and Hesitators. Contrasted to the age division in the entire sample (68% female), we can see that women are underrepresented in the Pioneer and Explorer segments. *Pioneers* and *Explorers* as segments are also characterized by

a proportionally high proportion of city-dwellers. In the *Skeptics* we find a relatively high proportion of women and a somewhat high proportion of persons who belong to the Swedish-speaking minority. *Hesitators* are predominantly female and have a higher percentage of individuals living alone.

Table 4: Demographic characteristics of the segments

	Female (%)	Living alone (%)	Large or medium city (%)	Swedish minority (%)
Pioneers (137)	55,6	20,9	**38,8**	*14,6*
Skeptics (83)	**69,9**	*17,5*	29,6	**22,9**
Hesitators (117)	**70,1**	**28,7**	30,2	*14,5*
Avoiders (80)	65,4	23,4	28,6	18,8
Explorers (121)	*52,5*	22,2	33,3	21,5

6   Discussion and conclusions

The aim of this study was to find meaningful segments within the young elderly cohort with regard to their technology readiness, and to contrast the segments found in this age group with the segments found by Parasuraman and Colby (2015) in a sample representing ages 18-65+. We underline the fact that Parasuraman and Colby collected their data in year 2012 whereas our data is collected in 2019-2020.

In this study we used the 16-item scale and could validate the presence of all four technology readiness dimensions in the dataset. We successfully identified five segments similar to the segments found by Parasuraman and Colby. This is noteworthy as our data covers the young elderly cohort; thus, this supports the allegation that the mature consumer market is not homogeneous and suggests that the mature technology consumers are similar to working age technology consumers regarding their technology readiness. However, the size of the segments differs. The Pioneers is the biggest segment in our study compared to the Skeptics in the Parasuraman and Colby study. The smallest group found by Parasuraman and Colby was the Hesitators, whereas the proportion of Hesitators in our material was higher. Previous research suggests that the level of innovativeness in information technology is overall lower for mature women, which might explain this finding. A plausible explanation for the higher proportion of Pioneers in our study could be

derived from the fact that Finns are frequent users of the internet. In year 2020, the proportion of internet users in the 16-89 years age group was 92% and of those 82% used the internet several times per day. The proportion of internet users among those aged 65-74 is 88% and of those 62% use the internet several times per day (Statistics Finland 2020).

The ranking of the segments was the same in both studies. When looking at the segment profiles, the Pioneer, Skeptic and Explorer segments are very similar. The Hesitators in both studies show a low level of innovativeness, but in our study also a markedly lower level of optimism. Conversely, the Avoiders in our study exhibit a higher level of innovativeness than in the Parasuraman and Colby (2015) study. Overall, the segments are surprisingly similar in the two studies; we would have expected the segments in the young elderly sample to show at least partially different profiles.

Apart from comparing our findings to Parasuraman and Colby (2015), we also contrast the results with a study by Rose and Fogarty (2010) as their study was specifically focused on mature consumers in Australia and also utilized a TRI-based segmentation. We observe some similarities and differences. Both studies found a five-segment solution, but the segments differ content wise and in size. Mature Australian consumers are contrary to our findings less likely to be explorers or pioneers (29,5%) but adopters at late growth or decline stage (57.7%). Skeptics and Laggards (equals to Avoiders) are the biggest segments in the Australian study and the smallest in our study. The results from these two studies on mature technology consumers differ significantly which warrants discussion. A partial explanation can be the differences in the technology development/infrastructure, as the 'second generation' mobile phone systems were introduced in 1990s and achieved early on a high penetration rate in the Finnish working age population, meaning that the Finnish young elderly in our study were technology users long before turning 60 years old. Also, the Australian study was done more than a decade ago; the mature technology consumer market changes rapidly as the proportion of technologically experienced individuals entering retirement constantly rises. Rose and Fogarty found that the Explorers and Pioneers are younger whereas the Laggards are older. This is a notable difference as age did not describe the segments we found. We propose that the importance of age as a predictor of technology use gradually loses its significance as the diffusion of technology reaches a certain level of maturity in a market.

The findings from our study support the four technology readiness dimensions and the five-segment solution. It provides interesting profiles of young elderly and a unique knowledge of this age group's technology beliefs in the different segments.

Note. The Technology Readiness Index 2.0 survey research scale is copyrighted by A. Parasuraman and Rockbridge Associates, Inc., 1999, and is used with written permission. TRI items from Parasuraman (2015) were translated into Finnish and Swedish.

References

Barnard Y, Bradley MD, Hogdson F, Lloyd AD (2013). Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. Computers in Human Behavior 29(4), 1715-1724.

Blut, M., Wang, C. (2020). Technology readiness: a meta-analysis of conceptualizations of the construct and its impact on technology usage. Journal of the Academy of Marketing Science, 48, 649–669.

Bouwman, H., Carlsson, C., de Reuver, M., Hampe, F. and Walden, P. (2014). Mobile R&D Prototypes – What is Hampering Market Implementation. International Journal of Innovation and Technology Management, 11(1).

Carlsson, C. and Walden, P. (2012). From MCOM Visions to Mobile Value Services. Roger Clarke, Andreja Puchar and Joze Gricar (eds.), The First 25 Years of the Bled eConference, University of Maribor, Bled, 69-91.

Czaja, S.J., Charness, N., Fisk, A.D., Hertzog, C., Nair, S.N., Rogers, W.A. and J. Sharit, (2006). Factors Predicting the Use of Technology: Findings From the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychology and Aging, 21:2, 333-352.

Czaja S. and Lee, C.C. (2007). The impact of aging on access to technology. Universal Access in the Information Society (UAIS), 5(4), 341-349.

Deng, Z., Mo, X. and S. Liu (2014). Comparison of the middle-aged and older users' adoption of mobile health services in China. International Journal of Medical Informatics, 83, 210-224.

European Commission (2020). Digital Economy and Society Index (DESI) 2020. https://ec.europa.eu/digital-single-market/en/digital-economy-and-society-index-desi

Fisk, R. P., Patricio, L., Lin, J. S. C., & Chang, H. C. (2011). The role of technology readiness in self-service technology acceptance. Managing Service Quality: An International Journal, 21(4), 424-444.

Heart, T., and E. Kalderon (2013). Older adults: Are they ready to adopt health-related ICT? International Journal of Medical Informatics, 82(11), e209-e231.

Hill, R., L.R. Betts, and S.E. Gardner (2015). Older adults' experiences and perceptions of digital technology: (Dis)empowerment, wellbeing, and inclusion. Computers in Human Behavior, 48, 415-423.

Kim, T., Chiu, W., & Chow, K. (2019). Sport technology consumers: segmenting users of sports wearable devices based on technology readiness. Sport, Business and Management, 9(2), 134-145.

Kim, M. J., Lee, C. K., & Preis, M. W. (2020). The impact of innovation and gratification on authentic experience, subjective well-being, and behavioral intention in tourism virtual reality: The moderating role of technology readiness. Telematics and Informatics, 49, 101349.

Kuo, T. H. (2011). The antecedents of customer relationship in e-banking industry. Journal of Computer Information Systems, 51(3), 57-66.

König, R., Seifert, A., & Doh, M. (2018). Internet use among older Europeans: an analysis based on SHARE data. Universal Access in the Information Society, 17(3), 621-633.

Maiden, N., D'Souza, S., Jones, S., Müller, L., Pannese, L., Pitts, K., ... & Zachos, K. (2013). Computing technologies for reflective, creative care of people with dementia. Communications of the ACM, 56(11), 60-67.

Massey, A. P., Khatri, V., & Montoya-Weiss, M. M. (2007). Usability of online services: The role of technology readiness and context. Decision Sciences, 38(2), 277-308.

Mercer, K., Giangregorio, L., Schneider, E., Chilana, P., Li, M. and Grindrod, K. (2016). Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR mHealth and uHealth, 4(1).

Mitzner, T. L., Boron, J. B., Fausset, C. B., Adams, A. E., Charness, N., Czaja, S. J., ... & Sharit, J. (2010). Older adults talk technology: Technology usage and attitudes. Computers in Human Behavior, 26(6), 1710-1721.

Morris, M. G., & Venkatesh, V. (2000). Age differences in technology adoption decisions: Implications for a changing work force. Personnel psychology, 53(2), 375-403.

Neves, B. B., & Amaro, F. (2012). Too old for technology? How the elderly of Lisbon use and perceive ICT. The Journal of Community Informatics, 8(1), 1-12.

Nägle, S., & Schmidt, L. (2012). Computer acceptance of older adults. Work, 41(Supplement 1), 3541-3548.

Parasuraman, A. (2000). Technology Readiness Index (Tri): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies. Journal of Service Research, 2(4), 307–320.

Parasuraman, A. and C.L. Colby (2001). Techno-ready marketing: How and why your customers adopt technology. New York: Free Press.

Parasuraman, A., & Colby, C. L. (2015). An updated and streamlined technology readiness index: TRI 2.0. Journal of service research, 18(1), 59-74.

Ramírez-Correa, P., Grandón, E. E., & Rondán-Cataluña, F. J. (2020). Users segmentation based on the Technological Readiness Adoption Index in emerging countries: The case of Chile. Technological Forecasting and Social Change, 155, 120035.

Rose, J., & Fogarty, G. J. (2010). Technology readiness and segmentation profile of mature consumers. In Proceedings of the 4th Biennial Conference of the Academy of World Business, Marketing and Management Development, 4(1), 57-65.

Scandurra, I., & Sjölinder, M. (2013). Participatory design with seniors: Design of future services and iterative refinements of interactive eHealth services for old citizens. Medicine 2.0, 2(2).

Sell A., Walden P. and Carlsson C. (2011). Segmentation Matters – An Exploratory Study of Mobile Service Users. International Journal of Systems and Service-Oriented Engineering, 2(3), 1-17.

Sell A. and Walden P. (2012). Segmentation bases in the mobile services market: Attitudes in, demographics out. The Forty-Fifth Annual Hawaii International Conference on System Sciences, IEEE Computer Society Press, Los Alamitos.

Sell A., Mezei J. and Walden P. (2014). An attitude-based latent class segmentation analysis of mobile phone users. Telematics & Informatics, 31(2), 209-219.

Sell, A., Walden, C. and Walden P., (2017). My Wellness as a Mobile App. Identifying Wellness Types among the Young Elderly. The Fiftieth Annual Hawaii International Conference on System Sciences. IEEE Computer Society Press, Los Alamitos.

Statistics Finland (2021). Use of information and communication technology by individuals. https://www.stat.fi/til/sutivi/index_en.html

Tsikriktsis, N. (2004). A technology readiness-based taxonomy of customers: A replication and extension. Journal of Service Research, 7(1), 42-52.

Upton, D., Upton, P., Jones, T., Jutlla, K., Brooker, D. and Grove, H. (2011). Evaluation of the impact of touch screen technology on people with dementia and their careers within care home settings. UK: University of Worcester

Victorino, L., Karniouchina, E., & Verma, R. (2009). Exploring the use of the abbreviated technology readiness index for hotel customer segmentation. Cornell Hospitality Quarterly, 50(3), 342-359.

Villarejo Ramos, Á. F., Peral Peral, B., & Arenas Gaitán, J. (2019). Latent segmentation of older adults in the use of social networks and e-banking services. Information Research, 24(4).

Vroman, K. G., Arthanat, S., & Lysack, C. (2015). "Who over 65 is online?" Older adults' dispositions toward information communication technology. Computers in Human Behavior, 43, 156-166.

Wiese, M., & Humbani, M. (2020). Exploring technology readiness for mobile payment app users. The International Review of Retail, Distribution and Consumer Research, 30(2), 123-142.

Yamagata, C., Coppola, J. F., Kowtko, M., & Joyce, S. (2013). Mobile app development and usability research to help dementia and Alzheimer patients. In 2013 IEEE Long Island Systems, Applications and Technology Conference (LISAT), 1-6.

Zachos, K., Maiden, N., Pitts, K., Jones, S., Turner, I., Rose, M., ... & MacManus, J. (2013). A software app to support creativity in dementia care. In Proceedings of the 9th ACM Conference on Creativity & Cognition

THE SHAPE OF BOTTOM-UP URBANISM PARTICIPATORY PLATFORMS: A CONCEPTUALISATION AND EMPIRICAL STUDY

PASCAL ABEL,[1] DENNIS MIETHER,[1] FLORIAN PLÖTZKY[2] &
SUSANNE ROBRA-BISSANTZ[1]

[1] Technische Universität Braunschweig, Chair of Information Management,
Braunschweig, Germany; e-mail: p.abel@tu-braunschweig.de,
d.miether@tu-braunschweig.de, s.robra-bissantz@tu-braunschweig.de
[2] Technische Universität Braunschweig, Institute for Information Systems,
Braunschweig, Germany; e-mail: ploetzky@ifis.cs.tu-bs.de

Abstract Citizens around the world are changing their urban environment through bottom-up projects. They are increasingly using digital platforms to come together. From the perspective of smart city research, this form of participation and interaction with city administrations has not yet been researched and defined. In our study we suggest a conceptualisation of bottom-up urbanism participatory platforms and analysed 143 platforms. We identified 23 platforms as our study sample. They vary in their focus from implementation to funding or discussion. Therefor we found a broad range of participation mechanisms. A wide range of employment or voluntary work of staff members was shown. A heterogeneous picture also emerged regarding other characteristics (e.g. funding size, users or number of projects). One thing they have in common is their good cooperation with cities and regional actors.

1   Introduction

In the notion of smart city initiatives local governments have an increasing interest in more citizen-centric approaches for future cities (Cardullo & Kitchin, 2019). They form "smart", new ways to tackle the challenges of the future. Those lay within the facets of economy, people, governance, mobility, environment and living (Lombardi, Giordano, Farouh, & Yousef, 2012). In order to adapt, local governments evolve with a growing importance for new governance strategies where the cities reflect their processes and role understanding. Gil-Garcia et al. (2020) address how governments of smart cities interact with their citizens and identify the dimensions of information availability, transparency, participation, collaboration and information technologies. Through participation the local governments perspective changes as does the citizens perspective. In Foth's (2017) Cities 4.0 concept the governments change from administrator to collaborator and citizens from residents to co-creators. From this point of view we see a collaborative approach with both partners meeting in a new "middle". It is a shift from designing for the citizens to giving them the right to change or even reinvent their urban environment.

Under the umbrella term of bottom-up urbanism citizens revive an unused building into a community cinema, organize street festivals or transform a fridge into a book-sharing shelve. Those and other activities are carried out by citizens from Detroit to Paris and from Christchurch to Vienna. Those initiatives are seen as a driver for urban innovation (Caragliu, Del Bo, & Nijkamp, 2011) by building an experimental environment (Anttiroiko, 2016). In this environment the cities can adapt through the actions of citizens in response of ongoing changes in society (Silva, 2016).

Since bottom-up urbanism is seen as an alternative to the top-down approach of planned environments (de Waal & de Lange, 2019) the role of city planners is changing: where planners previously developed projects for urban space, now the development of digital platforms for the engagement of the citizens is becoming a central task (Ertiö & Bhagwatwar, 2017). Those platforms differ in their functionalities (e.g. post ideas, discuss topics) and offer a broad range of participation levels (Falco & Kleinhans, 2018). Senbel & Church (2011) proposed a broadly used concept to distinguish participation levels on digital platforms.

Participation on the lowest level in this regard can be seen as simply being provided with information. The middle layers of the model by Senbel & Church (2011) allow citizens to participate by contribution of ideas and by getting inspiration for instance by using polls or inviting citizens directly for their opinions on a certain issue. The higher levels finally allow citizens to join the planning and design process along with the possibility of creating their own neighbourhood plans. However, the model misses citizen control (highest level) as proposed by Arnstein (1969), which is described as self-governance in smart city research (Zhilin, Klievink, & de Jong, 2019).

The lower, middle and higher level had been broadly researched (Desouza & Bhagwatwar, 2014; Ertiö & Bhagwatwar, 2017; Falco & Kleinhans, 2018; Gün, Demir, & Pak, 2019) but the implementation of platforms supporting higher levels of participation did not fulfil the users needs yet, as reported by empirical studies (Gün, Demir, & Pak, 2019). Mostly because the engaging mechanisms have not yet developed (Ertiö, 2015) and practitioners often fail to further improve and provide the funding for their platforms (Abel, Stuwe, & Robra-Bissantz, 2019). That leads us to the demand to further investigate and understand platforms of the highest level of participation with a focus on platforms supporting citizen projects. Researchers distinguish the importance of self-governance as a part of the broad concept of smart cities (Zhilin et al., 2019), which will be discussed in section 2.1 in detail. However, there is a lack of concrete concepts regarding self-governance in this context and the differentiation towards other concepts remains unclear (Rauws, 2016). That causes planners and cities to undervalue self-governance. The outcome of this study focuses on the research question:

How are bottom-up urbanism participatory platforms conceptualized?

In part A of this paper we derive a definition of bottom-up urbanism participatory platforms. Which, in part B, is evaluated in the field by analyzing 143 platforms. We then provide a detailed view on 26 platforms identified as bottom-up urbanism participatory platforms.

2   Conceptualisation

The percentage of people living in cities is growing and growing (Statista, 2020). Hence it is necessary to think about the future of cities which should be built to fulfill citizens needs and enable them to participate in their environment. In smart city concepts Gil-Garcia et al. (2020) identified interaction dimensions and Nam & Pardo (2011) offer three main components that seem to be at the core of it: technology factors, human factors and institutional factors. The connection and interplay of those factors is where investment into smart cities enhances quality of life and provides sustainable growth (Caragliu et al., 2011).

Zhilin et al. (2019) sees smart cities as an onion where the layers are connected and build apon each other. In our conceptualisation of bottom-up urbanism participatory platforms we are going to describe those layers in the following sections and bring them together in form of a definition in the last section (2.5).

2.1   Future cities and the public discourse

Future cities are often discussed as smart cities in the public and academic discourse. Oftentimes the definition is limited to technical solutionism like the smart city being a collection of services and the consumption of internet technologies (Walser & Haller, 2016). This focus is also described as the "Control Room" vision of a smart city where the focus of a city is laid on central optimization and the city as a service (de Waal & Dignum, 2017). However, the result of this de-subjectivism of citizens most likely leads to less participation because the only role for citizens is to be data provider for companies selling technology-centered smart city solutions (Keymolen & Voorwinden, 2020).

De Waal & Dignum (2017) also envision "Smart Citizens". The latter being a counter-argument regarding the Control Room vision described before. In it, citizens and civic organizations use digital technologies to mobilize themself, act together and claim self-governance (de Waal & Dignum, 2017). Additionally, the municipality uses digital technologies to optimize their citizen centered processes but is still the main regulator in the city. On the one hand we have citizen engagement (e.g. bottom-up urbanism) and on the other hand the municipality setting the legislative frame.

The concept of smart cities differs from concrete to vague ideas and can be better understood as "urban imaginaries" with a set of "visions, hopes, and fears – rational or irrational, fact based or emotionally appealing" (de Waal, 2011). Rather than having a clear agenda to reach a specific "what", smart cities evolve as "working arrangements" (Cowley, Joss, & Dayot, 2018) with multiple processes that need to be under constant evaluation and adapted if needed. A smart city is better seen as discourse with constant change that "may become the new 'normal'" and lead to "new rules and routines, in laws, in new business models, in new roles for actors, and even in newly shared values" (Hajer, 2016). However, this requires the realization of concrete projects rather than vague discussions of possibilities (Schinkel, Jain, & Schröder, 2014). In this regard, citizens all over the world already find new ways to take part in the discourse and prototype their understating of future by changing their surroundings.

2.2   Participation: When citizens really take their part

With Arnstein's (1969) Ladder of Participation a formulation to more power in urban planning for citizens began. She imagined a society that is more equal and saw the path to success by participating and transfering power to the citizens (Cardullo & Kitchin, 2019). At the same time Lefebvre (1968) criticized the development of cities with capitalism interests under control of the government and proclaimed the "right to the city" as a self-determined space for citizens.

In the notion of smart cities the self-determined space shifted towards the question of governance or who has to decide? Decision-making is traditionally lying in the hand of public actors but it is debated how a policy process is organised and how non-governmental actors such as citizens are involved.

Kooiman (2003) structures governance modes in hierarchical governance, self-governance and co-governance. The mode self-governance sets the non-governmental actors in the center and the government to the side. It can be divided in terms of actors, powers and rules (Arnouts, van der Zouwen, & Arts, 2012).

The Actors are citizens that actively participate to achieve a common goal, have the power in decision making and the form of internal coordination they choose (Zhilin et al., 2019). They operate in a high degree of freedom but are guided by rules of their own and by the government.

Acting within this mode of self-governance the individual intentions are transformed into a collective intent (Rauws, 2016). This transformation can be better seen as an ongoing process than a status quo. As the actors question and transform the urban environment the shape of the governance system itself is always questioned and iterated by non-governmental actors and the government.

Where Zhilin et al. (2019) sees self-governance as a top-down approach where the government empowers citizens, we argue that self-governance can as well arise from the bottom-up as a demand of the right to the city. It is emerging as an interplay from top-down and bottom-up approaches where both sides reimage their rules and roles on their way to a more equal future city.

2.3   Acting on the streets from the bottom-up towards big change

In recent years bottom-up urbanism became the umbrella term (Douglas, 2019) for several views of the transformation of the public space by citizens (Fabian & Samson, 2016). The focus of "bottom-up" represents the origin of the initiatives within the citizenry and the mode of self-governance of the actors (Kickert & Arefi, 2019).

The activities of citizens to transform the public space symbolise the difference between the city as a planned environment and as a lived place (Crawford, 2008). Citizens aim to solve unadressed problems (Finn, 2014) in a way of incremental improvements at smaller scale (Talen, 2015). In a do-it-yourself (DIY) manner they build projects and are seen as amateur designers which delimits their actions from planned urbanism (Iveson, 2013).

The outcome of those projects are very different and we find no project like another. They inhabit several perspectives and vary in their goals (Kickert & Arefi, 2019). That offers a contribution to the public discourse and planning processes. Since bottom-up urbanism is "a radical repositioning of the designer, a shifting of power

from the professional expert to the ordinary person" (Crawford, 2008) there are opportunities for planners to learn from citizen's projects. This contribution is inhabitant in the perspective of Tactical Urbanism with its mantra "short-term action for long-term change" (Lydon & Garcia, 2015). It is seen as a way to provide new insights of citizens through their activities and clarify the meanings by providing physical evidence (Silva, 2016). Even though their concrete projects are often of temporary nature. But like the transformation of parking lots with immediate results and the scope towards bigger change, it is the mentality to prototype an object which is transformed and tested while used to become a symbol for a future vision. A vision not of a concrete spatial situation but an opening for like-minded ideas.

2.4   Urban participation on participatory platforms

As shown before we have a good understanding of the governance mode and the activities of bottom-up urbanism. There has been serveral studies that show the usage of technology in participatory processes (Desouza & Bhagwatwar, 2014; Ertiö & Bhagwatwar, 2017; Falco & Kleinhans, 2018; Gün et al., 2019; Senbel & Church, 2011). But there has been no focus on the self-governance level on participatory platforms. In their work Desouza & Bhagwatwar (2014) studied 38 platforms of the biggest cities in the U.S. to reveal different archetypes in the lower (consultation, placatation) and higher level (partnership) of Arnstein's Ladder. Gün et al. (2019) analysed 25 platforms with only three platforms in the highest level of participation (e.g. self-goverance). Falco & Kleinhans (2018) provide a broad overview with 113 platforms and find 11 self- goverance platforms but not all of them in a public interest context. All of those empirical studies show a current status of all levels of participation on technology-enabled platforms and help to shape the understanding of the differences between the levels but did not specify the level of self-governance. This broad field in the full range of participation levels is more and more getting into the focus of researchers. And so are the definitions of participatory platforms in general: There are different types in the manner of levels of participation and the intensity of the actors' involvement (Falco & Kleinhans, 2018). Those actors are all individuals and organizations who interact with the city, e.g the residents, activists, public agencies, non-governmental organizations, businesses (Desouza & Bhagwatwar, 2014). A participatory platform has specific goals within its purpose and offers a range of attractors or functionalities to enable participation (e.g. information distribution, group organisation or idea voting) or data collection (e.g.

tracking apps). Those come in different mediums that differ in online and offline (Desouza & Bhagwatwar, 2014).

2.5   Definition

As shown in the sections before there are several views that led us to the concept of bottom-up urbanism platforms. To describe it with the onion metaphor of Zhilin et al. (2019) we believe that the layers of our concepts are interweaved into each other. With a closer look and the perspective of Nam & Pardo (2011) and the dimensions of Gil-Garcia et al. (2020) we see all components addressed.

The technology factor (or information technology) in form of a participatory platform is working as an enabler of participation, offers information availabilty and shows progress in a transparent manner and acts as a supportive structure for the other components. The human factors are covered in our concept through the focus on the citizens as the actors (from the bottom-up and DIY) and the transfer of power towards them or annexation of rights by them (self-governance) which is directly linked to the institutional factors as well. And we see contribution to the discourse of future cities, first, provided by the platform as a new governance system in constant development and, second, by the outcome in form of projects. Within the institutional factors our concept should be seen as a collaborative approach of cities and citizens. To merge the sections before we offer the following definition:

Bottom-up Urbanism platforms focus on providing power to the actors of cities. They are specific playgrounds of self-governance guided by rules where citizens propose, develop and implement their projects. The citizens and civic organizations build urban interventions as small scale and short-term solutions to address specific problems. This offers a tangible contribution to the ongoing discourse of future cities and a new mode how we want to shape the future of our cities. The main components of these platforms are online mechanisms providing participation through different levels (e.g. start a project, crowdfunding) to involve a broader part of the citizenry and offline components (e.g. workshops, local funding) to complement and enhance digital mechanisms.

P. Abel, D. Miether, F. Plötzky & S. Robra-Bissantz:
The Shape of Bottom-Up Urbanism Participatory Platforms: A Conceptualisation and Empirical Study

459

3   Methodology & study design

This study is divided into two parts. Part A, is dedicated to the question how a coherent conceptual definition can be developed from previous research contributions. Based on a systematic literature review, the described conceptualizations were realized by integrating different concepts and approaches.

Part B, includes the construction of a database of existing participation platforms, the development of a qualitative research design, as well as their subsequent evaluation and selection. Several different steps were performed in Part B.

Step 1 includes a second systematic literature review, a questioning of experts and a detailed internet research, which identified a total of 143 established participation platforms.

In step 2, the focus was to identify participation platforms that provide their users with the highest level of participation, the analysis was assessed by two independent raters. Using the conceptualization and the information publicly available on the platforms, the following questions had to be answered positive:

Does the platform provide functions that enable citizens to create their own projects for the public space? Does the platform empower citizens to implement these projects? Were most of the projects realized by the citizens themselves?
Only the platforms that met these criteria were included in the further analysis, which reduced the sample to 26 participation platforms.

Table 1: Sample table

Platform name			
Co-citoyens	Ecocrowd	Gapfiller	Gut für Nürnberg
Hannover machen	Ioby	Moveforhunger	Open Berlin
Place2help	**Patronicity***	**Platzprojekt***	**Raumpioniere***
Rabryka*	Schützenplatz	**Sandkasten***	Startnext
Spacehive	Sagerdersamler	Urbaneoasen	Urbangreenewcastel
Voorjebuurt*	Wechange	**Yooweedoo***	

Source: own elaboration; *completed the questionnaire

In step 3, the relevant characteristics of the participation platforms have been defined. In consideration of the existing variability, we selected only those categories, which reflect the most widespread similarities and differences (e.g. number of active users or focus of participation). The data was collected primarily from the main websites of the platforms, and less frequently from secondary sources (e.g. Internet archive – "wayback machine"). If important platform components have changed over time, the current information was used and earlier changes were not taken into account. Platforms not active anymore were not taken into account and reduced the study sample to 23 platforms which are listed in Table 1.

In step 4, a questionnaire was sent to the platform operators. A total of seven platforms completed the questionnaire, which corresponds to an average response rate of 33%. Two platform operators rejected a participation and 12 didn't response to our request. In addition to our previously data collection, the questionnaire included a query of non-free-access information related to organizational structure (e.g. funding, personnel). The analysis of this data was performed purely descriptively to gain an initial impression of the characteristics of existing participation platforms.

4   Results

The results are divided into two parts. The first section (A) presents the integrated results of the descriptive analysis of studied participation platforms. The second section (B) reports the results of the qualitative questionnaire survey.

First, it is important to describe relevant distinguishing dimensions for the analysed platforms ($n = 23$). As a recent phenomenon, the digital participation platforms within this sample were founded between 2009 and 2019. In fact, $n = 2$ (8.70%) platforms (Open Berlin, 2017; Place2help, 2020) are not active or going to be terminated. In addition, $n = 2$ (8.70%) platforms (Urbaneoasen, 2020; Gapfiller, 2019) were conceptually transformed into non-participatory platforms. The average duration of activity (cut-off date: 12/31/2020) comes to $M = 6.74$ ($SD = 3.55$) years.

P. Abel, D. Miether, F. Plötzky & S. Robra-Bissantz:
The Shape of Bottom-Up Urbanism Participatory Platforms: A Conceptualisation and Empirical Study

461

An important component of any digital participation platform is the number of participation options, referred in the following as participation mechanisms. One can distinguish between 9 various mechanisms, depending on the depth of participation: Information, Like, Follow, Comment, Share, Crowdfunding, (Offline) Participation/Assistance, Join and Start Own Project. On average, platforms provided multiple mechanisms to their users ($M = 5.17, SD = 1.99$). Considering the participation focus, three relevant main priorities could be identified. Thus, 8 platforms focused on the implementation of projects (37.78%), 4 platforms focused on discussion (17.39%) and 11 platforms focused on funding (47.83%). In terms of projects, an average of $M = 137.13$ ($Mdn = 37.00$) were initiated per year, although the number varied widely ($SD = 231.01, IQR = 134$). A total of 16 platforms (69.57%) support their users through professional support services (e.g., coaching; [self-] learning). The degree of networking varies within the sample ($n = 22$), averaging $M = 29.86$ ($SD = 30.57$) network partners. One platform ($n = 1$) was excluded from the network analysis due to lack of available data.

Second, to make the results more precise, data collection was carried out in the form of a self-developed questionnaire. A total of $n = 21$ platforms were surveyed, with a response rate of 33.33% ($n = 7$). Individual data points were missing. The exact sample size was reported in such cases. The questionnaire was rated ($n = 6$) on a scale of 1 ("Very poor") to 10 ("Very good") as good ($M = 7.33, SD = .82$). In the following, superordinate characteristics are presented first. Secondly, the qualitative results are reported separately by platform.

The participation platforms have an average of $M = 4.57$ ($SD = 3.82$) employees and $M = 20.92$ ($SD = 39.47$) other persons, e.g., voluntaries ($n = 6$). On average, $M = 631$ ($Mdn = 489$) projects were launched. The number of projects varied widely ($SD = 682; IQR = 863$). An average of $M = 544$ ($Mdn = 228$) were successfully implemented ($SD = 679; IQR = 776$). Regarding the analyses of all projects, the Rabryka platform indicated only a reference frame of the calendar year 2019. The number of makers ($M = 903, Mdn = 400$) varied widely ($SD = 987, IQR = 1730$). The average number of network partners is $M = 48$ ($SD = 39$). On a scale of 1 ("Very poor") to 10 ("Very good"), the willingness of public partners to cooperate ($M = 7.86, SD = 1.95$) as well as the

collaboration with regional actors ($n = 6, M = 8.00, SD = 2.10$) were rated as high.

5   Discussion and Conclusion

Our main contribution is the comprehensive conceptualization through the integration of existing research to promote the understanding of the functioning of bottom-up urbanism participatory platforms. In addition, an adequate description of the population is given. That leads to a foundation for future research and the identification of relevant topics for the practical domain.

With regard to the population, the heterogeneity of the digital participation platforms was particularly evident. Above all, this made it difficult to compare the platforms. For example, the number of projects in the sense of "food donation campaign" from Moveforhunger can only be compared with difficulty with the "installation of containers" for initiatives from the Platzprojekt. In the future, fundamental conceptual differences within the platforms should lead to the distinction between different subpopulations.

In our study, we examined platforms that were well funded and were able to retain several employees. But also platforms that have given up or turned away from a participatory concept. Platforms which offer crowdfunding seem to have a more solid business model but there are no clear indications to break it down to that point. We found innovative participation approaches e.g. the combination of crowdfunding for citizens and institutional funding as match funding from Patronicity or a mixed campaign to provide funds, helping hands, expert knowledge and material donations from Raumpioniere.

Future research should not only examine the view of the platforms and their founders but also the citizens themselves, the city government and other stakeholders to provide implications for business models and for platform design. A further point is the question of which participation mechanisms in practice exert the most influence on the participation experienced. It is also important to question whether more participation mechanisms automatically mean a positive effect.

In the broader context of smart cities we shed light on a practical phenomenon that offers the foundation for further discussions and could be an inspiration to take the

P. Abel, D. Miether, F. Plötzky & S. Robra-Bissantz:
The Shape of Bottom-Up Urbanism Participatory Platforms: A Conceptualisation and Empirical Study

463

discourse on a more concrete level. We believe that our conceptualization should be communicated into the practical domain of platform providers to help them get a better understanding of their role within the field of participatory platforms.

References

Abel, P., Stuwe, C., & Robra-Bissantz, S. (2019). A.1 Design Claims für Community-Plattformen im Kontext des bottom-up Urbanism. Communities in New Media: Researching the Digital Transformation in Science, Business, Education and Public Administration - Proceedings of 22nd Conference GeNeMe, 1–11.

Anttiroiko, A.-V. (2016). City-as-a-Platform: The Rise of Participatory Innovation Platforms in Finnish Cities. Sustainability, 8(9), 922. https://doi.org/10.3390/su8090922

Arnouts, R., van der Zouwen, M., & Arts, B. (2012). Analysing governance modes and shifts — Governance arrangements in Dutch nature policy. Forest Policy and Economics, 16, 43–50. https://doi.org/10.1016/j.forpol.2011.04.001

Arnstein, S. R. (1969). Journal of the American Planning Association A Ladder Of Citizen Participation. Journal of American Institute of Planners, 35(4), 216–224. https://doi.org/10.1080/01944366908977225

Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart Cities in Europe. Journal of Urban Technology, 18(2), 65–82. https://doi.org/10.1080/10630732.2011.601117

Cardullo, P., & Kitchin, R. (2019). Being a 'citizen' in the smart city: up and down the scaffold of smart citizen participation in Dublin, Ireland. GeoJournal, 84(1), 1–13. https://doi.org/10.1007/s10708-018-9845-8

Cowley, R., Joss, S., & Dayot, Y. (2018). The smart city and its publics: insights from across six UK cities. Urban Research & Practice, 11(1), 53–77. https://doi.org/10.1080/17535069.2017.1293150

Crawford, M. (2008). The current state of everyday urbanism. Everyday Urbanism: Expanded Edition, 12–15.

de Waal, M. (2011). The ideas and ideals in urban media. In M. Foth, L. Forlano, C. Satchell, & M. Gibbs (Eds.), From Social Butterfly to Engaged Citizen: Urban Informatics, Social Media, Ubiquitous Computing, and Mobile Technology to Support Citizen Engagement (pp. 5–20). Cambridge, Massachusetts: MIT Press.

de Waal, M., & de Lange, M. (2019). Introduction—The Hacker, the City and Their Institutions: From Grassroots Urbanism to Systemic Change. In The Hackable City (pp. 1–22). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-2694-3_1

de Waal, M., & Dignum, M. (2017). The citizen in the smart city. How the smart city could transform citizenship. It - Information Technology, 59(6). https://doi.org/10.1515/itit-2017-0012

Desouza, K. C., & Bhagwatwar, A. (2014). Technology-Enabled Participatory Platforms for Civic Engagement: The Case of U.S. Cities. Journal of Urban Technology, 21(4), 25–50. https://doi.org/10.1080/10630732.2014.954898

Douglas, G. C. C. (2019). Privilege and Participation: On the Implications and Social Contradictions of Bottom-Up Urbanisms. In The Palgrave Handbook of Bottom-Up Urbanism (pp. 305–321). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-90131-2_19

Ertiö, T.-P. (2015). Participatory Apps for Urban Planning—Space for Improvement. Planning Practice & Research, 30(3), 303–321. https://doi.org/10.1080/02697459.2015.1052942

Ertiö, T.-P., & Bhagwatwar, A. (2017). Citizens as planners: Harnessing information and values from the bottom-up. International Journal of Information Management, 37(3), 111–113. https://doi.org/10.1016/j.ijinfomgt.2017.01.001

Fabian, L., & Samson, K. (2016). Claiming participation – a comparative analysis of DIY urbanism in Denmark. Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 9(2), 166–184. https://doi.org/10.1080/17549175.2015.1056207

Falco, E., & Kleinhans, R. (2018). Beyond technology: Identifying local government challenges for using digital platforms for citizen engagement. International Journal of Information Management, 40, 17–20. https://doi.org/10.1016/j.ijinfomgt.2018.01.007

Finn, D. (2014). DIY urbanism: implications for cities. Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 7(4), 381–398. https://doi.org/10.1080/17549175.2014.891149

Foth, M. (2017). Participation, Co-Creation, and Public Space. The Journal of Public Space, 2(4), 21. https://doi.org/10.5204/jps.v2i4.139

Gil-Garcia, J. R., Gasco-Hernandez, M., & Pardo, T. A. (2020). Beyond Transparency, Participation, and Collaboration? A Reflection on the Dimensions of Open Government. Public Performance & Management Review, 43(3), 483–502. https://doi.org/10.1080/15309576.2020.1734726

Gün, A., Demir, Y., & Pak, B. (2019). Urban design empowerment through ICT-based platforms in Europe. International Journal of Urban Sciences, 1–27. https://doi.org/10.1080/12265934.2019.1604250

Hajer, M. A. (2016). On being smart about cities: Seven considerations for a new urban planning and design. In A. Allen, A. Lampis, & M. Swilling (Eds.), Untamed Urbanisms (pp. 50–63). London: Routledge.

Iveson, K. (2013). Cities within the City: Do-It-Yourself Urbanism and the Right to the City. International Journal of Urban and Regional Research, 37(3), 941–956. https://doi.org/10.1111/1468-2427.12053

Keymolen, E., & Voorwinden, A. (2020). Can we negotiate? Trust and the rule of law in the smart city paradigm. International Review of Law, Computers & Technology, 34(3), 233–253. https://doi.org/10.1080/13600869.2019.1588844

Kickert, C., & Arefi, M. (2019). Introduction. In M. Arefi & C. Kickert (Eds.), The Palgrave Handbook of Bottom-Up Urbanism (pp. 1–10). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-90131-2_1

Kooiman, J. (2003). Governing as Governance. London: Sage.

Lefebvre, H. (1968). The Right to the City. In E. Kofman & E. Lebas (1996) (Hg.) Writings on cities (pp. 63–181). Cambridge: Blackwell.

Lombardi, P., Giordano, S., Farouh, H., & Yousef, W. (2012). Modelling the smart city performance. Innovation: The European Journal of Social Science Research, 25(2), 137–149. https://doi.org/10.1080/13511610.2012.660325

Lydon, M., & Garcia, A. (2015). Tactical Urbanism. Tactical Urbanism. Washington, DC: Island Press/Center for Resource Economics. https://doi.org/10.5822/978-1-61091-567-0

Nam, T., & Pardo, T. A. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. In Proceedings of the 12th Annual International Digital Government Research Conference on Digital Government Innovation in Challenging Times - dg.o '11 (p. 282). New York, New York, USA: ACM Press. https://doi.org/10.1145/2037556.2037602

Rauws, W. (2016). Civic initiatives in urban development: self-governance versus self-organisation in planning practice. Town Planning Review, 87(3), 339–361. https://doi.org/10.3828/tpr.2016.23

Schinkel, U., Jain, A., & Schröder, S. (2014). Local Action and Participation. Berlin: Jovis.

Senbel, M., & Church, S. P. (2011). Design Empowerment. Journal of Planning Education and Research, 31(4), 423–437. https://doi.org/10.1177/0739456X11417830

Silva, P. (2016). Tactical urbanism: Towards an evolutionary cities' approach? Environment and Planning B: Planning and Design, 43(6), 1040–1051. https://doi.org/10.1177/0265813516657340

P. Abel, D. Miether, F. Plötzky & S. Robra-Bissantz:
The Shape of Bottom-Up Urbanism Participatory Platforms: A Conceptualisation and Empirical Study

465

Statista. (2020). Proportion of population in cities worldwide from 1985 to 2050. Retrieved March 19, 2021, from https://www.statista.com/statistics/264651/proportion-of-population-in-cities-worldwide/

Talen, E. (2015). Do-it-Yourself Urbanism. Journal of Planning History, 14(2), 135–148. https://doi.org/10.1177/1538513214549325

Walser, K., & Haller, S. (2016). Smart Governance in Smart Cities (pp. 19–46). https://doi.org/10.1007/978-3-658-15617-6_2

Zhilin, S., Klievink, B., & de Jong, M. (2019). Community Self-Governance in the Smart City: Towards a Typology (pp. 81–97). https://doi.org/10.1007/978-3-319-89474-4_5

Soft Skills of The Chief Information Security Officer

Jeroen M.J. van Yperen Hagedoorn,[1]
Richard Smit,[2] Patric Versteeg[1] &
Pascal Ravesteijn[1]

[1] HU University of Applied Sciences, Utrecht, Netherlands; e-mail:
jmj.van.yperen@kpnmail.nl, Info@vsec.nl, pascal.ravesteijn@hu.nl
[2] Amsterdam University of Applied Sciences, Amsterdam, Netherlands; e-mail:
r.smit@hva.nl

Abstract This study addresses the role of a Dutch chief information security officer (CISO) and the soft skills required in this leadership role. The overview of soft skills is the outcome of the CISO perspectives in a Delphi study combined with an analysis of soft skills mentioned in job ads. A comparison with an earlier US-based study revealed that soft skills are ranked differently by Dutch CISOs. Moreover, we found that soft skills are not clearly described in job ads – none of these ads had explicitly listed soft skills. The present study demonstrates that CISOs with soft skills are in demand. The development of soft skills starts at a young age through various social activities and is also the result of self-actuation. The practical implications of this study are that it offers insights into the soft skills required for the role and discusses best-fitting leadership styles and ways in which organisations should include soft skills in recruitment.

University of Maribor Press

1 Introduction

IT systems or IT-powered solutions are prolific nowadays. Organisations use IT systems to support the information flows in their business processes, and IT is the primary process in digitally enabled companies (The Open University, 2019). The threat of misuse or abuse of IT systems causes high risks to organisations, and the need to protect IT systems and the information they contain has therefore increased significantly (ENISA, 2019).

A chief information security officer (CISO) is responsible for an organisation's information security programme (IGguru , 2019). The CISO manifesto defines a CISO as 'a senior-level executive who has the responsibility to establish and maintain the organisation's security program' (Hayslip, 2019), and according to Death (2019), the right combination of hard and soft skills is the key to being a successful CISO. Searching for the keywords 'CISO' and 'soft skills' in academic libraries at the HU University of Applied Sciences Utrecht, the University of Amsterdam and the Amsterdam University of Applied Sciences returned a single book: CISO Soft Skills by Collette, Gentile and Gentile (2008). This observation aligns with our problem statement that there is little to no academic research on the topic of CISOs, leadership and soft skills. Our research answers the following main research question: *What soft skills positively influence the CISO leadership position in Dutch organisations with more than 500 employees?*

2 Soft skills for leaders

Putrus (2019, p. 29) has found that the role of the CISO shifts from a technical implementer of security hardware and software to a more business-focused executive or leader. Multiple books are available both on CISO leadership, for example Essential principles for success (Fitzgerald & Krausse, 2007) and The CISO handbook (Gentile & Ron Collette, 2016), and on the CISO leadership role, risk management and CISO positions. However, here is little to no relevant academic literature that describes soft skills in relation to CISOs and their leadership position.

J. M.J. van Yperen Hagedoorn, R. Smit, P. Versteeg & P. Ravesteijn:
Soft Skills of The Chief Information Security Officer

469

Our literature review revealed a broad academic view on soft skills. Defining the term 'soft skills' is complicated and influenced by multiple factors (Chimatti, 2016), and Matteson et al. (2016) state that soft skills are often a catch-all category for non-technical skills. Technical skills are the skills required to perform a job, and soft skills are for interpersonal relations. Moreover, interactions with people (via parents, school, sports and other activities) fuel the learning process for developing soft skills.

Van Laar et al. (2017) describe a conceptual framework of 21st-century skills (learning skills, digital skills and life skills), as they suggest that employees' skills extend beyond their professional knowledge. Solely learning hard skills or gaining only professional knowledge is not sufficient for the ideal profile of the modern employee; the addition of soft skills to the skill set is emphasised by Cano et al. (2013). Their conclusion is that on one's career journey, the development of soft skills should occur in parallel to the development of the necessary hard skills. Furthermore, the development of professional IT skills is essential, and the development of personal skills will be even more significant.

Our research into soft skills in general yielded results of Weber et al, (2011); Zhang (2012) and Mar (2016). Part of the research question of this study pertains to the soft skills that positively influence the CISO leadership position. Therefore, the soft skills identified by Robles (2012, p. 455) were chosen as a foundation, as they are aimed at business leaders. Robles' list of 10 most relevant soft skills is presented below and was confirmed by our response group. No additional soft skills were introduced by the response group.

The 10 most relevant soft skills, according to Robles (2012, p. 445), are as follows:

- Integrity
- Communication
- Courtesy
- Responsibility
- Interpersonal skills
- Professionalism
- Positive attitude
- Teamwork skills

- Flexibility
- Work ethic

3   Methodology used and research undertaken

The goal of this research is to determine the soft skills that CISOs need to be able to attain a leadership role. For this, we used a multisource research approach (Zohrabi, 2013, p. 259) consisting of a literature review, an analysis of CISO job offers published on Dutch recruitment websites and a Delphi study with CISOs (Figure 1).
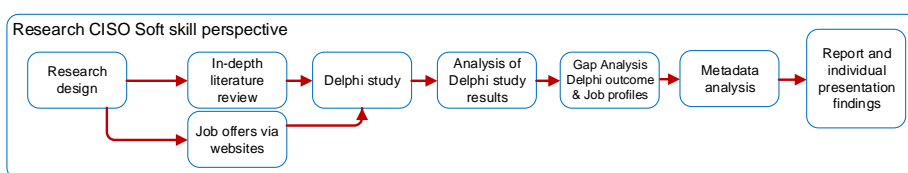
Figure 2: Consolidated research process

3.1   Job offer analysis

In 2020, over the course of 5 months, eight different Dutch recruitment websites were scanned for publicly published CISO job ads. The collected job offers were the input for a quantitative content analysis (Bryman & Bell, 2015, pp. 558-559) via text coding, as described in Section 4.1. The goal was to understand the demand side of job descriptions, and these insights were used as input in the Delphi study.

3.2   Delphi study

A Delphi study was conducted in the summer of 2020 to determine the soft skills from the CISO perspective. The Delphi method is an approach to capture existing knowledge and pinpoint areas of agreement or disagreement within a group of experts (Iqbal & Pipon-Young, 2009, p. 600). The target for the response group was set at 15 participants, which is a sufficient size, according to Okoli and Pwalowski (2004, p. 18).

J. M.J. van Yperen Hagedoorn, R. Smit, P. Versteeg & P. Ravesteijn:
Soft Skills of The Chief Information Security Officer

471

3.2.1 Composition of the CISO response group and Delphi design choices

The participants in Round 1 comprised 24 out of 31 invitees who met the criteria (CISOs in a Dutch organisation with 500 or more employees), and 23 CISOs took part in Round 2. The CISO response group works in eight different categories of organisations, such as education, information and communications, and human health, and the majority of the participants have an IT technology background. The key Delphi study design decisions are presented in Table 1.

Table 1: Taxonomy of Delphi design choices

Criteria	Choices for Delphi study
Number of rounds	Two rounds Reason: Round 1 was to capture the expert field input, and Round 2 was to reach consensus from the respondent group (Hasson, Keeney, & McKenna, 2000)
Consensus in the Delphi study	Consensus was reached when 70% or more of the respondents agreed on statements in Round 1 or when they agreed on the majority of their feedback in Round 2, based on The Delphi Technique: Making Sense of Consensus, Hsu and Sandford (2007).
Consensus count in the Delphi study	The use of combined consensus is based on Börger (2012, p. 157), who stated that respondents using Likert scales tend to answer moderately and avoid extreme answers. Therefore, both positive answers (agree and strongly agree) were combined into one positive answer. A similar structure was used on the negative answers (disagree and strongly disagree).
Mode of operation	Remote survey Reason: difficulty in planning the availability of CISOs to participate in a group debate during the pandemic of 2020.
Anonymity	Anonymous Reason: no need for personally identifiable information or opinions during analysis of data.
Locality	In the Netherlands, Dutch organisations with more than 500 employees are considered to be an enterprise and are expected to have a mature security organisation led by a CISO.
Media	Electronic survey Reason: ease of processing for participants and researchers.
Validation of survey questions	Survey questions were validated by Gordon B Willis' (1999) Question Appraisal System (QAS-99) and with two test rounds.

Criteria	Choices for Delphi study
	Reason: QAS-99 provides a questionnaire preparation checklist, and field testing individual rounds prevents unclear questions and research bias.
Socially desirable answers	The questions were structured based on a five-scale Likert question with the neutral answer in Position 3. With this construct, positive answers were in Positions 1 and 2, while negative answers were in Positions 4 and 5. Using this structure, we included both the socially desirable and the extreme answers of respondents (Börger, 2012, p. 157).

4 Results and interpretation

This section describes the three key areas of our research results. These are the outcome of the job ads analysis, the identified soft skills from a CISO perspective based on the Delphi study and the findings of the combined meta-analysis.

4.1 CISO job ads – a qualitative data analysis

In total, 77 CISO job ads were analysed to gain insights into the soft skills that Dutch organisations demand from the CISO role. Most of the advertisements clearly described the desired hard skills in bulleted lists. However, the descriptions of soft skills in the job ads were less clear. In the 77 analysed ads, the need for soft skills was mostly formulated in descriptive sentences and not explicitly listed. Writing and communicating appeared most frequently when analysing the job ads, for example '*Excellent written & spoken English essential. Multiple languages preferred*' (ING, 2019).

Text analysis provides information on the frequency of quotations, or text elements, that align with a label (groundedness of codes). In the data analysis, labelling was applied using the soft skills identified by Robles (2012), with the addition of 'leadership'. Leadership was added because this study researches the influence of soft skills on the CISO leadership position, and the adopted soft skills list does not include leadership. Our analysis of the demand for soft skills in job ads revealed that communication is the top soft skill based on the groundedness of codes, as illustrated in
Figure .

J. M.J. van Yperen Hagedoorn, R. Smit, P. Versteeg & P. Ravesteijn:
Soft Skills of The Chief Information Security Officer

473
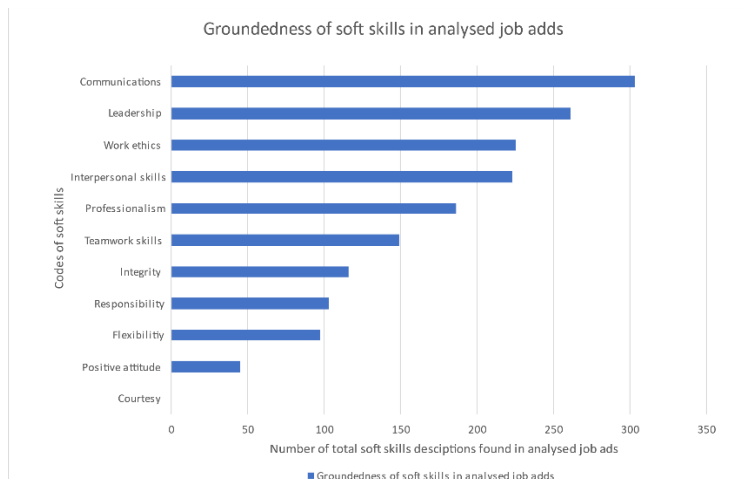
Figure 2: Groundedness of soft skills in job ads

4.2   Outcome of the Delphi study – CISOs

All of the CISOs in the response group (N = 24) are familiar with soft skills (12.5% extremely familiar, 66.7% very familiar) and the concept of leadership style skills (50% very familiar, 50% moderately familiar).

The CISO response group, however, values the relevance of soft skills and rated the use of variations of soft skills in different organisations and with different audiences as highly relevant. Based on the soft skills identified by Robles (2012), the participants agreed (strongly agree / agree, range of means: 1.43–1.78, *N* = 23) that soft skills are relevant for a CISO. Although the Dutch CISOs recognised those soft skills, they had a different view on their order of importance. They agreed on the ranking, based on Round 1, with a 73.9% consensus in Round 2 (*N* = 23). Table displays the soft skills' rankings.

Table 2: Soft skills confirmation – CISO response group

Rank	Soft skill	Ranking by Robles (2012)
1.	Communication	2
2.	Leadership style	-
3.	Integrity	1

Rank	Soft skill	Ranking by Robles (2012)
4.	Interpersonal skills	5
5.	Professionalism	6
6.	Work ethic	10
7.	Responsibility	4
8.	Teamwork skills	8
9.	Positive attitude	7
10.	Flexibility	9
11.	Courtesy	3
Note that the soft skill 'leadership' is not included in Robles' (2012) overview.		

Regarding leadership style, there was consensus among the CISO response group: Round 2 revealed that 78.3% (*N* = 23) *strongly agreed* and *agreed* on the question, 'do you agree with the consensus of the responders that the combination of transactional and transformational leadership is the most relevant leadership style a CISO must have?'. The transactional and transformational leadership style is based on research by Gurl et al. (2019), who relate three different leadership styles (management, transactional and transformational leadership) to users' compliance intentions.

The questions regarding leadership position, leadership style and the effect of softs skills on leadership (presented in Appendix A) were valued as highly relevant in the response group. The mean – ranging from 1.57 to 1.96 (*N* = 23) depending on the question – suggests that the respondents agree that the effect of soft skills is relevant for a CISO leadership position.

The CISOs in our study reached a high level of consensus (+70%, *N* = 23) on the relevance of soft skills in the recruiting phase. They have found that these skills are a priority in selecting the best-fitting CISO and must be described in CISO-related job ads.

The data set of the Delphi CISO response group was analysed for correlation using Spearman's rank correlation coefficient, where *d* = the difference between ranks and *n* = the amount of data points:

$$\left(\rho s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \right).$$

J. M.J. van Yperen Hagedoorn, R. Smit, P. Versteeg & P. Ravesteijn:
Soft Skills of The Chief Information Security Officer

475

This non-parametric measure of statistical dependence between the rankings of two variables is optimal for the ordinal data set of the CISO Round 2 responses (Baarda & Dijkum, 2014, p. 121). We now discuss the two most significant, positive correlations. The correlation between the relevance of CISO softs skills and CISO leadership skills is significantly positive, as shown in Appendix B. The respondents have found that a CISO applying soft skills has a better leadership position ($r = 0.721$, $p < 0.001$). The relevance is in soft skills in general for a CISO and in applying a different set of soft skills depending on the audience. Soft skills also contribute to various leadership styles in different circumstances. With a score of 0.527 on Spearman's rho, the second significant, positive correlation is between soft skills' contribution to various leadership styles in different circumstances and the prioritisation of soft skills in recruitment. Here, the CISOs emphasised the importance of the effect that soft skills have on the leadership position, and they value the focus of soft skills in the recruitment phase.

5 Discussion and conclusion

To answer our main research question – What soft skills positively influence the chief information security officer leadership position in Dutch organisations with more than 500 employees? – we list the relevant soft skills below. The ranking is based on both the CISO response group and the ranking from the collected job ads (both weighted equally):

1. Communication
2. Leadership – not included in Robles' (2012) overview
3. Interpersonal skills
4. Professionalism
5. Integrity
6. Work ethic
7. Responsibility
8. Teamwork skills
9. Positive attitude
10. Flexibility
11. Courtesy

This list is the outcome of the research. It is our conclusion that a CISO leadership position can be improved with these soft skills.

5.1 Practical implications

The primary goal of this study is to contribute to academic evidence regarding the soft skills that positively influence the CISO leadership position. We found that CISO job descriptions should contain clearer details regarding the necessary soft skills, taking into account the leadership style for a specific organisation. Better descriptions can improve the selection process, which in turn can improve alignment between the tasks and responsibilities of a CISO and the demand from business. In addition, based on the outcome that soft skills are highly relevant, education curricula and frameworks for personal development should focus on those much-needed skills. Lavasseur (2013) demonstrated that hard skills are developed through education and training, whereas soft skills are acquired through self-actuation; therefore, CISO-related curricula should include a focus on self-actuation.

From the analysis of the Delphi study, we found that soft skills should be a topic when recruiting a CISO. The practical implication is that an assessment of soft skills should be part of the recruitment process and balanced with the desired hard skills.

Combining the outcome of the job ads analysis and the experts' input that soft skills should be a priority in recruitment, the conclusion is that soft skills are not clearly listed in job ads. If soft skills are clearly articulated in job descriptions, then the recruitment of the most suitable candidate as well as the development of current CISOs' soft skills will be positively influenced. Furthermore, CISOs should be self-actuated for soft-skill development (Levasseur, 2013).

References

B.Willis, G., & Lessler, J. T. (1999). Question Appraisal System QAS-99. Researchgate, 1-86.
Baarda, B., & Dijkum, C. v. (2014). Statistiek met SPSS. Groningen: Noordhof Uitgevers bv.
Börger, T. (2012). Social desirablility and environmental valuation. Frankfurt am Main: Peter Lang.
Bryman, A., & Bell, E. (2015). Business Research Methods. New York: Oxford University Press.
Cano, C., Fernández-Sanz, L., & Misra, S. (2013, January - March). Featuring CIO: Roles, Skills and Soft Skills. International Journal of Human Capital and Information Technology Professionals, pp. 22-33.

Chimatti, B. (2016). Definition, development, assessment of soft skills and their role for the quality of organizations and enterprises. International Journal for Quality Research v10 n1, 97-130.

Collette, R., Gentile, M., & Gentile, S. (2008). CISO Soft Skills. New York: Auerbach Publications.

Death, D. (2019, February 18). Security Current Resources. Retrieved from Security Current: https://securitycurrent.com/must-have-skills-for-cisos-a-cisos-connect-report/

ENISA. (2019, January 28). ENISA Publications. Retrieved from ENISA: https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018

Fitzgerald, T., & Krausse, M. (2007). CISO Leadership: essential principles for success. New York: Auerback Publications.

Gentile, M., & Ron Collette, T. D. (2016). The CISO Handbook. New York: CRC Press.

Guhr, N., Lebek, B., & Breitner, M. H. (2019, March). The impact of leadership on employees' intended information security behaviour: An examination of the full-range leadership theory. Information Systems Journal, pp. 340-362.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. Journal of advanced nursing, 1008-1015.

Hayslip, G. (2019, May 5). CISO Manifesto: Recruiting CISOs not Unicorns. Retrieved from Peerlyst: https://www.peerlyst.com/posts/ciso-manifesto-recruiting-cisos-not-unicorns-gary-hayslip-cissp-cisa-crisc-ccsk?utm_source=peerlyst_perspective&utm_medium=email&utm_content=peerlyst_post&utm_campaign=top_posts_on_peerlyst_this_week_05142019

Hsu, C.-C., & Sandford, B. A. (2007). The Delphi Technique: Making Sense Of Consensus. Practical Assessment, Research & Evaluation, 1-8.

IGguru . (2019, March 12). IGguru News. Retrieved from IGguru Information Governance News & Community: https://igguru.net/2019/03/12/ciso-the-c-level-executive-missing-from-your-board/

ING. (2019, 04 28). Indeed. Retrieved from Indeed: indeed.nl

Iqbal, S., & Pipon-Young, L. (2009, July). The Delphi method. The Psychologist, pp. 598-601.

Levasseur, R. E. (2013). People Skills: Developing Soft Skills— A Change Management Perspective. Interfaces, 566-571.

Mar, A. (2016, July 7). 87 Soft Skills (The Big List). Retrieved from Simplicable: https://training.simplicable.com/training/new/87-soft-skills

Matteson, M. L., Anderson, L., & Boyden, C. (2016). "Soft Skills": A Phrase in Search of Meaning. portal: Libraries and the Academy, 16(1), 71-88.

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. Information & Management, 15-29.

Putrus, R. (2019). The role of the CISO and the digital security landscape. ISACA Journal, 2, 22-29.

Robles, M. M. (2012). Executive perceptions of the top 10 soft skills needed in today's workplace. Business Communication Quarterly, 453-465.

The Open University. (2019, 05 18). Information technology: A new era? Retrieved from OpenLearn: https://www.open.edu/openlearn/people-politics-law/politics-policy-people/sociology/information-technology-new-era/content-section-0?active-tab=description-tab

Van Laar, E., van Deursen, A. J., van Dijk, J. A., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. Elsevier Computers in Human Behavior, pp. 577-588.

Weber, M. R., Crawford, A., Rivera Jr., D., & Finley, D. A. (2011). Using Delphi Panels to Assess Soft Skill Competencies in Entry Level Managers. Journal of Tourism Insights, 1(1), 98-106.

Zhang, A. (2012). Peer Assessment of Soft Skills and Hard Skills. Journal of Information Technology Education: Research, 11, 155-168.

Zohrabi, M. (2013, Februari). Mixed Method Research: Instruments, Validity, Reliability and Reporting Findings . Theory and Practice in Language Studies, pp. 254-262.

Appendix A

The questions regarding leadership position, leadership style and the effect of softs skills on leadership as part of the first Delphi round are presented below.

Leadership style and effect	
A CISO has a leadership position. These questions are to get an insight in the leadership style and position of the CISO.	
Q19. Leadership styles Are you familiar with variations in leadership styles?	**Answer options:** Extremely Very Moderately Slightly Not at all Do not know
Q20. Which type of leadership style should a CISO have? Leadership by management position (leadership style based on a hierarchy position); Transactional leadership (leadership style based on rewards and controls in exchange for achieving previously defined governance goals); Transformational leadership (leadership style based on lead by good example and motivate followers to identify with them); Combination of Transactional and Transformational leadership; Other (please specify):	**Answer option:** Select best fit leadership style.
Q21. Do you agree with the following statements on CISO leadership and soft skills. The CISO will function at her/his best when they have a leadership position. The CISO should be able to demonstrate specific leadership styles in various situations. The CISO has a better leadership experience with soft skills. The CISO should be able to demonstrate specific soft skills in various situations.	**Answer options:** Extremely Very Moderately Slightly Not at all Do not know

J. M.J. van Yperen Hagedoorn, R. Smit, P. Versteeg & P. Ravesteijn:
Soft Skills of The Chief Information Security Officer

479

Appendix B: Table displaying Spearman's correlation coefficient – CISO response group

Survey questions (top row / left column)	Spearman's correlation coefficient – CISO response group							
	Relevance CISO_SK	Q10.1.	Leadership CISO_SK	Q6.1.	Q9.1.	Q11.1.	Q11.2.	Q12.1.
RelevanceCISO_SK (Q4.1–Q4.5)	1.000	-0.013	**0.721****	0.198	**0.219**	**0.269**	**-0.202**	-0.020
		0.954	0.000	0.365	0.315	0.214	0.355	0.929
Q10.1. Do you agree with the responders' consensus of very relevant on the question: 'the CISO will function at her/his best when they have a leadership position?'	-0.013	1.000	0.050	**0.251**	0.092	-0.138	-0.073	**-0.364**
	0.954		0.819	0.248	0.678	0.530	0.741	0.088
LeadershipCISO_SK (Q10.2–Q10.4)	**0.721****	0.050	1.000	0.173	**0.268**	**0.527****	0.029	0.196
	0.000	0.819		0.430	0.217	0.010	0.896	0.371
Q6.1. Do you agree with the ranking?	0.198	**0.251**	0.173	1.000	0.039	-0.168	-0.198	0.047
	0.365	0.248	0.430		0.861	0.442	0.365	0.833
Q9.1. Do you agree with the consensus of the responders that the combination of transactional and transformational leadership is the most relevant leadership style a CISO must have?	**0.219**	0.092	**0.268**	0.039	1.000	0.128	**-0.318**	**-0.315**
	0.315	0.678	0.217	0.861		0.561	0.139	0.143
Q11.1. Do you agree with the responders' consensus of	**0.269**	-0.138	**0.527****	-0.168	0.128	1.000	0.103	**0.500***
	0.214	0.530	0.010	0.442	0.561		0.640	0.015

very relevant on the question: 'do soft skills have priority in selecting the best fit CISO?'								
Q11.2. Do you agree with the responders' consensus of very relevant on the question: 'is it relevant that the CISO job ad describes the soft skills needed?'	**-0.202**	-0.073	0.029	-0.198	**-0.318**	0.103	1.000	0.155
	0.355	0.741	0.896	0.365	0.139	0.640		0.479
Q12.1. Do you agree with the responders' consensus that it is better to select the candidate with best-developed soft skills, but trainable for hard skills?	-0.020	**-0.364**	0.196	0.047	**-0.315**	**0.500***	0.155	1.000
	0.929	0.088	0.371	0.833	0.143	0.015	0.479	

. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).**	c. Listwise $N = 23$

Legend:	**Significant positive correlation**	**Moderate significant positive correlation**	**Moderate significant negative correlation**	**Mediocre significant negative correlation**	Significance 2-tailed.

The correlation scoring on Spearman's rho is based on the following ratio:

o   0.4 and above: Significant positive correlation (negative is the opposite scale of the positive scoring)

o   0.3–0.4: Moderate significant positive correlation (negative is the opposite scale of the positive scoring)

o   0.2–0.3: Mediocre significant positive correlation (negative is the opposite scale of the positive scoring)

o   0.2 and below: Irrelevant positive correlation (negative is the opposite scale of the positive scoring)

QUANTUM COMPUTER RESISTANT CRYPTOGRAPHIC METHODS AND THEIR SUITABILITY FOR LONG-TERM PRESERVATION OF EVIDENTIAL VALUE

THIEL CHRISTIAN[1] & THIEL CHRISTOPH[2]

[1] OST Ostschweizer Fachhochschule, School of Management, Rosenbergstrasse 59, Postfach, 9001 St.Gallen, Switzerland; e-mail: christian.thiel@ost.ch
[2] FH Bielefeld University of Applied Sciences, Faculty of Minden Campus, 32427 Minden, Germany; e-mail: christoph.thiel@fh-bielefeld.de

Abstract In the areas of electronic identification and electronic trust services, the Regulation No. 910/2014 of the European Parliament and of the Council on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC (eIDAS) creates uniform regulations for electronic signatures, seals, time stamps, registered mail and website certificates in the European single market. All developments that affect the security of signature procedures have an impact. In this study, we consider the candidates for quantum computer-resistant asymmetric cryptographic (PQC) methods currently under investigation in international research and standardization and evaluate their suitability for PKI systems with a focus on long-term preservation of evidential value, as is the case in particular with eIDAS-compliant signature solutions. Based on an evaluation system proposed by us - an adaptation of the system from [2] - we compare the application requirements with the properties of the candidates and recommend suitable methods.

1   Introduction

This study focuses on quantum computer-resistant crypto methods, also called post-quantum cryptography (PQC) after J.D. Bernstein (in particular in asymmetric methods). It is not comprehensive and does not list every quantum computer-resistant asymmetric method ever proposed. Instead, it lists a representative sample (as of End 2020) of cryptographic techniques that are being discussed in academia, are supported by currently active research teams, may be viable for real-world applications, and are therefore suitable candidates for consideration by various standardization organizations for standardization. Beyond NIST's PQC standardization, we also consider extensions of classical algorithms as well as quantum-assisted algorithms (i.e., the use of quantum technology to augment classical systems, see also [10]) with respect to the possibility of providing sufficient quantum computing resistance.

2   Overview of the procedures

In this study, we define PQC methods as cryptographic methods (in particular asymmetric cryptographic methods) which, according to the current state of research, can possibly provide sufficient security against attacks that use the capabilities and properties of quantum computers, i.e. are "quantum computer resistant". In this context, the procedures themselves do not use any support from quantum computers for preparation and execution.

The underlying principle of continuing to use the previously employed public-key methods such as RSA and ECDSA (Elliptic Curve Digital Signature Algorithm) with significantly larger keys than is currently customary in the post-quantum era is obvious at first glance. On the one hand, the approach of increasing the key sizes of RSA and ECDSA to cope with ever-improving cryptanalysis and newly discovered attacks is already a tradition (see, e.g., evolution of NIST's SP 800-57 Part 1[11]). In the context of quantum computers, this principle would very quickly lead to large and unwieldy key sizes that corresponding keys might not be usable in practice:

Quantum computers are based on the concept of qubits (quantum bit), where each qubit exists simultaneously as a superposition (superposition or also called coherence) of the states 1 and 0 and all those in between. The number of qubits needed on a quantum computer to break RSA[1] is estimated to be 2n+3 [12] and 2n+2 [13], which means that a quantum computer with about 4,000 qubits is needed to break an RSA-2048 signature (further algorithm optimizations are expected, so the actual number of qubits needed is expected to be lower). Shor's QFT algorithm can also be adapted to solve the discrete logarithm problem. The number of qubits to break ECDSA is "approximately" 6n [6]. This means that a quantum computer with about 1,500 qubits can break an ECC-P256 signature. Following the assumption of Neven's law [14] (the quantum equivalent of Moore's law), one can estimate that the computational power of quantum computers increases at a "double exponential rate" compared to classical computers.

If we start with 100 qubits in a given year and double the qubits every 18 months, 9 years later we will probably have computers with over 6000 qubits and in 32 years we will be able to break a 1-million-bit RSA key. Post-qubit RSA (i.e., RSA with such large key lengths) was studied by Bernstein [15], who showed the technical feasibility of implementing a terabit key using 231 4096-bit primes as factors. At these key sizes, each RSA operation amounted to tens or hundreds of hours. In practice, such a system can thus probably be ruled out. It should be noted, incidentally, that post-quantum RSA was in Round 1 of the NIST PQC competition but was not selected for Round 2.

Currently, it is unclear how many qubits the most powerful quantum computers have at the time of writing. The company IQM FINLAND OY is to build a quantum computer which is to have 50-qubits by the end of the third phase in 2024 ([36]). Google LLC, IBM, and others have developed machines with about 50 or more high-quality qubits (see [34], [35]). IBM is planning (even faster than Neven's law would suggest) more quantum computers with 127 qubits in 2021, 433 qubits in 2022, and over 1000 qubits in 2023 [35].

---

[1] i.e. breaking any private key

If IBM's development speed remains the same, we could expect the above-mentioned 6000 qubits to be reached around 2026 to 2027. Even with somewhat slower developments, one must assume that this will be the case from the year 2030. Although attacks against symmetric cryptosystems using quantum computers and algorithms by Grover or Simon (see [3] and [4]) are more effective than attacks using conventional computers, it is currently assumed that doubling the effective key length cancels out this advantage of quantum computers. Thus, for example, AES256 would be about as secure against a quantum computer as AES128 is against conventional computers.

Assuming the availability of sufficiently powerful quantum computers in the near future, it is obvious to use them not only as a tool to attack classical crypto methods, but also to investigate how quantum computer-resistant crypto methods could be realized with their help. The use of quantum computers to perform certain cryptographic operations is called quantum cryptography. Corresponding operations typically exploit the quantum properties of superposition, interference, and entanglement, which are not reproducible by classical computers. Quantum-enhanced security [17] is then understood to be the extension of classical non-quantum systems that make use of or are augmented by quantum technology to improve their ability to secure their data and transactions against adversaries that may be fully quantum capable.

While quantum key distribution (QKD) (see [18], [19]) is often equated with (general) quantum cryptography, QKD is based on the Vernam one-time pad and is therefore more suitable only for key exchange and encryption. Quantum researchers have introduced several quantum digital signature schemes (see [20] - [22]), but since they typically refer to QKD, they would be better referred to as data authentication schemes. As of this writing, we are unable to identify any quantum digital signature schemes in the literature that actually have the necessary constructs of a digital signature scheme and are EUF-CMA secure (existentially unforgeable under chosen message attack), let alone post-quantum secure.

Based on the above considerations classic cryptographic methods such as RSA and ECDSA with very large keys are ruled out (in the medium term) and can at best be used for a short transition phase (i.e., for the next 9 years at most). Signatures generally have a rather short lifetime and in principle only need to be secure up to

the time of their verification. If a signature procedure can be broken by a quantum computer in the future, today's signature certificates will probably already have expired. Only in the case of very long validity periods for signature keys should caution already be exercised. According to the current state of research, quantum-enhanced processes do not (yet) play a role specifically for electronic signatures. In the medium and long term, therefore, the focus should be on PQC processes.

3   Parameterized evaluation of PQC methods and applications

The objective of this study is not to replicate NIST's research in the NIST PQC competition (see [23], [24])., but to build on it and make it more concrete in order to find a basis for assessing the concrete practical applicability of a procedure in building blocks of e-business applications. In doing so, we extend the evaluation scheme from [2]. We define the three value ranges Small (S), Medium (M), and Large (L) for different parameters of the procedures, respectively. Specifically, we consider the following parameters.

- Key Generation Resources (KeyGen() Resources)
- Key sizes of the public and private keys
- Key Lifetime: Certain signature processes only allow the private signature key to be used for a limited number of signature creations. We record this using the "key lifetime".
- Resources for signature creation (Sign() resources) or encryption (Crypt() resources).
- Size of a signature (Signature Size) or size of a ciphertext (Cipher Size)
- Time for the creation of a signature (Signature Time) or the creation of a ciphertext (Crypt Time)
- Resources for signature verification (Ver() resources) or decryption (Decrypt() resources).

The parameters are categorized as follows in table 2 (assuming a single core of a current Intel I7 processor for mobile devices running at 3.2 Ghz, as in [2] and [25]):

Table 2: Parameters and their categories for evaluating the practical usability of PQC methods in applications

	Small (Optimal)	Medium	Large
KeyGen() Resources **Sign() Resources** **Crypt() Resources** **Ver() Resources** **Decrypt() Resources**	can be executed on a chip card. (< 3M cycles)	executable on a terminal/mobile phone (< 30 M cycles)	Requires operation on a powerful laptop (> 30 M cycles)
Key Size **Signature Size** **Cipher Size**	< 2Kbits (e.g. ECC-P256)	< 2Kbytes (e.g. RSA-8192)	> 2Kbytes
Key Lifetime	< 1000 signatures per key	< 10000000 signatures per key	unlimited
Signature Time	< 1ms per signature	< 100ms per signature	> 100ms per signature
Crypt Time	< 1ms per encryption	< 100ms per encryption	> 100ms per encryption

In order to evaluate the suitability of different PQC methods for concrete applications, we first look at the applications from the ETSI (see [26]) and now use the parameters described above as the requirements of the applications for a PQC procedure to be deployed (the parameters are therefore no longer descriptive in nature but have a requirement character). Of course, there are other use cases for asymmetric (signature) procedures, but the selection considered covers common scenarios from the areas of finance (for business), infrastructure (for people and devices), cloud & Internet (for business-to-business, business-to-consumer, peer-to-peer, and Internet-of-Things interactions), and enterprise (for companies). Based on [2] and [26], the following picture emerges in Table 3.

Table 3: Parameter evaluation of typical use cases of asymmetric signature solutions

	KeyGen() Resource	Private Key Size	Public Key Size	Key Lifetime	Sign() Resource	Signature Size	Signature Time	Ver() Resource
3SKey	S	M	M	M	S	M	M	L
EMVSDA	L	L	S	L	L	S	L	M
EMVDDA	S	S	S	M	S	S	S	M
CA Key [2]	L	L	M	M	L	L	L	M
ICAO 9303	L	L	S	S	L	S	L	M
GSM eSIM	L	S	M	S	S	M	M	S
TLS server	L	L	L	L	L	L	S	M
TLS client	M	L	L	L	M	L	M	L
Bitcoin M	L	L	L	M	M	M	M	L
FIDO [3]	M	L	L	M	M	L	M	L
USB signature token [4]	M	L	L	M	M	L	M	L
PGP/ SMIME	M	L	L	L	M	L	M	M
PAdES / AES [5]	L	L	M	L	L	M	M	M
QES [6]	S	M	M	M	S	M	M	M
Code Sign	L	L	L	M	L	L	L	M

4   Status of standardization

To facilitate the development of new quantum computer-resistant and practical methods, the National Institute of Standards and Technology (NIST) initiated a standardization process in 2016 (see [7]). After an evaluation and selection process based on public feedback and internal review by NIST, those methods were identified to move to the third round of review as finalists [16]: The encryption and key agreement/transmission methods are Classic McEliece [30], CRYSTALS-

---

2 Simplified consideration for qualified trust service providers

3 We consider FIDO and other tokens with comparable computational power and memory for strong authentication

4 Here we consider signature tokens that are more powerful than common smart cards.

5 advanced electronic signatures when using a document server with HSM to sign documents

6 qualified electronic signatures when using a signature creation device such as a smart card or USB token

KYBER [31]], NTRU [32], [33] and SABER [25]. The finalists for digital signatures are CRYSTALS-DILITHIUM [27], FALCON [28] and Rainbow [29].

A special feature are so-called stateful hash-based signatures, a special class of signature schemes with certain restrictions, from which currently XMSS (eXtended Merkle Signature Scheme) [8] and LMS (Leighton-Micali Signatures) [9] are in the process of standardization at the Internet Engineering Task Force (IETF) and at NIST, so that standards can be expected earlier than in the above-mentioned PQC process at NIST. The use cases mentioned are code signing and issuing PKI root certificates from certification authorities.

The standardization organizations ETSI and ISO are also involved in PQC standardization with their own working groups. At present, however, it looks as if ETSI and ISO will rely on NIST for the initial selection of procedures. At the moment it seems rather unlikely that other fundamentally new procedures not yet considered by NIST will emerge as part of the (international) standardization effort. In this study, we therefore restrict ourselves to the above mentioned candidates and go on to investigate their suitability for e-business applications.

5   Evaluation of the procedures

We apply the parameter description introduced in Section 3 to the procedures listed above. According to [2], we obtain the following parameter profiles for the current favorites of the NIST and IETF standardization of PQC signature methods in Table 4:

Table 4: Parameter profiles for PQC signature methods

	CRYSTALS-DILITHIUM	FALCON	Rainbow	XMSS	LMS
KeyGen() resource	S	M	L	L	L
Private Key Size	L	M	L	S	S
Public Key Size	M	M	L	M	S
Key Lifetime	L	L	L	M	M
Sign() resource	M	S	S	M	S
Signature Size	M	S	S	M	M
Signature Time	S	S	S	M	S
Ver() Resources	S	S	S	S	S

For encryption methods and key exchange or key transport (KEM) methods, we combine the results from [38, Table 3] with the evaluation method from [2] and obtain the following parameter profiles for the current favorites of NIST's standardization in Table 5:

Table 5: Parameter profiles for PQC encryption methods and key exchange/key transport methods

	Classic McEliece	CRYSTAL-KYBER	NTRU	SABER
KeyGen() resource	S	L	L	L
Private Key Size	L	S	L	S
Public Key Size	M	S	L	M
Crypt() resource	M	S	S	M
Cipher Size	M	S	S	M
Crypt Time	S	S	S	M
Decrypt() resources	S	S	S	S

If we contrast the parameterization of the procedures with the parameterization of the applications from Table 3, we can derive an evaluation scheme as in [2] based on a point assignment for the suitability of the procedures for the respective application. The basis of scoring is as follows: If the procedure provides a score for a single parameter that is equal to or better than what the application provides, then the score remains unchanged. If the procedure for a parameter is worse by a range (e.g. M instead of S) than what the application allows, then 1 is subtracted from the score for each such parameter[7]. If there is a parameter for which the procedure is two ranges worse (e.g., L instead of S) than what the application allows, then we consider the procedure to be not fit (NF = not fit). For quantitative purposes, we assign a score of -100 for each NF. Then the individual ratings of the parameters are summed up. The most suitable procedures can now be found for each application. A score of zero means that no changes are required and the process can most likely be used for the application. A negative score means that the procedure is not completely suitable, but that optimizations for the procedure may need to be found. After zero, the algorithm with the highest score (i.e., with the lowest negative score)

---

[7] For each individual parameter, the context determines whether a larger or smaller value is better. For example, a larger memory requirement is worse, but a longer lifetime of a key may be better.

is the next most suitable, as it requires the least number of changes to be used by the application.

Table 6: Selection of suitable processes per application

	CRYSTALS-DILITHIUM	FALCON	Rainbow	XMSS	LMS
3SKey	-2	**-1**	NF	NF	NF
EMV-SDA	-2	**-1**	NF	-2	-2
EMV DDA	NF	**-3**	NF	NF	NF
CA Key	**0**	**0**	-1	**0**	**0**
ICAO 9303	-2	**-1**	NF	-2	**-1**
GSM eSIM	NF	-1	NF	-1	**0**
TLS server	**0**	**0**	**0**	-2	-1
TLS Client	**0**	**0**	-1	-2	-2
Bitcoin	**0**	**0**	-1	-2	-2
FIDO	**0**	**0**	-1	-1	-1
USB signature	**0**	**0**	-1	-1	-1
PGP	**0**	**0**	-1	-2	-2
PDF-AES [8]	**0**	**0**	-1	-1	-1
PDF QES [9]	-2	**-1**	NF	NF	NF
Code sign	**0**	**0**	**0**	**0**	**0**
Points	**-208**	**-8**	**-606**	**-315**	**-312**

As a result no PQC method currently considered is suitable for all mentioned use cases in Table 3 (in particular for replacing RSA and EC in all use cases). For various use cases, such as for root CA keys, for code signing or for applications where signature creation and verification are performed on a powerful PC, the PQC procedures currently considered in the NIST standardization can be used. This also applies, with minor restrictions, to the use of tokens that are more powerful than "usual" smart cards such as signature cards. However, it becomes critical if the procedure is to be executed on hardware with limited computing power, such as a smart card. Thus, there are at least approaches for a first solution in the eIDAS context if not a completely satisfactory answer to the upcoming developments.

---

[8] when using a document server with HSM for signing documents
[9] when using a signature creation device such as a smartcard or USB token

6   Recommendations

Post-quantum cryptography will become the standard in the long term [1]. Consideration should be given at an early stage, as part of a measured risk management process, as to whether and when a switch to quantum computing resistant methods should be made (depending on the application) [1]. Especially in connection with signatures with a medium validity period of the certificates (3-5 years), there is no need to rush. For cryptographic applications that process information with long secrecy periods and high protection requirements, however, there may already be a need for action now [1]. The danger here is that messages for key negotiation and the data encrypted with the negotiated keys are collected in advance and decrypted in the future with the aid of a quantum computer ("store now, decrypt later"). Caution is also required with very long validity periods for signature keys. It is therefore already necessary to discuss how a migration to post-quantum cryptography to a Fully Quantum Safe Cryptographic State (FQSCS) for e-business applications can be initiated today.

References

Migration zu Post-Quanten-Kryptografie, Handlungsempfehlungen des BSI, August 2020

Teik Guan Tan und Jianying Zhou, "A Survey of Digital Signing in the Post Quantum Era", Cryptology ePrint Archive, Report 2019/1374, 2019, https://eprint.iacr.org/2019/1374

L. K. Grover, "Quantum mechanics helps in searching for a needle in a haystack," Physical review letters, vol. 79, no. 2, p. 325, 1997.

Simon, D.R. (1994), "On the power of quantum computation", Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on: 116–123, retrieved 2011-06-06

D. J. Bernstein, "Comparing proofs of security for lattice-based encryption", Second PQC Standardization Conference, Cryptol. ePrint Arch., 2019, https://eprint.iacr.org/2019/691

Announcing Request for Nominations for Public-Key Post-Quantum Cryptographic Algorithms", 81 Federal Register 92787 (December 20, 2016), pp. 92787 92788. https://federalregister.gov/d/2016-30615

A. Hülsing, D. Butin, S. Gazdag, J. Rijneveld, and A. Mohaisen, "XMSS: eXtended Merkle signature scheme," Available Online: https://tools.ietf.org/html/rfc8391, 2018 [last accessed: September 2020].

F. T. Leighton and S. Micali, "Large provably fast and secure digital signature schemes based on secure hash functions," 1995, uS Patent 5,432,852.

P. Wallden and E. Kashefi, "Cyber security in the quantum era", Commun. ACM, vol. 62, no. 4, p. 120, 2019.

E. Barker, "SP 800-57 part 1 rev. 4 recommendation for key management part 1: General," NIST special publication, vol. 800, p. 57, 2016.

Y. Takahashi and N. Kunihiro, "A quantum circuit for shor's factoring algorithm using 2n+2 qubits," Quantum Information & Computation, vol. 6, no. 2, pp. 184–192, 2006.

S. Beauregard, "Circuit for shor's algorithm using 2n+3 qubits," arXiv preprint quant-ph/0205095, 2002.

K. Hartnett, "A new law to describe quantum computing's rise?" Available
Online: https://www.quantamagazine.org/does-nevens-lawdescribe-quantum-computings-rise-20190618/, 2019 [last accessed: September 2020].

D. J. Bernstein, N. Heninger, P. Lou, and L. Valenta, "Post-quantum RSA," in International Workshop on Post-Quantum Cryptography. Springer, 2017, pp. 311–329.

Moody, D. , Alagic, G. , Apon, D. , Cooper, D. , Dang, Q. , Kelsey, J. , Liu, Y. , Miller, C. , Peralta, R. , Perlner, R. , Robinson, A. , Smith-Tone, D. and Alperin-Sheriff, J. (2020), Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process, NISTIR, National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.IR.8309

P. Wallden and E. Kashefi, "Cyber security in the quantum era", Commun. ACM, vol. 62, no. 4, p. 120, 2019.

C. H. Bennett and G. Brassard, "Quantum cryptography: public key distribution and coin tossing," Theor. Comput. Sci., vol. 560, no. 12, pp. 7–11, 2014.

A. K. Ekert, "Quantum cryptography based on bell's theorem", Physical review letters, vol. 67, no. 6, p. 661, 1991.

D. Gottesman and I. Chuang, "Quantum digital signatures," arXiv preprint quant-ph/0105032, 2001.

M.-Q. Wang, X. Wang, and T. Zhan, "An efficient quantum digital signature for classical messages," Quantum Information Processing, vol. 17, no. 10, p. 275, 2018.

W. Li, R. Shi, and Y. Guo, "Blind quantum signature with blind quantum computation," International Journal of Theoretical Physics, vol. 56, no. 4, pp. 1108–1115, 2017.

NIST, "Submission Requirements and Evaluation Criteria for the Post-Quantum Cryptography Standardization Process," Available
Online: https://csrc.nist.gov/CSRC/media/Projects/Post-Quantum- Cryptography/documents/call-for-proposals-final-dec-2016.pdf, 2016 [last accessed: September 2020].

"Post-Quantum Cryptography Round 2 Submissions," Available
Online: https://csrc.nist.gov/Projects/Post-Quantum- Cryptography/Round-2-Submissions, 2019 [last accessed: September 2020].

"Saber: Module-LWR based key exchange, CPA-secure encryption and CCA-secure KEM", Available Online: https://eprint.iacr.org/2018/230.pdf, 2018 [last accessed: September 2020].

ETSI, "Quantum safe cryptography; case studies and deployment scenarios etsi gr qsc 003 v1.1.1," Available Online: https://www.etsi.org/deliver/etsi gr/QSC/001 099/003/01.01.01 60/ gr QSC003v010101p.pdf, 2017 [last accessed: September 2020].

Ducas, L., Kiltz, E., Lepoint, T., Lyubashevsky, V., Schwabe, P., Seiler, G., and D. Stehle, "Crystals-dilithium: A lattice-based digital signature scheme", IACR Cryptology ePrint Archive , 2017, https://eprint.iacr.org/2017/633

Fouque, P-A., Hoffstein, J., Kirchner, P., Lyubashevsky, V., Pornin, T., Prest, T., Ricosset, T., Seiler, G., Whyte, W., and Z. Zhang, "Falcon: Fast-Fourier lattice- based compact signatures over NTRU", IACR Cryptology ePrint Archive , 2018, https://falcon-sign.info/

Ding, J. and D. Schmidt, "Rainbow, a New Multivariable Polynomial Signature Scheme", 2005, https://link.springer.com/chapter/10.1007/11496137_12

McEliece R (1978) A public-key cryptosystem based on algebraic coding theory. The Deep Space Network Progress Report, DSN PR 42-44. NASA. Available at https://tmo.jpl.nasa.gov/progress_report2/42-44/44N.PDF

CRYSTALS-Kyber (version 2.0) – Submission to round 2 of the NIST post-quantum project. Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. Specification document ,2019, Available at https://pq-crystals.org/kyber/data/kyber-specification-round2.pdf [last accessed: September 2020].

Institute of Electrical and Electronics Engineers (2009) IEEE Standard 1363.1-2008 - Specification for Public Key Cryptographic Techniques Based on Hard Problems over Lattices (IEEE, Piscataway, New Jersey, United States). https://doi.org/10.1109/IEEESTD.2009.4800404

American National Standards Institute (2010) ANSI X9.98-2010 -Lattice-Based Polynomial Public Key Establishment Algorithm for the Financial Services Industry (ANSI, New York City, United States). Available at https://webstore.ansi.org/standards/ascx9/ansix9982010r2017

Arute, F., Arya, K., Babbush, R. et al.: Quantum supremacy using a programmable superconducting processor. Nature 574, 505–510 (2019). https://doi.org/10.1038/s41586-019-1666-5

Gambetta, J.: „IBM's Roadmap For Scaling Quantum Technolog, IBM Research", IBM Research Blog, September 15, 2020, https://www.ibm.com/blogs/research/2020/09/ibm-quantum roadmap/, [last accessed: November 2020].

Finland selects IQM to build its first quantum computer; to deliver a 50-qubit machine by 2024., November 2020,https://www.meetiqm.com/articles/press-releases/finland-selects-iqm-to-build-its-first-quantum-computer-to-deliver-a-50-qubit-machine-by-2024, [last accessed: November 2020].

Smart Services for Energy Communities: Insights on Options and Priorities from a Multicriteria Mapping Study in Germany

Julio Viana,[1] Rainer Alt[2] & Olaf Reinhold[1]

[1] Social CRM Research Center/Leipzig, Germany; e-mail: julio.viana@scrc-leipzig.de, olaf.reinhold@scrc-leipzig.de
[2] Leipzig University, Information Systems Institute/Leipzig, Germany; e-mail: rainer.alt@uni-leipzig.de

Abstract Energy Communities are finding their way into the local energy systems as new regulations surge. However, they often lack resources due to their limited size, and depend heavily on subsidies for providing competitive offerings. In parallel, new technologies support the development of smart services for the energy market and provide chances for increasing the competitiveness of energy communities. This paper utilizes the multi-criteria mapping (MCM) method to discuss with stakeholders from energy communities in Germany the relevance and priorities for realizing specific smart services. A general ranking, as well as four perspective-based rankings, are analyzed by discussing contrasts and uncertainties. The results provide relevant insights on potentials from each service and a basis for the design of new information systems and architectures for energy communities.

1   Introduction

New regulations are setting ways for further development of the energy market worldwide (Spence 2019; Vasily Kupriyanovsky et al. 2019). In Germany, the energy industry is strongly regulated and generates a positive economic impact after the German Renewable Energy Sources Act (EEG), increasing investments and employment (Hillebrand et al. 2006), as well as serving as a model for legislation in other countries (Lehr et al. 2008). Consequently, projects on the demand side emerged (Palensky and Dietrich 2011) and new concepts, such as Smart Energy Communities (SECs), found their way into the regulatory systems. SECs consist of a group of households with different forms of electric loads and technologies integrated into a control system, which actively manages generation and demand in the community (Fazeli et al. 2011). Recently, citizens started to engage in local energy systems due to community identity, social norms, trust and environmental concern (Kalkbrenner and Roosen 2016; Fazeli et al. 2011; Massey et al. 2018).

In parallel, technological advancements pave the way for 'smart' energy services (SES) (Mathiesen et al. 2015; van Dinther et al. 2021) using a smart grid architecture based on 'prosumers' - users that consume and produce energy (Grijalva and Tariq 2011). Services are smart when based on hard field intelligence, and are processing a large amount of data and giving decision-makers more visibility into their business (Allmendinger and Lombreglia 2005), using interconnected Information Systems (IS) for data acquisition, algorithms, data reports and interfaces for visualization and configuration (Palensky and Dietrich 2011; Beverungen et al. 2019).

While some studies discuss energy communities with a focus on choosing the type of renewable energies (Karunathilake et al. 2019), their ecosystems (Vernay and Sebi 2020) or social innovation aspects (Caramizaru and Uihlein 2020), current literature on SECs focuses on specific services and processes, such as big data analysis (Zhou et al. 2016), smart meters (Anda and Temmen 2014), peer-to-peers interconnected smart homes (Steinheimer et al. 2012) and smart Internet of Things (IoT) (Giordano et al. 2020). New business models based on SESs, including peer-to-peer (P2P) marketplaces, microgrids or virtual power plants derive from energy generated intelligently and optimized to balance with its demand (Paukstadt and Becker 2019). Such models use smart systems, providing a more holistic approach rather than focusing on specific services or only on smart grids (Lund et al. 2017).

SECs benefit from SES as they secure reliability, enhance market service, minimize environmental impact, reduce costs and improve the use of renewable energy (Wang et al. 2015), following the development goals from United Nations (Leal Filho et al. 2021). These local networks have limited resources to invest in many technologies as they count mainly on investments from citizens in the region (Dóci et al. 2015). No study has provided yet an overview of SES, highlighting their potentials and drawbacks, facilitating their prioritization by SECs.

This paper addresses this gap by discussing smart options according to their potential of contribution to improve services and processes within SECs. The study answers the following research questions:

- What are the options and priorities for smart services applied to SECs in Germany?
- What challenges and opportunities for SECs and IS solutions derive from these options?

Figure 1 depicts the research agenda and expected outcomes. After an expert group defined the options of smart services and four necessary perspectives to assess them, stakeholders were selected according to these perspectives. The options were then assessed (ranking) and discussed (appraisal) during guided interviews. This assessment contributes to the prioritization of smart services and provides an analysis of current uncertainties and potentials in the energy market for smart communities.
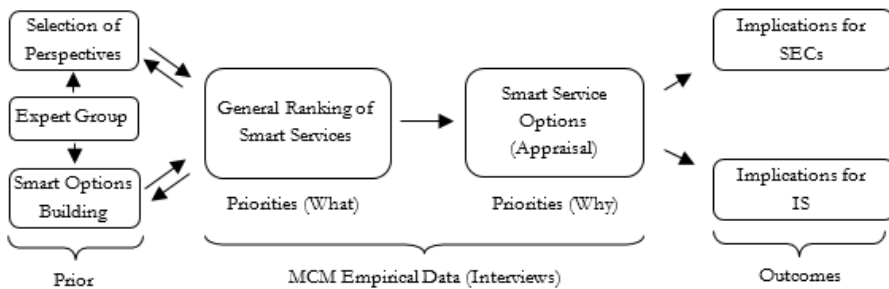
Figure 1: Research Agenda

The remainder of the paper is structured as follows: Chapter 2 explains the methodology and its steps to assess smart options. Chapter 3 discusses the results derived from the analysis, while Chapter 4 provides the main conclusions and contributions for SECs and IS design, as well as insights on further research.

2 Methodology

Researchers, industry and policymakers have assessed risk related to decisions and technologies (Waterstone 1992), fostering studies of their risk perception (Slovic 1987). Adopting technologies require investments and these assessments provide positive and negative aspects of the evaluated objects, reducing risks. Different methods to assess and appraise risk have been developed (Covello and Merkhofer 1993; Horvath and Zuckerman 1993), especially related to technology (Lefley 1997; Stirling 2008). Methods, such as RT Delphi (Gordon and Pease 2006) or cost-effective models (Hubbard 2014), have been applied to the prioritization of technologies. However, these methods fall short when assessing the uncertainties of new technological developments. In this sense, the Multicriteria Mapping Method (MCM) provides an extensive view of potential options.

MCM provides a structured analysis of uncertainties applied to various domains (Stirling and Mayer 2001; Hansen 2010; Shankar et al. 2002). This analysis is based on insights and information from stakeholders of a given industry (Shankar et al. 2002; Donaldson and Preston 1995; Carpenter et al. 2003). Researchers using MCM refer to these stakeholders to provide an analysis with different views and perspectives on the same subject taking into consideration these uncertainties (Hansen 2010; Shankar et al. 2002; McDowall and Eames 2007).

This study used pre-structured options introduced and assessed numerically. Pre-structured surveys are applied to study diversity, defining the objects of analysis beforehand (Jansen 2010). This descriptive analysis aims to prioritize existing options empirically within certain stakeholder groups. A survey is qualitative if it does not count the frequencies of categories, but searches for empirical diversity in the analyzed objects, even if these results are expressed in numbers (Jansen 2010). MCM combines a numeric assessment to rank the options and visualize uncertainty, but focuses on discussions why some options are considered more relevant.

The analysis of this paper follows the steps suggested by Coburn (2016) (see Figure 2). MCM provides an online platform[1] to guide the interview process and support researchers in setting up the interview environment, allowing stakeholders to understand the pre-defined options and move along the research steps. A prior preparation phase took place to define these options and the stakeholder groups by inviting experts in the field to discuss and define the options for SECs.

Step 1 Choose Options		**Step 2** Define Criteria		**Step 3** Assess Scores		**Step 4** Assign Weights		**Step 5** Calculate Ranks
Define the options that are the subject of the analysis.		Interviewees define the criteria for assessment.		Interviewees assign each option a pessimistic and optimistic score.		Interviewees assign a weight to each criterion.		The final rankings are calculated and analyzed.

Figure 2: Research Steps from MCM
Source: Stirling and Mayer (2000)

2.1 Selection of Stakeholders (Perspectives) and Smart Options

Ten experts related to the energy market and IS field discussed and developed a list of smart options that affect SEC performance, as a list of smart services for SECs was not found in the existing literature. The expert group included leaders and representatives from energy (three) and IS-related (three) research institutes, energy communities (two), and software companies (two). They defined the smart options based on their expertise, focusing on services that can be improved using current technologies (see Table 1) and the stakeholder groups (see Table 2) to combine different perspectives on the topic. There was no overlap between the experts and the stakeholders. The options below are coded in three-letter acronyms for later visualization and discussions.

---

[1] Multicriteria Mapping - https://www.multicriteriamapping.com/

Table 1: Options and Descriptions for Smart Services

Option	Description (D) and Contribution (C) to SECs
Applications based on Measured Data (AMD)	(D) User behavior information and usage anomaly ground the development of various applications, such as gamification (power savings comparison), individual billing per device, etc.
	(C) Provision of data visualization to show consumers their exact power consumption and provide forecasts.
Peer-to-peer Trade (P2P)	(D) Private individuals, small businesses and producing companies trade electricity. Consumers, producers and storage facilities are networked to communities and trade locally generated electricity with each other.
	(C) Development of trading platforms.
Selection of Energy Mix (SEM)	(D) Different systems are combined and the consumption profile transparently connected with generation capacities. This improves the location planning of companies as they could define it based on preferences in the energy mix from local producers.
	(C) Consumers select their energy mix systematically.
Proof of Origin (POO)	(D) Electricity is transformed from a commodity to an emotional product by proving when and where it comes from.
	(C) Information regarding less-burdened networks is provided and improves the local matching of supply and demand.
Consumption and Production Optimization (CPO)	(D) Consumption and production could be as close as possible to local communities. IS could support the timetable optimization from flexible producers, consumers and energy storages based on very accurate forecasts and equipment management.
	(C) Timetable optimization from flexible producers, consumers and energy storages based on accurate forecasts and equipment management.
Virtual Power Plants (VPP)	(D) SECs aggregate their flexibility to market their surpluses directly.
	(C) Development of a virtual power plant.
New Tariffs (NTA)	(D) New flexible tariffs (dynamic fares) are adapted and provided to users (prosumers and flexible consumers).
	(C) Development of an incentive system to relieve the local power grid and balance the community's residual load, increasing the local matching of production and consumption.
Investment Opportunities (IOP)	(D) People living on low-invested land and in rented houses/flats could participate financially and generate returns through investments.
	(C) Investments from users are part of the electricity costs. Consumers gradually buy shares of a production plant and participate in the revenue, while SECs invest in production and storage as needed.

According to the reality of energy communities in Germany, the expert group selected four perspectives (stakeholder groups) and indicated several stakeholders to take part in an interview session, which lasted between 60 and 90 minutes. In total, 15 stakeholders participated in the study.

Table 2: Group of Stakeholders (Perspectives)

Perspectives / Stakeholders	Participants
Energy Cooperatives	5
Municipal Utilities as Energy Suppliers	5
Energy Providers with New Disruptive Business Models	3
Technology/Software	2
Total	15

2.2 Assigning Scores and Weights based on Criteria

During the interviews, participants were encouraged to create up to three criteria and assess each option according to them. This allows a degree of freedom for stakeholders to indicate the aspects that are important to them when assessing the given options, as the group consists of different expertise. The criteria were grouped into five topics: (1) perspective from consumers (costs and acceptance), (2) external factors (feasibility, regulatory requirements), (3) level of innovation, (4) economical and (5) ecological aspects. Subsequently, participants assigned a pessimistic and an optimistic score to each option on a scale from zero to 100, and weighted each criterion to improve the analysis of uncertainty.
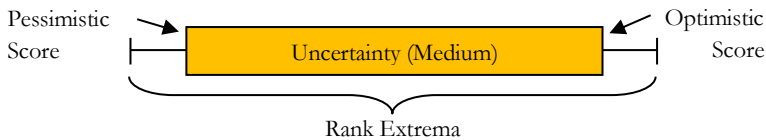
Figure 3: Chart Analysis

This assessment produced a chart and Figure 3 depicts how it displays the results. Options can rank high or low and the difference between the optimistic and pessimistic scores reflects the level of uncertainty. For that, the medium of the scores was considered. The highest and lowest scores are reflected in the extrema line.

3   Results

The general chart provides an overview of how all options are ranking considering their medium values. Stakeholders ranked higher the option Applications based on measured data (AMD), while they were more pessimistic regarding the selection of energy mix (SEM). P2P Trade has a higher level of uncertainty, indicating disagreements among them regarding the developments of such a trading scheme.
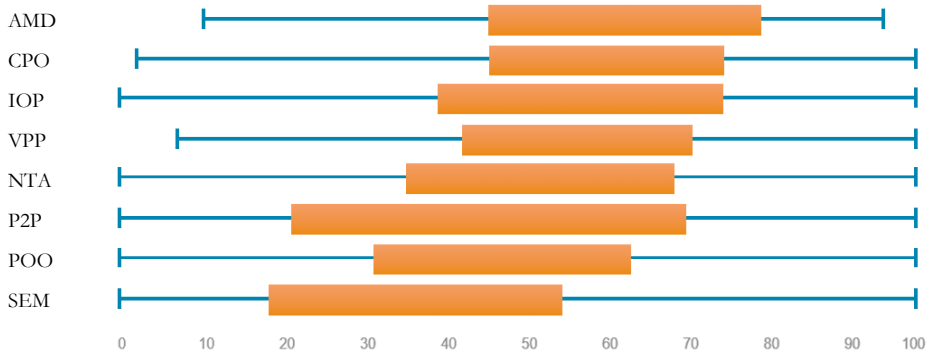
Figure 4: General Ranking of Options (n=15)

Despite the indications of this first chart, a view according to the perspectives highlights the differences in stakeholder groups (see Figure 5). For example, technology-related stakeholders are less optimistic about the optimization of consumption and production (CPO) than the other groups. Additionally, Municipal Utilities are more positive about the option of Virtual Power Plants (VPP).
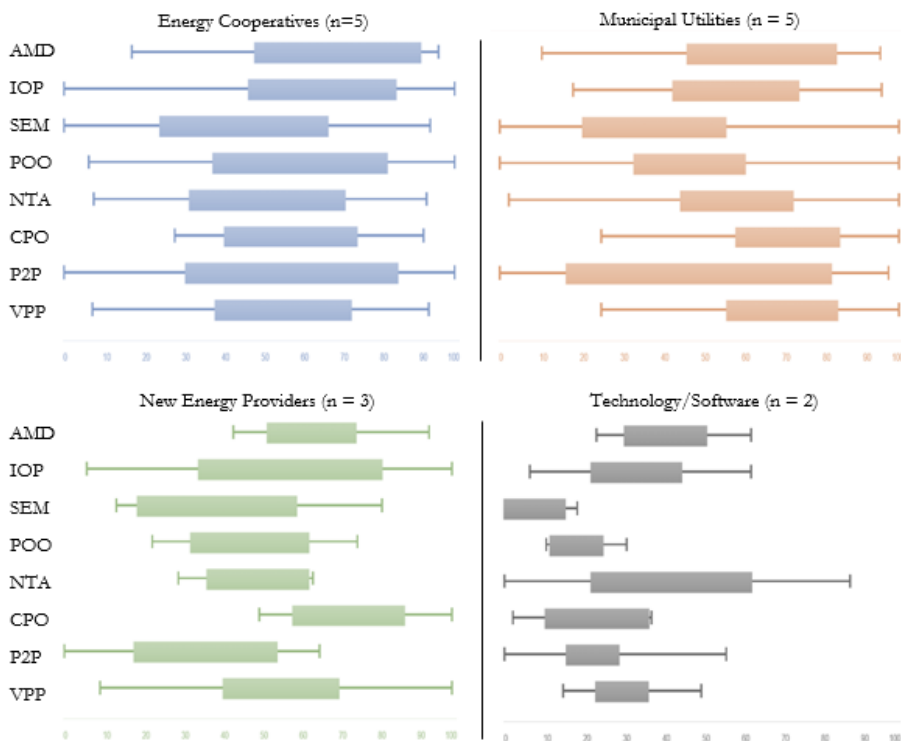
Figure 5: Ranking of Perspectives: Stakeholder Groups.

The observed high level of uncertainty derives from the optimistic and pessimistic scores, assigned and justified by the stakeholders. Besides the numerical assessment, the reasons for the given scores were discussed, as shown in the table below.

Table 3: Optimistic and Pessimistic Views on the Options

Opt.	Optimistic View	Pessimistic View
AMD	- Smart meters produce data for the predictive models of consumption to provide an appropriate supply. - The option provides consumption indication for users.	- Data transfer between gateway and terminals is not yet standardized. - Different interfaces make access non-discriminatory and only companies with the same technology can act as the provider. - Data protection concerns limit the analysis of measured data
IOP	- Participation from citizens increase acceptance and accelerate the energy transition.	- New regulation concerning direct transactions in the market brings concerns that no return on equity investments would payout.

		- SECs must receive support from local stakeholders.
SEM	- Relevant option for the future, but the technical feasibility is very difficult in contrast with the benefits.	- Most customers find it sufficient to obtain green electricity via certificates. - Consumers have emotional attachments to producers and types of energy production.
POO	- Strongly related to P2P trading as the origin is clear in such transactions.	- Green electricity certificates are rather opaque and guarantees of origin need to become more accurate. - Most customers are sensitive to price and do not understand the issue.
NTA	- Uncertainty whether incentive systems with variable prices lead to a behavioral adjustment.	- Relocation of power consumption is difficult for many consumers in private and commercial areas. - Electricity would continue to be consumed when needed, without short-term price elasticity. - A necessary reform of network charges might not take place soon.
CPO	- Seen as the main reason for starting a community. - Balancing generation and consumption done at a regional level.	- It is necessary to define what to connect and record. - Privacy issues can hinder the implementation. - Dependent on the application based on measured data (AMD) and smart meters.
P2P	- Considered the future of the energy market. However, it requires a regionalization of trade and marketplaces.	- Current market is too complex and not transparent. - A community could be reached through a pooling of actors and, therefore, be organized in a common control group without real P2P trading. - A community in a control group is already working today, so true innovation could surge from the emersion of a genuine regional marketplace.
VPP	- For a community, the offering of flexibility is interesting. - The option is reasonable from the physical point of view and logical for the network.	- There is a lack of a clear framework to market it locally. - The individual producer or consumption lacks expertise.

The indication of pessimistic and optimistic aspects also included challenges and potential applications for the options. New challenges concern AMD, such as local injection peaks or high withdrawal peaks due to e-mobility. Nevertheless, smart meters help to predict such consumption. According to stakeholders, the benefit of cooperative electricity could surge through "add-ons" after refinement of the electricity product. Modular product architecture supports the development of interchangeable options (Dahmus et al. 2001). However, technology-related stakeholders are less optimistic because of the lack of standardization. In addition, data protection could hinder such analysis, which goes in line with recent data-protection concerns regarding the deployment of smart meters (Erkin et al. 2013).

Regarding IOP, the participation from citizens is relevant, but requires support from local stakeholders, municipal utilities, investment banks, government, etc. New

energy providers are concerned with return on investments due to new regulations of the direct market. This affects the uncertainty of this option, despite the relevance of citizen participation to finance renewable energy in Germany (Yildiz 2014).

SEM ranks low as stakeholders believe customers are satisfied with current certificates for green electricity. Its technical feasibility is difficult, and consumers are, sometimes, emotionally attached to certain types of electricity. Despite the willingness of energy cooperatives to source their electricity from renewable energies, if that incurs higher electricity costs, these would have to be justified. In this line, POO performed differently in the stakeholder groups as they differ in the level of concern from consumers regarding energy origin. However, energy consumers, when asked to make an active choice between a green and a standard energy provider, choose mostly a green program (Hedlin and Sunstein 2016).

Concerning NTA, participants differed on its impact on behavioral adjustment. Many interviewees claimed electricity would continue to be consumed when needed, diminishing the chances for short-term price elasticity. Nevertheless, they claim CPO is necessary to start an SEC. Efforts should focus on balancing generation and consumption as much as possible at a regional level and should also be network-optimized. New energy providers claim installations should be built where the consumption is located, defining what to connect and record.

Energy cooperatives indicated the motivation from their members not only economically, but also intrinsically or ideationally towards P2P. However, stakeholders from the municipal utility group are uncertain about the need for such trading as a community in an existing control group might be sufficient and this implementation can be costly. ICT and control systems are necessary to enable P2P energy trading in local energy markets (Zhang et al. 2017).

Participants agree that a market possibility is necessary for communities to act as a VPP. As renewables become more prevalent, the need for local governance increase. Representants from municipal utilities were more optimistic about this option and claimed that, even though the market is not ready yet, the shift to the end consumer's perspective is shaping the energy transition. From the technological aspect, some

studies developed algorithms able to aggregate the capacity of different energy resources (Pudjianto et al. 2007; Ruiz et al. 2009; Pandžić et al. 2013).

The views on the options support SECs to decide on the adoption of smart services, defining priorities and investments based on specific needs and market reality. The next chapter presents the implications for research and practice.

4. Conclusions and Implications for Research and Practice

Stakeholders examined technology-based options that influence processes within SECs in Germany through the MCM method. Eight options were developed by an expert group and stakeholders assessed them, indicating priorities. Participants indicated AMD as a high priority for SECs once it also grounds the development of further smart options based on the application of smart meters. The indication and prioritization of smart services answer the first research question, contributing to future solutions for digital ecosystems platforms in the energy industry. Furthermore, stakeholders discussed optimistic and pessimistic aspects for each option, answering the second research question regarding challenges and opportunities. In addition, they indicated that regulatory challenges, data privacy, and the cost-benefit of available technologies are able to hinder the application or reduce the relevance of some options for German SECs.

The predominance of positive aspects around applications based on smart meter data indicates a potential for research on data generated in SECs as a way to determine optimization practices and balance between energy production and consumption.

Stakeholders pointed out that energy communities need to integrate processes and dispersed data to a high degree, as well as to integrate and coordinate different actors in a cross-organizational environment. Although smart meters are not yet widely used, most options benefit directly from their availability. Taking into account the current resource limitation of the energy communities in Germany, either service platforms for several communities or decentralized architectures seem necessary for realizing SECs. Stakeholders shared their opinion on technological developments, the behavior of electricity consumers and current regulations, supporting SECs in

their strategic planning and providing directions on technological demands for smart services in this industry.

SECs benefit from the development and improvement of smart services based on recent IS technologies and the indication and assessment of SES are able to guide IS designers to prioritize their offerings in the field. However, systems should be designed to allow the future aggregation of new functionalities into a complete service system (Lund et al. 2017). That requires an integration of information systems across different organizations. Furthermore, new systems could assume functionalities that are typically performed by intermediaries, co-evolving towards decentralized solutions matching buyers and sellers (Alt 2018) or, in the case of SECs, matching the prosumers. Adopting innovative technologies, such as Blockchain, can support cooperative principles in marketplaces (Kollmann et al. 2020) and foster this change towards decentralized systems.

IS solutions should address the challenges and potentials of technology implementation to support SECs in the optimization of the community. SECs can benefit from the development and improvement of smart services based on recent IS technologies. Among the various contributions, systems can (1) support to predict demands, manage supply and ground investments; (2) draw relevant analysis based on the data; (3) support the emission of certificates regarding the energy, allowing a selection of energy mix; (4) support relieving the power grid and balancing residual load; (5) support the matching of supply and demand, optimizing production; (6) provide and manage incentives through new tariffs, according to usage; (7) support the virtual trade of generated energy; (8) support smart contracts and peer-to-peer trading.

Furthermore, public policies could support the implementation of such technologies applied to the energy market, influencing how SECs will adapt to the recent regulatory changes.

Although the results of the MCM provide indications, conclusions about stakeholders' preferences should be made with caution due to the small number of interviewees. This also affected the balance between the groups, as the total of

stakeholders per perspective is not equal. The inclusion of large energy companies could improve the assessment due to the risk SECs might represent to their business.

Researchers can use these results and methodology to investigate further the options of smart services and to identify possible demands for new integrated information systems in the energy market. Moreover, further research could use MCM to provide a deeper analysis of the source of uncertainty for each stakeholder group and weighting justifications. As some of the options for smart services are already available, market-related information for these options could be explored in addition to the analysis of this paper. Although the paper focuses on the current scenario of German SECs, communities in other regions can benefit from the analysis.

Acknowledgement

References

Allmendinger, G., Lombreglia, R., 2005. Four Strategies for the Age of Smart Services. https://hbr.org/2005/10/four-strategies-for-the-age-of-smart-services.

Alt, R., 2018. Electronic Markets and current general research. Electronic Markets 28 (2): 123–128. 10.1007/s12525-018-0299-0.

Anda, M., Temmen, J., 2014. Smart metering for residential energy efficiency: The use of community based social marketing for behavioural change and smart grid introduction. Renewable Energy 67: 119–127. 10.1016/j.renene.2013.11.020.

Beverungen, D., Müller, O., Matzner, M., Mendling, J., vom Brocke, J., 2019. Conceptualizing smart service systems. Electronic Markets 29 (1): 7–18. 10.1007/s12525-017-0270-5.

Caramizaru, A., Uihlein, A., 2020. Energy communities: an overview of energy and social innovation: European Commission.

Carpenter, M. A., Pollock, T. G., Leary, M. M., 2003. Testing a model of reasoned risk-taking: governance, the experience of principals and agents, and global strategy in high-technology IPO firms. Strategic Management Journal 24 (9): 803–820. 10.1002/smj.338.

Coburn, J., 2016. Multicriteria Mapping Manual - Version 2.0. SSRN Electronic Journal. 10.2139/ssrn.2873524.

Covello, V. T., Merkhofer, M. W., 1993. Risk Assessment Methods: Approaches for Assessing Health and Environmental Risks. Boston, MA: Springer.

Dahmus, J. B., Gonzalez-Zugasti, J. P., Otto, K. N., 2001. Modular product architecture. Design Studies 22 (5): 409–424. 10.1016/S0142-694X(01)00004-7.

Dóci, G., Vasileiadou, E., Petersen, A. C., 2015. Exploring the transition potential of renewable energy communities. Futures 66: 85–95. 10.1016/j.futures.2015.01.002.

Donaldson, T., Preston, L. E., 1995. The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. The Academy of Management Review 20 (1): 65. 10.2307/258887.

Erkin, Z., Troncoso-pastoriza, J. R., Lagendijk, R. L., Perez-Gonzalez, F., 2013. Privacy-preserving data aggregation in smart metering systems: an overview. IEEE Signal Processing Magazine 30 (2): 75–86. 10.1109/MSP.2012.2228343.

Fazeli, A., Christopher, E., Johnson, C., Gillott, M., Sumner, M., 2011. Investigating the effects of dynamic demand side management within intelligent Smart Energy communities of future decentralized power system. In: 2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies, 1–8: IEEE.

Giordano, A., Mastroianni, C., Scarcello, L., 2020. Optimization Model for IoT-Aware Energy Exchange in Energy Communities for Residential Users. Electronics 9 (6): 1003. 10.3390/electronics9061003.

Gordon, T., Pease, A., 2006. RT Delphi: An efficient, "round-less" almost real time Delphi method ☆. undefined.

Grijalva, S., Tariq, M. U., 2011. Prosumer-based smart grid architecture enables a flat, sustainable electricity industry. In: 2011 IEEE PES Innovative Smart Grid Technologies, 1–6. Piscataway: IEEE.

Hansen, S. F., 2010. Multicriteria mapping of stakeholder preferences in regulating nanotechnology. Journal of nanoparticle research an interdisciplinary forum for nanoscale science and technology 12 (6): 1959–1970. 10.1007/s11051-010-0006-3.

Hedlin, S., Sunstein, C. R., 2016. Does active choosing promote green energy use: Experimental evidence. Ecology Law Quarterly 43 (1): 107–142.

Hillebrand, B., Buttermann, H. G., Behringer, J. M., Bleuel, M., 2006. The expansion of renewable energies and employment effects in Germany. Energy Policy 34 (18): 3484–3494. 10.1016/j.enpol.2005.06.017.

Horvath, P., Zuckerman, M., 1993. Sensation seeking, risk appraisal, and risky behavior. Personality and Individual Differences 14 (1): 41–52. 10.1016/0191-8869(93)90173-Z.

Hubbard, D. W., 2014. How to Measure Anything: Finding the Value of Intangibles in Business. Hoboken: Wiley.

Jansen, H., 2010. The Logic of Qualitative Survey Research and its Position in the Field of Social Research Methods. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, Vol 11, No 2 (2010): Visualising Migration and Social Division: Insights From Social Sciences and the Visual Arts / Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, Vol 11, No 2 (2010): Visualising Migration and Social Division: Insights From Social Sciences and the Visual Arts. 10.17169/FQS-11.2.1450.

Kalkbrenner, B. J., Roosen, J., 2016. Citizens' willingness to participate in local renewable energy projects: The role of community and trust in Germany. Energy Research & Social Science 13: 60–70. 10.1016/j.erss.2015.12.006.

Karunathilake, H., Hewage, K., Mérida, W., Sadiq, R., 2019. Renewable energy selection for net-zero energy communities: Life cycle based decision making under uncertainty. Renewable Energy 130: 558–573. 10.1016/j.renene.2018.06.086.

Kollmann, T., Hensellek, S., Cruppe, K. de, Sirges, A., 2020. Toward a renaissance of cooperatives fostered by Blockchain on electronic marketplaces: a theory-driven case study approach. Electronic Markets 30 (2): 273–284. 10.1007/s12525-019-00369-4.

Leal Filho, W., Azul, A. M., Brandli, L., Lange Salvia, A., Wall, T., 2021. Affordable and Clean Energy. Cham: Springer International Publishing; Imprint: Springer.

Lefley, F., 1997. Approaches to risk and uncertainty in the appraisal of new technology capital projects. International Journal of Production Economics 53 (1): 21–33. 10.1016/S0925-5273(97)00106-0.

Lehr, U., Nitsch, J., Kratzat, M., Lutz, C., Edler, D., 2008. Renewable energy and employment in Germany. Energy Policy 36 (1): 108–117. 10.1016/j.enpol.2007.09.004.

Lund, H., Østergaard, P. A., Connolly, D., Mathiesen, B. V., 2017. Smart energy and smart energy systems. Energy 137: 556–565. 10.1016/j.energy.2017.05.123.

Massey, B., Verma, P., Khadem, S., 2018. Citizen Engagement as a Business Model for Smart Energy Communities. In: 2018 5th International Symposium on Environment-Friendly Energies and Applications (EFEA), 1–6: IEEE.

Mathiesen, B. V., Lund, H., d. Connolly, Wenzel, H., Østergaard, P. A., Möller, B., Nielsen, S., Ridjan, I., Karnøe, P., Sperling, K., Hvelplund, F. K., 2015. Smart Energy Systems for coherent 100% renewable energy and transport solutions. Applied Energy 145: 139–154. 10.1016/j.apenergy.2015.01.075.

McDowall, W., Eames, M., 2007. Towards a sustainable hydrogen economy: A multi-criteria sustainability appraisal of competing hydrogen futures. International Journal of Hydrogen Energy 32 (18): 4611–4626. 10.1016/j.ijhydene.2007.06.020.

Palensky, P., Dietrich, D., 2011. Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads. IEEE Transactions on Industrial Informatics 7 (3): 381–388. 10.1109/TII.2011.2158841.

Pandžić, H., Morales, J. M., Conejo, A. J., Kuzle, I., 2013. Offering model for a virtual power plant based on stochastic programming. Applied Energy 105: 282–292. 10.1016/j.apenergy.2012.12.077.

Paukstadt, U., Becker, J., 2019. Uncovering the business value of the internet of things in the energy domain – a review of smart energy business models. Electronic Markets: 1–16. 10.1007/s12525-019-00381-8.

Pudjianto, D., Ramsay, C., Strbac, G., 2007. Virtual power plant and system integration of distributed energy resources. IET Renewable Power Generation 1 (1): 10. 10.1049/iet-rpg:20060023.

Ruiz, N., Cobelo, I., Oyarzabal, J., 2009. A Direct Load Control Model for Virtual Power Plant Management. IEEE Transactions on Power Systems 24 (2): 959–966. 10.1109/TPWRS.2009.2016607.

Shankar, V., Urban, G. L., Sultan, F., 2002. Online trust: a stakeholder perspective, concepts, implications, and future directions. The Journal of Strategic Information Systems 11 (3-4): 325–344. 10.1016/S0963-8687(02)00022-7.

Slovic, P., 1987. Perception of risk. Science 236 (4799): 280–285. 10.1126/science.3563507.

Spence, D. B., 2019. Regulation and the New Politics of (Energy) Market Entry. Notre Dame Law Review 95: 327.

Steinheimer, M., Trick, U., Ruhrig, P., 2012. Energy communities in Smart Markets for optimisation of peer-to-peer interconnected Smart Homes. In: 2012 8th International Symposium on Communication Systems, Network & Digital Signal Processing (CSNDSP 2012): Poznan, Poland, 18 - 20 July 2012, 1–6. Piscataway, NJ: IEEE.

Stirling, A., 2008. "Opening Up" and "Closing Down". Science, Technology, & Human Values 33 (2): 262–294. 10.1177/0162243907311265.

Stirling, A., Mayer, S., 2000. A Precautionary Approach to Technology Appraisal? - A multi-criteria mapping of genetic modification in UK agriculture. Journal of Technology Assessment in Theory and Practice 3 (9. Jahrgang): 39–51.

Stirling, A., Mayer, S., 2001. A Novel Approach to the Appraisal of Technological Risk: A Multicriteria Mapping Study of a Genetically Modified Crop. Environment and Planning C: Government and Policy 19 (4): 529–555. 10.1068/c8s.

van Dinther, C., Flath, C. M., Gaerttner, J., Huber, J., Mengelkamp, E., Schuller, A., Staudt, P., Weidlich, A., 2021. Engineering Energy Markets: The Past, the Present, and the Future. In: Gimpel, H., Krämer, J., Neumann, D., Pfeiffer, J., Seifert, S., Teubner, T. Market Engineering: Insights from Two Decades of Research on Markets and Information, 113–134. Cham: Springer International Publishing; Imprint: Springer.

Vasily Kupriyanovsky, Alexey Konev, Oleg Grinko, Oleg Pokusaev, Dmitry Namiot, 2019. On the way to the energy Internet: new regulations, business models, economic and technical background. International Journal of Open Information Technologies 7 (3): 60–70.

Vernay, A.-L., Sebi, C., 2020. Energy communities and their ecosystems: A comparison of France and the Netherlands. Technological Forecasting and Social Change 158: 120123. 10.1016/j.techfore.2020.120123.

Wang, H., Abdollahi, E., Lahdelma, R., Jiao, W., Zhou, Z., 2015. Modelling and optimization of the smart hybrid renewable energy for communities (SHREC). Renewable Energy 84: 114–123. 10.1016/j.renene.2015.05.036.

Waterstone, M., 1992. Risk and Society: The Interaction of Science, Technology and Public Policy. Dordrecht: Springer.

Yildiz, Ö., 2014. Financing renewable energy infrastructures via financial citizen participation – The case of Germany. Renewable Energy 68: 677–685. 10.1016/j.renene.2014.02.038.

Zhang, C., Wu, J., Long, C., Cheng, M., 2017. Review of Existing Peer-to-Peer Energy Trading Projects. Energy Procedia 105: 2563–2568. 10.1016/j.egypro.2017.03.737.

Zhou, K., Fu, C., Yang, S., 2016. Big data driven smart energy management: From big data to big insights. Renewable and Sustainable Energy Reviews 56: 215–225. 10.1016/j.rser.2015.11.050

Social Robots in Elderly Healthcare: A Burden or a Gift?

Stefan van den Eijkel, [2] Dorien Foppen-de Graaf, [2] Robbert Schuurmans, [2] Stefan van Genderen, [2] Koen Smit[1] & Sam Leewis[1]

[1] HU University of Applied Sciences Utrecht, Institute for ICT, Digital Ethics, Utrecht, Netherland; e-mail: koen.smit@hu.nl, Sam.leewis@hu.nl

[2] HU University of Applied Sciences Utrecht, Institute for ICT, Utrecht, Netherland; e-mail: name.surname@student.hu.nl

Abstract The healthcare sector is currently under enormous pressure and the COVID-19 pandemic does not improve this situation. The quality of healthcare will be negatively impacted when this pressure continues in the longer term. In 2050 it is expected that a total of 2.1 billion people will be aged 60+ years old. To overcome the increasing demand for healthcare by this age group, various studies are being conducted into various technological solutions, such as social robots. In this study, the Alpha Mini social robot was used in an experiment to research which tasks a social robot could assist with, to reduce the work pressure of healthcare professionals and to help the elderly live longer at their own homes. The experiment was carried out using interviews with healthcare professionals and informal caregivers about the demonstrated Alpha Mini. In addition to the experiment and interviews a survey was sent out to 237 healthcare organizations in the Netherlands to identify the 1) work pressure, 2) daily tasks, 3) social robot experiences, and 4) the features a social robot should have to gather requirements. The experiment failed due to work pressure at the healthcare organization. The survey resulted in 181 respondents. The results suggest that tasks such as reminders, setting alarms and physiotherapy have a great potential to help the healthcare professional in reducing their work pressure and tasks, and the elderly to be able to stay living longer at their own home.

1 Introduction

A serious deficiency in the number of healthcare professionals is becoming a worldwide issue and the current COVID-19 pandemic does not improve this situation (Greenberg et al., 2020; Henkel et al., 2020; World Health Organization, 2020). The healthcare sector is under enormous pressure, and if it endures, it will badly impact the quality of healthcare worldwide (World Health Organization, 2020). The forecast is that in 2030 there will be 1,4 billion elderly people of 60+ years old worldwide and in 2050 this number will increase by approximately 66% to 2,1 billion elderly people of 60+ years old (World Health Organization, 2020). The Netherlands is among the top countries (along with e.g. Switzerland, Germany, France, Austria, Finland, Norway) of Europe in terms of the quality of healthcare rated in a yearly report published since 2005 (Arne Björnberg & Ann Yung Phang, 2019). In line with the global growth of the number of elderly people, the ageing population of the Netherlands also continues to increase (NOS, 2020), which is why it is expected that in 2041 there will be 4.7 million people over the age of 65 compared to the current 3.2 million in 2021. This increase in the coming 20 years will only cause extra demand for healthcare in the Netherlands (Schumacher, 2017), raising the total healthcare costs to 19-31% of the annually gross domestic product (GDP) of the Netherlands (Albert van der Horst et al., 2011). In comparison and for illustrative purposes 31% of the GDP of the Netherlands is the entire GDP of Hungary (The World Bank Group, 2021). To overcome the increasing problem of demand in healthcare, studies are being conducted into various technological solutions such as a smart pill dispenser (Medido, 2021), sensors (Joshi et al., 2014), smartwatch (Vivago, 2021), social robot (Hoorn, 2017), and home automation (Harmo et al., 2005) to help elderly live longer at home. Previous studies have shown that social robots can have a great potential to assist in addressing the current issues in healthcare (Abdi et al., 2018) (Bemelmans et al., 2012), (Kachouie et al., 2014) (Broadbent et al., 2009) (Breazeal, 2011) (Phu & Garbrah, 2020). With the social robot, care could be performed more efficiently and effectively by healthcare professionals (Forlizzi et al., 2004). The social robot could help ensure that elderly persons in healthcare continue to receive good care, adding the possibility for elderly persons to live longer at home (Forlizzi et al., 2004).

S. van den Eijkel, D. Foppen-de Graaf, R. Schuurmans, S. van Genderen, K. Smit & S. Leewis:
Social Robots in Elderly Healthcare: A Burden or a Gift?

515

A social robot is a robot that, through the usage of various technologies, such as speech recognition, face recognition, and emotion recognition, can perform non-physical tasks like providing reminders, providing information like the news or the weather and stimulate physical activity (Joshi et al., 2014). Research shows that a daily structure is very important for elderly people because it provides a sense of tranquillity (Góngora Alonso et al., 2019). It has been demonstrated that social robots were able to assist elderly persons with their daily structure (Góngora Alonso et al., 2018). More specifically, in elderly care, the social robot can assist in tasks such as reminding of medicine usage, act as an alarm, connecting with family and friends, and help with the maintenance of physical activity (Forlizzi et al., 2004). In this study, a social robot is defined as a humanoid robot (Duffy, 2003) in the role of an assistant in healthcare. The social robot is not meant to replace the healthcare professionals, but to assist them with their daily tasks (Góngora Alonso et al., 2019; Robinson et al., 2014). With the assistance of social robots, healthcare professionals will have more time for other tasks (e.g. that focus on safety and hygiene and medication adherence) (Robinson et al., 2014).

The goal of this research is to identify which tasks a social robot can assist with and how the social robot could accomplish that, in providing care for elderly persons in healthcare at home, helping healthcare professionals to work more efficiently and help elderly persons to live longer at home. To achieve this goal, an answer is needed to the following research question: '*How can a social robot help provide care more effectively, so that healthcare professionals can spend more time on tasks that they would like to perform, but now do not have enough time for and help the elderly to be able to live longer at home?*'. This was done by conducting expert interviews and a survey.

The paper is structured as follows: in the Background and Related work, previous research on social robots and elderly care related healthcare will be discussed. Next, in the Research Method section, the used methods will be detailed used to provide an answer to the posed research question. This is followed by the Data Collection and Analysis section where the qualitative and quantitative data and analysis are elaborated. Then, in the Results section, the findings of this research will be presented, which are followed by the Discussion section, the Conclusion section, and lastly, directions for Future Research are presented.

2 Background and Related work

Social robots are becoming more popular among researchers and in practice (Campa & Campa, 2016; Share & Pender, 2018). Various applications of social robots exist in healthcare (Share & Pender, 2018), education (Belpaeme et al., 2018), and hospitality (de Kervenoael et al., 2020). An example of a social robot that has been used in home healthcare observations (Bouwhuis, 2016), is the Tessa Robot (Tinyrobot, 2015). This social robot is designed as a flowerpot. Due to its design, it is small, practical, and affordable. The eyes consist of led lights which it uses as facial expressions. The Tessa Robot can play music, provide reminders, can tell the weather forecast and can ask the user questions. However, the response is limited to "yes" and "no" (Tinyrobot, 2015). Another social robot, the NAO robot, is often used in groups, where talks and exercises are done with the social robot (SoftBank Robotics, 2020). Its design focusses heavily on human interaction through the use of camera's, microphones, speech recognition, and touch sensors. It has the ability to walk, sit, and move its arms and head. There is also the Alpha Mini, which is the same size as the Tessa Robot, but has abilities like walking and moves its arms like the NAO (UB Tech, 2021). The Alpha Mini robot has been released just for a little over a year now (2020), so not a lot of research is conducted with this robot. Therefore, in this study, the Alpha Mini robot will be used. The Alpha Mini is more humanized and has more movement abilities which differentiates it from the Tessa robot. While one could argue that other social robots do exist in practice (e.g., Pepper, Sophia, Asimo), the social robots described in this paper are more suitable for home care by healthcare workers due to their size, practicality and (deployment) costs for both elderly people as well as healthcare organizations.

2.1 Acceptance & Adoption of Social Robots

One of the reasons that social robots have not been widely adopted is that users are not involved during development, causing their requirements and wishes not being accounted for (Turja et al., 2018). Another reason is that a lot of users have never had any or low experience with social robots (Turja et al., 2018). Studies (Flandorfer, 2012; Frennert & Östlund, 2014) show that people with experience with social robots are more positive towards the idea of using them.

S. van den Eijkel, D. Foppen-de Graaf, R. Schuurmans, S. van Genderen, K. Smit & S. Leewis:
Social Robots in Elderly Healthcare: A Burden or a Gift?

517

2.2 Elderly People and Social Robots

The elderly are influenced by the usage of new technology by people in their near vicinity, such as healthcare professionals, family members and friends (Tempels, 2016). Elderly blame themselves if they have issues using new technology and do not want to be a strain to their near vicinity (Tempels, 2016). This behaviour represents an important factor in the adoption of new technology by elderly people. However, this is not the only factor that influences the adoption of the elderly. In total, there are 13 factors to decide if the elderly are positive or negative towards new technology such as the social robot (Tempels, 2016). The 13 factors are as followes: The positive factors are 1) independence, 2) daily life, 3) trust, 4) safety, 5) benefits, 6) ease of use, and 7) observed features. The negative factors are 8) knowledge, 9) privacy, 10) fear, 11) relations, 12) practical doubts, and 13) health and demographic factors. Other research confirm Tempels' findings (Tempels, 2016) for social robots (Alaiad & Zhou, 2014; Robillard et al., 2018). These contributions show that elderly people can be motivated to use new technology if people in their near vicinity assist and motivate them in using and trying it (Tempels, 2016). In general, all these aspects raise the importance of including all stakeholders during the design process and elderly people gaining more experience with social robots by using them with the assistance of others close to them.

3 Research Method

For this research, a mixed-method approach is utilized containing qualitative data (Hennink et al., 2020) and quantitative data (Sofaer, 2002). The Mixed-method approach integrates the data during data collection, analysis, or discussion and allows for the creation of a more holistic view of the problem space.

3.1 Experiment and Interviews

Before the start of each experiment, a semi-structured interview (Qu & Dumay, 2011) was held with a healthcare professional and informal caregivers, in order to gather their opinion on social robots in healthcare. This was used to define a baseline for comparison against the final interview at the end of each experiment. After each demonstration, a second semi-structured (Qu & Dumay, 2011) interview was held with the same interviewee, in order to identify potential benefits and limitations

concerning the usage of the Alpha Mini robot. All interviews utilized an interview protocol and were recorded with the informed consent of the interviewees. The interviews were transcribed and coded in order to identify the possible benefits and limitations of the utilization of a social robot in healthcare. For each interview, two coders coded the transcribed interview separately. Next, the two coders compared and discussed the coding results and combined them into one final version. For the demonstration, healthcare professionals were asked to send a daily structure for each elderly person. The daily structure for each elderly person contains timestaps and actions. An example of a daily structure can be found here. The daily structure was programmed on the robotsindezorg.nl (Interactive Robotics, 2021) platform for each individual elderly person. To demonstrate the Alpha Mini, it was installed at the home of the selected elderly person through a supplier of the Alpha Mini robot. The Alpha Mini robot was used in the home of the elderly person for a period of seven weeks with the assistance of healthcare professionals and informal care givers. During this timeframe, the Alpha Mini robot tried to assist the elderly persons retaining their daily structure trough reminders and personal additions, such as a hairdresser appointment and family visitations.

3.2   Survey

A survey was created and validated, based on the current body of knowledge as well as input from a healthcare professional that was not involved in the social robot experiment, conducted trough a separate interview to identify elderly care characteristics and the usage of social robots. These characteristics included 1) work limitations, 2) work pressure, 3) daily tasks, 4) past experiences with social robots, 5) embodiment preferences of the social robots, and 6) their view on the usage of social robots in healthcare. The survey contained a total of 25 questions divided in six sections: 1) introduction, 2) general questions, 3) tasks, 4) social robot characteristics, 5) social robot and reminders, and 6) social robot appearance. After validation and verification by the healthcare professional, the survey was sent via email to healthcare organizations specialized in elderly care for expert sampling. The reason that expert sampling is chosen for this survey, is because it has a better way of constructing the views of experts in elderly healthcare (Etikan, 2017).

S. van den Eijkel, D. Foppen-de Graaf, R. Schuurmans, S. van Genderen, K. Smit & S. Leewis:
Social Robots in Elderly Healthcare: A Burden or a Gift?

519

4 Data Collection and Analysis

The data collection for this study occurred over a period of three months, between November 2020 till January 2021. The implementation of the Alpha Mini robots was planned for November 2020. The interviews were held from the first weeks in November 2020. The survey was published from December 2020 till January 2021.

4.1 Demonstration of the Alpha Mini Robots and Interviews

In order to demonstrate the Alpha Mini robot, a request of participation was sent to healthcare organizations that were in direct contact with the supplier of the Alpha Mini robot. If a healthcare professional was interested and the elderly person gave their approval, a request was submitted for the implementation of the Alpha Mini robot.

Out of the five requested implementations, only one was successful. The main issues that caused unsuccessful implementations were caused due to misplacement, deterioration of the elderly persons' health, and misuse at the side of the elderly person. Due to this, the results of the experiment are deemed invalid and not taken into account in the results of this study. However, for transparency reasons, this activity is described in the paper.

4.2 Survey

Based on the input of healthcare professionals, a list of healthcare organizations (*n* = 237) specialized in elderly care, nursing homes and elderly home care in the Netherlands was formed. Additionally, the researchers utilized their network to distribute the survey to healthcare organizations such as mentioned earlier in this paper. The data analysis of the survey data was conducted using SPSS v27.

5 Results

In this section, the results from the data collection and analysis are presented and will be further discussed. The results will be divided into four sub-sections: tasks, work pressure, elderly people, and social robot requirements. In total, a total of 181

participants submitted the survey out of which 12 were excluded because they were not healthcare professionals, resulting in 169 valid responses

5.1   Tasks

The category 'Tasks' covers the tasks healthcare professionals could not perform due to lack of time. This also included the tasks the social robot could assist with. The top five tasks that are not performed due to lack of time are (as shown in figure 1): 1) making conversation, 2) listening to music, 3) physiotherapy, 4) helping elderly persons with reminders, and 5) extra tasks (calling the general practitioner or the pharmacy). When asked in the survey, which tasks the robot could assist with, the top five most given answers were: listening to music, making conversation, providing reminders, physiotherapy and preparing medication, as shown in figure 2.
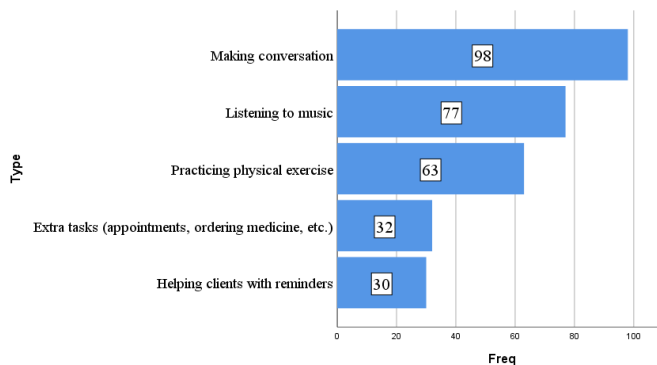
Figure 1: Top 5 tasks that are not performed when there is not enough time left
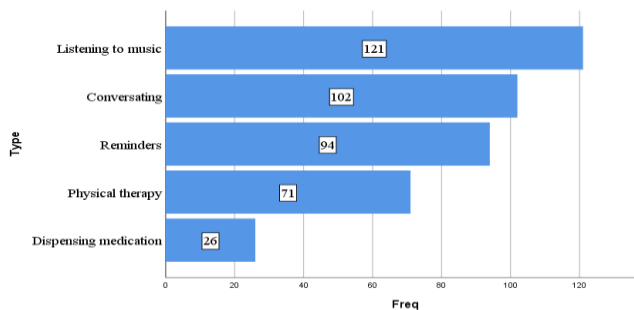
Figure 2: Five most given answers where the robot could assist with

S. van den Eijkel, D. Foppen-de Graaf, R. Schuurmans, S. van Genderen, K. Smit & S. Leewis:
Social Robots in Elderly Healthcare: A Burden or a Gift?

521

5.2   Work pressure

The category 'Work pressure' covers the work pressure experienced by healthcare professionals or informal caregivers. The majority (*n* = 154) of the healthcare professionals indicated that they experience work pressure. In the survey, the following question was posed, "*Do you think a social robot can assist you with certain tasks?*". The majority (*n* = 134) indicated that a social robot could help reducing work pressure.

5.3   Elderly people

The category 'Elderly people' covers whether the usage of reminders could help the elderly live longer at home and what other functions a social robot could help the elderly to live longer at home. In the survey, the majority (*n* = 148) of the healthcare professionals indicated that a social robot could assist the elderly with reminders, as shown in figure 3.
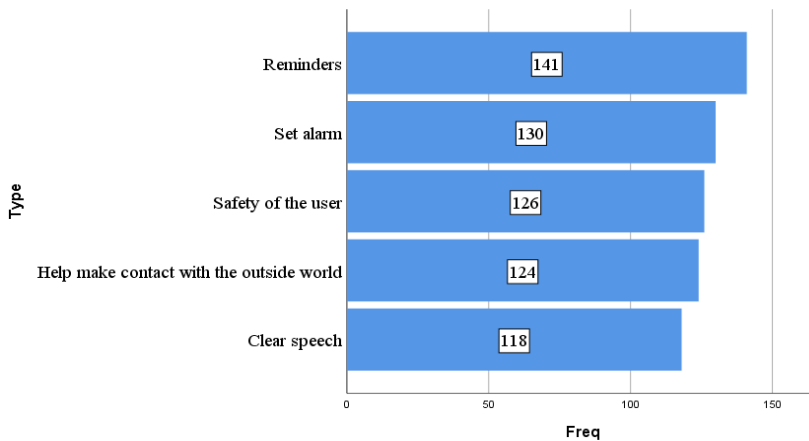
Figure 3: Features that could help elderly people stay home longer

5.4   Social robot requirements

The category 'Social robot requirements' covers the requirements of a social robot from the perspective of the healthcare professionals, as it is important for the social robot to be accepted by the users (Tempels, 2016). The answers from the healthcare workers on the questions about certain requirements and embodiment features of a

social robot are: if the eyes of a social robot mimic human-like behaviour (blinking, look at the elderly person and winking) would scare the elderly person (Yes $n = 21$, No $n = 149$), if the social robot should move its torso and arms when speaking (Yes $n = 142$, No $n = 27$), what type of voice a social robot should have (Male $n = 7$, Female $n = 133$, Sexless $n = 29$), and if a notification sound should be played before a reminder (Yes $n = 126$, No $n = 43$).

6   Discussion

The interviews, experiments and survey results revealed a number of interesting findings for implementation of social robots in elderly healthcare. These concerns, for example, the acceptance of the social robot among the elderly and healthcare workers, but also about the tasks that a robot could assist with. These topics will be discussed in more detail below.

6.1   Tasks and work pressure

It is considered worrisome that, sometimes, certain tasks such as washing elderly people or physiotherapy are not performed due to a lack of time. The majority ($n = 154$) of the healthcare professionals already indicate that they experience work pressure. The research showed that the social robot can reduce workload in the daily structure, such as reducing work pressure, assist with their daily tasks, and further improve the healthcare system. Healthcare professionals indicated that the social robot has the potential to assist them in tasks like listening to music, conversating, reminders, and physiotherapy. This supports Forlizzi's study, which indicates that a social robot can support the care worker in providing care (Forlizzi et al., 2004). Interestingly, these are also most of the tasks that are not performed when there is a lack of time.

6.2   Elderly people

Implementing the Alpha Mini robot could lead to elderly people live longer at home trough features such as reminders, alarms, and ensuring the safety of the user. Healthcare professionals indicated that there is currently no way to check if an elderly person took their medicine. The healthcare professional also indicated that it is important that the healthcare professional or informal caregiver and the elderly person are prepared for the implementation of a social robot to further improve the

S. van den Eijkel, D. Foppen-de Graaf, R. Schuurmans, S. van Genderen, K. Smit & S. Leewis:
Social Robots in Elderly Healthcare: A Burden or a Gift?

523

success of such implementations at the home of the elderly person. This proves the findings of Tempels that people in the near vincinity of elderly people strongly impact the success of an implementation of new technology (Tempels, 2016).

6.3 Social robot requirements

The findings of Turja indicated that it's important for stakeholders to be involved in the development of a social robot to improve adaption of social robots in the healthcare sector (Turja et al., 2018). Our findings showed that healthcare professionals have several requirements for social robots in elderly healthcare. The vast majority ($n = 134$) of healthcare professionals indicated that the social robot could assist elderly people with their daily structure through reminders. Furthermore, the healthcare professionals prefer ($n = 133$) a social robot with a female voice and that the social robot should first play a notification sound before telling a reminder ($n = 126$). Other features such as the eyes, ability to move its torso/arms while talking answers were divided between the healthcare professionals. Some ($n = 21$) did find that, for example, the eyes would be scary for an elderly person but the common idea was that it greatly depends on the individual if it is perceived as scary. Adding the possibility to change or disable certain features, for example, the eyes on an individual basis would be a feature to prevent this issue and to make it more compatible for a specific elderly person.

7 Conclusion

In this research, we aim to answer the following research question: 'How can a social robot help provide care more effectively, so that healthcare professionals can spend more time on tasks that they would like to perform, but now do not have enough time for and help the elderly to be able to live longer at home?' In order to do so, the goal of this research is to identify which tasks a social robot can assist with and how the social robot could accomplish that, in providing care for elderly persons in healthcare at home, helping healthcare professionals to work more efficiently and help elderly persons to live longer at home. Trough the survey data we identified that there is a big potential for the use of social robots in elderly healthcare, especially in its use for assistance in retaining their daily structure trough providing reminders. It seems that the reminders have great potential to help healthcare professionals reducing their work pressure and tasks. Also, it seems that the elderly will be able to

live longer at home, but the data collected does not provide a definitive answer to this. However, the data does show how a social robot could assist healthcare professionals and identifies which requirements should be taken into account for the further development of the Alpha Mini robot.

7.1   Limitations

This research has multiple limitations. The first limitation was caused by Covid-19 restrictions like restricted access to the elderly and impacted the implementation of social robots at the homes of elderly persons. Therefore, the researchers couldn't install, demonstrate and observe the social robot in combination with the elderly, so it had to be outsourced to the supplier of the Alpha Mini robot. This meant that the observation data was all second-hand data gathered through intermediaries. This is a threat to validity and reliability, which resulted in omitting the data from the results and the conclusions in this paper and study. The second limitation was also caused by Covid-19 restrictions. A lot of nursing homes and healthcare professionals did not have sufficient time for the experiment, causing the social robot implementations to be cancelled. The third limitation is related to the survey design, which caused the researchers to not be able to do certain analyses required to prove the trustworthiness and the significance of the results, therefore we solely discuss the descriptive data of the survey in this paper. However, the literature (Moharana et al., 2019; Turja et al., 2018) suggests that there needs to be more attention towards the requirements for healthcare professionals while designing a social robot, which this research still has contributed to and our results provide sufficient insights into. Although it seems that the use of a social robot has potential in elderly healthcare, there is still plenty of research left to conduct.

7.2   Future research

Future research should focus on observations with social robots for reminders and assisting elderly people with their daily structure, which was intended in this study. In future research, elderly people need to be involved in the design process of the social robot, because the elderly are the user of the social robot, next to healthcare workers. This should be directly observed in future research to get more reliable results on the efficacy of a social robot in this context. Future research should also focus on creating architectures based on requirement categories so that a framework

S. van den Eijkel, D. Foppen-de Graaf, R. Schuurmans, S. van Genderen, K. Smit & S. Leewis:
Social Robots in Elderly Healthcare: A Burden or a Gift?

525

can be created. Such a reference framework could then be utilized to address requirements in different situations, making knowledge on requirements for utilizing social robots situationally applicable.

References

Abdi, J., Al-Hindawi, A., Ng, T., & Vizcaychipi, M. P. (2018). Scoping review on the use of socially assistive robot technology in elderly care. BMJ Open, 8(2). https://doi.org/10.1136/bmjopen-2017-018815

Alaiad, A., & Zhou, L. (2014). The determinants of home healthcare robots adoption: An empirical investigation. International Journal of Medical Informatics, 83(11), 825–840.

Albert van der Horst, Frank van Erp, & Jasper de Jong. (2011). Trends in gezondheid en zorg. Centraal Planbureau.

Arne Björnberg, & Ann Yung Phang, RN, B.A. (2019). Euro health consumer index 2018. Health Consumer Powerhouse.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. In Science Robotics (Vol. 3, Issue 21, p. eaat5954).

Bemelmans, R., Gelderblom, G. J., Jonker, P., & de Witte, L. (2012). Socially Assistive Robots in Elderly Care: A Systematic Review into Effects and Effectiveness. Journal of the American Medical Directors Association, 13(2), 114–120.e1. https://doi.org/10.1016/j.jamda.2010.10.002

Bouwhuis, D. G. (2016). Current use and possibilities of robots in care. Gerontechnology, 15(4). 0

Boynton, P. M., & Greenhalgh, T. (2004). Hands-on guide to questionnaire research: Selecting, designing, and developing your questionnaire. In British Medical Journal (Vol. 328, Issue 7451, pp. 1312–1315). https://doi.org/10.1136/bmj.328.7451.1312

Breazeal, C. (2011). Social robots for health applications. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 5368–5371.

Broadbent, E., Stafford, R., & MacDonald, B. (2009). Acceptance of healthcare robots for the older population: Review and future directions. International Journal of Social Robotics, 1(4), 319–330.

Bryant, A., & Charmaz, K. (2012). The SAGE Handbook of Grounded Theory. In The SAGE Handbook of Grounded Theory. SAGE Publications Ltd.
https://doi.org/10.4135/9781848607941

Campa, R., & Campa, R. (2016). The rise of social robots : a review of the recent literature. Journal of Evolution and Technology, 26(1).

de Kervenoael, R., Hasan, R., Schwob, A., & Goh, E. (2020). Leveraging human-robot interaction in hospitality services: Incorporating the role of perceived value, empathy, and information sharing into visitors' intentions to use social robots. Tourism Management, 78, 104042.

Duffy, B. R. (2003). Anthropomorphism and the social robot. Robotics and Autonomous Systems, 42(3–4), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

Eijkel, S. van den. (2021). OSF | Social robots in elderly healthcare: a burden or a gift?
https://osf.io/65zua/?view_only=795dc300c6ab472db0a98c5011f9b850

Etikan, I. (2017). Sampling and Sampling Methods. Biometrics & Biostatistics International Journal, 5(6).

Field, A. (2013). Andy Field - Discovering Statistics Using IBM SPSS Statistics. In Lavoisier.Fr (5e ed., pp. 1–816). SAGE Publications Ltd.

Flandorfer, P. (2012). Population Ageing and Socially Assistive Robots for Elderly Persons: The Importance of Sociodemographic Factors for User Acceptance. International Journal of Population Research, 2012, 1–13. https://doi.org/10.1155/2012/829835

Forlizzi, J., DiSalvo, C., & Gemperle, F. (2004). Assistive robotics and an ecology of elders living independently in their homes. Human-Computer Interaction, 19(1–2), 25–59.

Frennert, S., & Östlund, B. (2014). Review: Seven Matters of Concern of Social Robots and Older People. International Journal of Social Robotics, 6(2), 299–310.

Góngora Alonso, S., Hamrioui, S., De La Torre Díez, I., Motta Cruz, E., López-Coronado, M., & Franco, M. (2019). Social Robots for People with Aging and Dementia: A Systematic Review of Literature. In Telemedicine and e-Health (Vol. 25, Issue 7, pp. 533–540).

Greenberg, N. E., Wallick, A., & Brown, L. M. (2020). Impact of COVID-19 Pandemic Restrictions on Community-Dwelling Caregivers and Persons With Dementia. Psychological Trauma: Theory, Research, Practice, and Policy. https://doi.org/10.1037/tra0000793

Harmo, P., Taipalus, T., Knuuttila, J., Vallet, J., & Halme, A. (2005). Needs and solutions - Home automation and service robots for the elderly and disabled. https://doi.org/10.1109/IROS.2005.1545387

Henkel, A. P., Čaić, M., Blaurock, M., & Okan, M. (2020). Robotic transformative service research: deploying social robots for consumer well-being during COVID-19 and beyond. Journal of Service Management, 31(6), 1131–1148. https://doi.org/10.1108/JOSM-05-2020-0145

Hennink, M., Hutter, I., & Bailey, A. (2020). Qualitative Research Methods (A. Owens (Ed.); 2nd ed.). SAGE Publications Ltd.

Hoorn, J. (2017). De wereld heeft zorgrobots nodig. In Wetenschappelijk Bureau GroenLinks. https://wetenschappelijkbureaugroenlinks.nl/artikel/de-wereld-heeft-zorgrobots-nodig

Huq Khandkar, S. (2009). Open Coding.

Interactive Robotics. (2021). Innovatieve en interactieve robots in het onderwijs en de zorg. https://www.interactive-robotics.com/

Joshi, G. P., Acharya, S., Kim, C. S., Kim, B. S., & Kim, S. W. (2014). Smart Solutions in Elderly Care Facilities with RFID System and Its Integration with Wireless Sensor Networks. International Journal of Distributed Sensor Networks, 2014(8), 713946. https://doi.org/10.1155/2014/713946

Kachouie, R., Sedighadeli, S., Khosla, R., & Chu, M. T. (2014). Socially Assistive Robots in Elderly Care: A Mixed-Method Systematic Literature Review. International Journal of Human-Computer Interaction, 30(5), 369–393. https://doi.org/10.1080/10447318.2013.873278

Medido. (2021). Medido. https://medido.com

Moharana, S., Panduro, A. E., Lee, H. R., & Riek, L. D. (2019). Robots for Joy, Robots for Sorrow: Community Based Robot Design for Dementia Caregivers. ACM/IEEE International Conference on Human-Robot Interaction, 2019-March, 458–467. https://doi.org/10.1109/HRI.2019.8673206

NOS. (2020, September). SER slaat alarm: groot personeelstekort in de zorg. https://nos.nl/l/2337769

Phu, U., & Garbrah, W. (2020). The Impact of Artificial Intelligence on Healthcare for Elderly A Literature Review The Impact of Artificial Intelligence on Healthcare for Elderly A Literature Review Degree programme in Nursing. http://www.theseus.fi/handle/10024/343950

Qu, S. Q., & Dumay, J. (2011). The qualitative research interview. In Qualitative Research in Accounting and Management (Vol. 8, Issue 3, pp. 238–264). https://doi.org/10.1108/11766091111162070

Robillard, J. M., Cleland, I., Hoey, J., & Nugent, C. (2018). Ethical adoption: A new imperative in the development of technology for dementia. Alzheimer's and Dementia, 14(9), 1104–1113.

Robinson, H., MacDonald, B., & Broadbent, E. (2014). The Role of Healthcare Robots for Older People at Home: A Review. International Journal of Social Robotics, 6(4), 575–591.

Schumacher, J. (2017, September). Cijfers: vergrijzing en toenemende zorg | Kennisplein Zorg voor Beter. 22-05. https://www.zorgvoorbeter.nl/veranderingen-langdurige-zorg/cijfers vergrijzing

Share, P., & Pender, J. (2018). Preparing for a Robot Future? Social Professions, Social Robotics and the Challenges Ahead. Irish Journal of Applied Social Studies, 18(1), 4. https://doi.org/10.21427/D7472M

Sofaer, S. (2002). Qualitative research methods. International Journal for Quality in Health Care, 14(4), 329–336. https://doi.org/10.1093/intqhc/14.4.329

S. van den Eijkel, D. Foppen-de Graaf, R. Schuurmans, S. van Genderen, K. Smit & S. Leewis:
Social Robots in Elderly Healthcare: A Burden or a Gift?

527

SoftBank Robotics. (2020). Pepper and NAO in the service of the Healthcare sector. https://www.softbankrobotics.com/emea/en/industries/healthcare

Tempels, J. (2016). De beleving van technologie door ouderen bij hun gezondheidszorg en de rol van hun omgeving [Universiteit Gent]. https://lib.ugent.be/fulltxt/RUG01/002/350/335/RUG01-002350335_2017_0001_AC.pdf

The World Bank Group. (2021). GDP per capita (current US$) - Netherlands | Data. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=NL

Tinyrobot. (2015). Tessa ondersteunt zelfregie, zelfstandig wonen en biedt zorg.

Turja, T., Van Aerschot, L., Särkikoski, T., & Oksanen, A. (2018). Finnish healthcare professionals' attitudes towards robots: Reflections on a population sample. Nursing Open, 5(3), 300–309.

UB Tech. (2021). Alpha Mini. https://www.ubtrobot.com/products/alpha-mini?ls=en

Vivago. (2021). Vivago. ZorgHorloge. http://zorghorloge.nl/

World Health Organization, W. (2020). Environments Decade of Healthy Ageing. World Health Organisation. https://www.who.int/docs/default-source/decade-of-healthy-ageing/full decade-proposal/decade-proposal-fulldraft-en.pdf?sfvrsn=8ad3385d_6

Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

Koen Smit,[1] Matthijs Smakman,[1] Sil Bakker,[2]
Jurgen Blokhuis,[2] Guido Evertzen[2] &
Lars Polman[2]

[1] HU University of Applied Sciences Utrecht, Digital Ethics, the Netherlands; e-mail: koen.smit@hu.nl, matthijs.smakman@hu.nl
[2] HU University of Applied Sciences Utrecht, the Netherlands; e-mail: sil.bakker@student.hu.nl, jurgen.blokhuis@student.hu.nl, guido.evertzen@student.hu.nl, lars.polman@student.hu.nl

Abstract Social Robotics is becoming more relevant for the healthcare sector as an increasing amount of research and development is invested by researchers and practice. One research area where additional research would help the acceptation and adoption of social robots is intramural care where people with dementia live. The current body of knowledge on this topic can be described as nascent. In this study, we add to the body of knowledge regarding the design and enactment of social robots like the one used in this study, the Tessa robot, with the goal to improve acceptation and adoption of social robots in dementia care. To do so, we conducted a case study at a healthcare organization, featuring semi-structured interviews, observations and talking mats. During this case study, an experiment was carried out in which a Tessa robot was used in intramural care with three clients suffering from dementia. The most important finding of this study is that for the robot to be accepted and effective it must be implemented properly in the existing healthcare processes, otherwise it might serve as a companion, but will not relieve the workload of healthcare workers.

1   Introduction

The overall life expectancy in the world is increasing steadily for the last decades (Roser, Ortiz-Ospina, & Ritchie, 2019). For example, since the 1950s, average life expectancy in The Netherlands has increased by 10 years from 71 to 81 years (van der Aalst, 2019), which is similar to other western countries. An aging population introduces new social challenges related to physical and mental issues, one of these issues is dementia. The estimated number of people with dementia worldwide is ~ 46.8 million people. Every 3.2 seconds someone is diagnosed with dementia (Bouwhuis, 2016). It is estimated that this number will double every 20 years and by 2030 there will be ~ 75 million people living with dementia worldwide (Domínguez-Rué & Nierling, 2016). At the time of writing this paper there is no medicine available that completely supresses or eliminates dementia symptoms nor does a cure exist. However, there are forms of treatments that can partially suppress the symptoms and thus improve the quality of life of a client. An example of this is Animal Assisted Therapy (AAT), in which animals are used in therapy sessions (Downes, Dean, & Bath-Hextall, 2013).

Research indicates that, in some countries, there are not enough health-care workers (HW's) available. Which, in turn, could result in a decline in the quality of care for people with dementia (Domínguez-Rué & Nierling, 2016). Due to the shortage, a majority of HW's (~60%), has to work extra shifts, often with fewer people (van der Aalst, 2019). The number of on-call workers is also increasing due to these shortages. More than two thirds of the on-call workers indicate that the workload has increased in the past year. This puts the quality of care under pressure, and it affects the mental and physical well-being of HW's. Already 71% of HW's indicate experiencing more stress (van der Aalst, 2019).

Social robots could support HW's and potentially aid people with dementia. In this paper, the definition of a social robot is used: "*a physically embodied robot that communicates autonomously with humans and other autonomous physical robots in a way that is conducive to its own goals and those of its environment*" (Duffy, 2003). In the situation of an aging population and shortages in health care workforce, social robots can attribute to the well-being of the aging population suffering from dementia (from here on referred to as: clients) by supporting and taking over certain tasks from HW's. Social robots can be used, for example, for monitoring clients and for the use of therapy

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

531

(Valentí Soler et al., 2015a). The sensors of social robots can respond to changes in the environment (movements, sound), allowing them to interact with clients. Potential advantages that complement HW's are, for example, that robots can work longer and take up less space and need less care (Valentí Soler et al., 2015a).

The Dutch government made funds available to improve the quality of care in nursing homes, with funds rising to a structural amount of € 2.1 billion per year in 2021 (van der Aalst, 2019). Part of this budget is spent on home automation, which includes social robots. This financial stimulus led to multiple nursing homes in the Netherlands that started experimenting with social robots in their healthcare processes. One of these social robots is called Tessa, a flowerpot like robot (Robotzorg, 2021). The Tessa robot is a social robot that talks and is designed to support the daily structure of people with a cognitive disability and to provide suggestions for certain activities to them (Robotzorg, 2021).

In the light of the aforementioned social challenges, Tessa could assist and take over care tasks from HW's, thereby reducing current high workload's, and improving the quality of care. Although social robots hold great potential, there are still challenges regarding the acceptance and effectiveness of social robots in dementia care. Therefore, this exploratory study aims to answer the following research question: '*How can the acceptance and effectiveness of the Tessa robot for both HW's and clients with dementia be improved?*'

This paper is structured as follows. In the next section the background and related work are discussed. Then, in section three, the utilized research methods are detailed. In section four, the data analysis is explained. This is followed by the results in section five. The paper is concluded by presenting the discussion in section six, which is followed by the conclusions in section seven. Finally, future directions for future research are presented in section eight.

2 Background and Related Work

In this section, we further ground the potential of social robots in the context of HW's that care for elderly people with dementia. Acceptance and involvement of stakeholders is one of the key aspects when responsibly designing and implanting technology (Friedman, Kahn, Borning, & Huldtgren, 2013). Earlier research has

shown that usefulness, adaptability, enjoyment, sociability, companionship, and perceived behavioural control are important variables in social robotic acceptance (de Graaf & Ben Allouch, 2013). Furthermore, to determine whether and how social robots actually meets real-world needs, it is important to study these robots in ecologically valid settings (de Graaf, Somaya, & van Dijk, 2016). In general, people have positive attitudes towards social robots and are willing to interact with them, according to a large, standardized study with a combined sample of over 13,000 participants (Naneva, Sarda Gou, Webb, & Prescott, 2020). Research in elderly care reveals that the attitudes of elderly towards social robots are more often positive than negative (Savela, Turja, & Oksanen, 2018), the same holds for people with dementia (Whelan et al., 2018). These findings are promising for our study as we aim to experiment with social robots in the natural environment, whereby the perspective of stakeholders is important for the successful implementation of this new technology.

2.1   The needs of clients suffering from dementia

The physical needs of people with dementia are more often met than the emotional needs (Miguel et al., 2016), (Visser & Vandemeulebroucke, 2018). Personal contact with care workers or family has a positive effect on people with dementia, it reduces the chance of loneliness and depression (Miguel et al., 2016). Other studies show that, by increasing the daily structure and amount of personal contact, dementia symptoms are less likely to develop further (van Beek, Frijters, Wagner, Groenewegen, & Ribbe, 2011), (Mordoch, Osterreicher, Guse, Roger, & Thompson, 2013). With these contributions, we would like to bring attention to the influence of dementia associated motivational and emotional disorders on the positive affective state that interactions with social robots are able to prompt. Social robots could utilize existing methods that support these issues such as exercises, images, and sound (Qwiek, 2021).

2.2   The needs and acceptance of (in)formal HW's

The (technological) support that is most used in dementia care is often limited to a mobile or desktop interface (Alnanih & Ormandjieva, 2016). The main focus of these tools is to help entertain and keep people with dementia active. However, there are few supportive solutions that relieve the (in)formal HW's (Tyack & Camic, 2017),

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

533

(Cheng, 2017). Until now, social robots have not been used extensively in healthcare (Turja, Van Aerschot, Särkikoski, & Oksanen, 2018).

For robots to be accepted and used, it is essential that HW's are included in the design process of the social robots (Góngora Alonso et al., 2019a), (Moharana, Panduro, Lee, & Riek, 2019), (Robillard, Cleland, Hoey, & Nugent, 2018). Compared to elderly people investigated in different studies, HW's are generally less positive about using social robots in their context, however, after exposing HW's to a social robot, positive attitude towards a social robot seems to be higher (Savela et al., 2018). Although, not all HW's have a positive attitude regarding social robots; the acceptance of robots in healthcare seems to be strongly linked to the HW's moral considerations (van Kemenade, Hoorn, & Konijn, 2018).

In the context of this research, social robots are studied to relieve HW's from repetitive tasks such as reminding clients with regards to, e.g., exercises, food or drink moments, and social activities, which matches the work of (Góngora Alonso et al., 2019; Valentí Soler et al., 2015). This also stems from the fact that healthcare organizations are not seeking to replace HW's by social robots, but to lower work pressure for them (Valentí Soler et al., 2015).

2.3   Social Robot: Tessa

Previous studied related to the Tessa robot in a dementia care context have been conducted. In 2019, (Casaccia et al., 2019) set up an eWare platform for the Tessa robot based on qualitative and quantitative data collection from dementia clients, (in)formal HW's and healthcare managers. The functionalities of the Tessa robot seem to meet the needs of HW's and people with dementia (Casaccia et al., 2019), (Miguel et al., 2016), (Johnson et al., 2020). For example, a recurring theme is daily structure, something that all stakeholders and dementia clients in particular benefit from (van Beek et al., 2011), (Mordoch et al., 2013), (Miguel et al., 2016).

Scientific research has yet to be conducted into the effectiveness of the Tessa robot. As mentioned in the previous section, the increase in the number of people with dementia (caused by aging) will not decrease (Domínguez-Rué & Nierling, 2016), which means that research on technological support, such as the Tessa robot, is necessary. Effectiveness of Tessa will be measured by taking into account both the

interaction between the robot and the client, and the reduction of workload for HW's.

3   Research Method

The research field of social robotics related to the application of social robots in nursing homes is relatively nascent; either relatively small qualitative samples are analysed, or large meta-level reviews are conducted (Savela et al., 2018; Whelan et al., 2018). When a research field is nascent, new constructs still need to be identified and relations between these constructs should be established (Edmondson & Mcmanus, 2007), which is often characterized by exploratory (qualitative) research. To achieve this, a case study at a nursing home was executed. Case study research is a technique that can be used to explore a broad scope of complex issues, particularly when human behaviour and social interactions are of importance (Pervan & Maimbo, 2005). This study comprised a holistic case study approach (Runeson & Höst, 2008), focusses on the context of nursing homes for people with dementia (permanent and closed care facility). This way, the intervention (Tessa) can be evaluated in the natural context for which it is designed. Also, when the boundaries between the intervention (Tessa) and the context are not clearly evident, multiple sources of empirical evidence are used (Pervan & Maimbo, 2005), which is taken into account in our case study approach.

The people with dementia in our case study, gave consent to participate in this research themselves, as well as via their responsible healthcare worker and their direct family members. In this cases study we will focus on both the people with dementia (clients) as well as the HW's related to these clients in combination with the Tessa robot. During this case-study, data was collected using three different methods: 1) semi-structured interviews, 2) naturalistic observations, and 3) Talking Mats. All data was collected in The Netherlands between August 2020 and December 2020.

3.1   Case Study

This case study was conducted in close cooperation with a large healthcare organization in the Netherlands that operates multiple facilities for different types of healthcare with a total of 777 FTE's, which was based on convenience sampling. The clients that participated on our study were selected and approached by the

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

535

healthcare organisation. To study the social robot and its users in the most natural context, each client was provided a Tessa robot in their living room for twelve days.

3.2   Semi-structured interviews

The first method of data collection in this case study are semi-structured interviews, which allow the research team to gather qualitative data during two phases. The first phase is utilized to elicitate requirements from HW's regarding the workings of the Tessa robot. The second phase is utilized to gather data on the acceptance and the effectiveness of the Tessa robot. During the course of eight weeks, four digital and seven physical interviews were conducted with HW's, which were selected in cooperation with the management of the organization, taking into account that the HW's are connected to the selected clients in the experiment and known each other. The goals of the interviews with HW's were to learn about the various daily activities and the needs of both the HW's and the clients. Whilst conducting the interview an interview protocol was utilized, which increases repeatability and comparability of the results (Castillo-Montoya, 2016). Also, different protocols were used before and after the experiment. The protocol's themes and corresponding questions focused on: 1) current experience with social robots, 2) perceived value of social robots, 3) added value of Tessa in the context of the HW, 4) involving clients and HW's in the development proces, 5) pro's and con's of Tessa after the experiment, 6) preceived value for the client, and 7) experienced values for HW's. The questions were open-ended, allowing for discussion and relevant deviation when deemed nessesary by the researchers. The interviews have been audio recorded and transcribed after the interview, for which all participants provided verbal consent. The average length of the interviews was 28 minutes, the longest being 44 minutes and shortest being 8 minutes.

3.3   Observations

The second method of data collection in this case-study are naturalistic observations. The goal of these observations was perceiving and recording the effects that the Tessa robot had on the clients. One observation has been conducted per client. To ensure reliability, it was made sure that all clients had been using the Tessa robot during a timeframe of equal length (twelve days). Each observation was held on the Monday of the second week of use, after seven days of usage, where two observers independently filled in the protocol. All three duo's of researchers were different as

well to mitigate potential bias. Every client had a daily schedule with activities that take place on each day. This schedule was manually programmed onto each Tessa-robot by the research team, based on the schedule received by the responsible HW. Programming the robot was performed using mytinybo.io, which is a platform used to synchronize commands with the robot, via Wi-Fi. Based on this, it would announce these activities at predetermined times, during the experiment. The announcements start at 8AM and end at 10PM, with at least one announcement every hour. The observations were organized during 10AM and 2:30PM. During the observations, a predetermined observation protocol was used. The protocol was created using the Interactive Behaviour Codification System (Andrés, Pardo, Díaz, & Angulo, 2015). This system is used for grading the interactive behaviour between humans and robots, which makes it very suitable for the observations as part of this case study. The form consists of eight distinct categories (perceived emotions, proxemics, gaze, etc.), which are then each divided in sub-categories (such as joy, focused gaze, etc.). For each observation moment, two research team members individually filled in the protocol, to improve reliability.

3.4 Talking Mats

Talking Mats is a technique which helps people with cognitive disabilities communicate during interview sessions (Murphy, Gray, Cox, & Joseph Rowntree Foundation, 2007). This method is easy to prepare, suitable to be conducted during the covid-19 pandemic and meets the criteria for better communication for the clients. Talking Mats consists of three kinds of cards:

1. Subject – what the conversation is about and what the options are paired to. In this case: the Tessa robot.
2. Options – in this case, the announcements, which Tessa made to the clients. A total of eight images have been made and used, one for each type of announcement.
3. Scale – the clients were able to communicate their opinion of the different options by pairing the cards with the respective grades, which reflect their emotions and feelings about a subject: negative, neutral or positive.

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

537

The cards with the different options were then presented to the client. The session was conducted in cooperation with a HW, who was experienced in communicating with the clients. Due to covid-19 restrictions, the cards were not placed by the client themselves since this would cause unnecessary physical contact. After a card was placed at one of the three gradings, a picture was taken to capture the results. This was repeated for each subject.

4   Data Analysis

All interviews were recorded, transcribed and coded independently by three researchers. This process was conducted redundantly to eliminate coding bias as well as to improve the validity of the results (Armstrong, Gosling, Weinman, & Marteau, 1997). Coding was performed in AtlasTI. To analyse the transcriptions, the Toulmin's framework was utilized (Toulmin, 2003), which consists of three elements: 1) Claims, 2) Grounds, and 3) Warrants. Finally, all codes were merged, and an assessment of the intercoder-agreement was made. Where no agreement was initially reached, the coders partook in a session where the codes were discussed, and consensus was reached, also described by Campbell as a "negotiated agreement" (Campbell, Quincy, Osserman, & Pedersen, 2013).

To study the interaction between the client and the Tessa robot, an observation-protocol (Andrés et al., 2015) was used. The protocol utilizes the following variables that are recorded by each observant; 1) type of instruction, 2) emotions, 3) proxemics, 4) gaze, 5) communication, 6) facial expression, 7) body gestures, and 8) interaction with the robot.

The data resulting from the observations was different than expected, because clients often were not in the room when the robot made an announcement. Therefore, in addition to the interaction with the robot, the presence of the client when the robot made an announcement was measured in percentages and included as well. This created a new angle to be explored during the interviews with HW's after the experiment.

The recorded responses of the clients resulting from the Talking Mats method were compared. This allowed for a comparison of clients' views. The views served as additional input for the interview protocol used for the interviews that were held after the experiment with HW's. The questions discussed as subjects during the

talking mats sessions with the clients were based on the full range of pre-programmed announcements executed by Tessa during the experiment, which in turn were based on the personal day-to-day programmes.

5 Results

In the following sections we will present the results of each technique utilized seperatly.

5.1 Semi-structured interviews (before the experiment)

This sub-section focuses on the interviews before the experiment. Of the 1062 total codes, the coders independently reached an agreement of 871 eligible codes. Which means an initial percentage match of 82.02%. In the observer agreement model of Landis and Koch (Landis & Koch, 1977) this falls into the "Almost Perfect" category, making the coding process reliable.

Prior to the experiment, separate interviews were held with five HW's working at the healthcare organisation. Additionally, an interview was held with a HW that had over two years of working experience with the Tessa in extramural dementia care. The knowledge gained from these interviews has been translated into functional requirements for the experiment and Tessa robot. Below we present the predominant requirements mentioned in the data.

Input requirements:

- (In)formal HW's must be involved in the set-up of the robot because a personalized Tessa gets more response from the client compared to a standard Tessa;
- The physical and mental condition of the client must be good enough that he or she can hear and understand the Tessa properly. Formal HW's must be included in the selection process to derive suitable clients for the experiment;
- Formal and informal HW's must be instructed in the form of a training and / or by providing a manual. This enables HW's to become properly prepared and self-reliant regarding how to handle the Tessa;

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

539

- Other minor requirements that detail the day-programme of a particular client. For example, one of three clients required announcements to smoke, while another client required announcements about coming for lunch as the client usually forgets that particular moment.

5.2 Observations

In total, three researchers made observations during four and a half hours in the living room of the client. This method aimed at measuring the effectiveness of the Tessa robot by observing the response of the client.

The Tessa robot succeeded in provoking a reaction out of one of the three clients. This was concluded because one of the clients responded twice to an announcement given by the robot. In an average of 88% of the announcements intended to relieve the HW's, the HW's themselves had already verbally given an announcement to the client before Tessa's announcement had triggered. As a result, both the client and the HW were regularly not in the room when Tessa made an announcement. On average, the clients were only present during roughly half of the announcement made by Tessa. Furthermore, only client 1 reacted 2 out of 6 times, in a neutral sense. The results of this process are presented in Table 1.

Table 1: Observation results

Client	Present	Absent	Attendance %
1	6 (no reaction)	5	55%
2	5 (no reaction)	2	71.4%
3	2 (2 reactions)	6	25%
Total	**50%**	**50%**	

In the cases where the clients were present at the announcements, two of the three clients did not give a verbal response to the announcements from the Tessa robot. One type of announcement required verbal confirmation (the request for music) from the clients before the Tessa could proceed with the action. When the client gave an answer, the robot's microphone activated too late, which happened twice in the situation of client 3.

5.3 Talking Mats

The talking mats resulted to be effective in measuring the opinion of the clients; the research team managed to address all topics and gather a value for each of them. The results of the Talking Mats were as follows: 5.6% of the questions were answered with a negative value. The remainder of values registered consist of 47.2% neutral responses and 47.2%positive responses. Figure 1 shows the results per client. The horizontal axis represents the subject of the Tessa activity. Not all clients smoked cigarettes. Not everyone had to set the table either, so the number of questions per client differed by a maximum of two (minimum 11, maximum 13 questions).

Client 1 had no further comments during the Talking Mats interview, she did not want to part ways with Tessa after the experiment. The HW told her she will receive a new robot one week after the experiment was completed. She told us client 1 was glad to have the robot back. **Client 2** indicated twice that he views the Tessa robot the same as the HW's. He also said that he will miss the music when Tessa is gone. This was later discussed with his HW, which indicated that she had never heard him respond to Tessa and music has never been played by the robot. All negative results were posed by **client 3**. During the talking mats session, this client indicated that 1) he thinks Tessa is too noisy and 2) he thinks Tessa, in general, is nonsense. In some cases, the client experienced the robot as disturbing because he was busy with something else. The HW indicated that the client does not allow himself to be commanded by a speech robot and that this is probably the cause of the negative values. The client indicated that he does what he wants and does not have to listen to the Tessa robot. The client did indicate that he understands the purpose and good intentions of the Tessa.

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

541
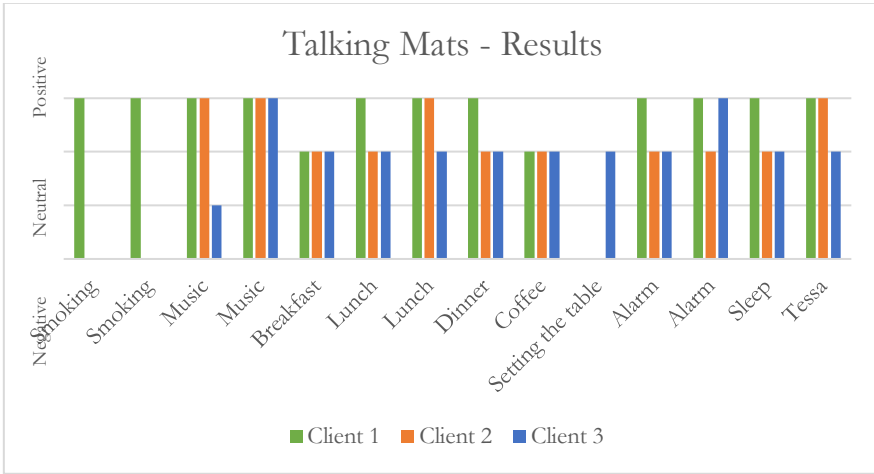
Figure 1: Talking mats results per subject

5.4   Semi-structured interviews (after the experiment)

After the experiment, interviews were held with seven HW's who worked with the Tessa robot, being three digital and four physical interviews. This led to additional requirements with regards to the context the Tessa is used in. Due to space constraints, not all requirements could be listed. Below we present the predominant requirements mentioned in the data.

Validation requirements:

- Involving HW's: Understanding the physical and mental health of clients is not only important. It is also important for creating support. One of the HW's had doubts beforehand about the usefulness of the robot. After seeing and experiencing the robot in use, she understood its purpose and was willing to use it.
- The HW's' trust in the Tessa robot: HW's continue to check on the clients whether they actually go to the living room, after the Tessa gives instructions to do this. In fact, HW's often already instructed clients from the room before the robot can give the instruction.
- The Tessa robot is not enough for every client to be stimulated, especially those who suffer from a more advanced stage of dementia. For example,

some clients need physical support in order to follow up on an instruction, which they do not always have access to.

6   Discussion

This study and its results could be influenced by several limitations. The first limitation concerns the influence of the HW's, which had often already reminded the client to get lunch or had already given a cigarette before Tessa provided an announcement. The HW's thereby limited the possibility for the Tessa robot to support the clients. As a result, it can be concluded that the Tessa robot must be implemented properly in the existing healthcare processes in order to determine its effectiveness. This meant that the effect of the announcements from Tessa could not always be measured properly. This could affect the overall validity of the measured effect Tessa might have had in this context. The results of the observations were different than expected but contributed to one of the most important findings of the study: the implementation of the Tessa robot in the existing healthcare processes in which HW's are an essential stakeholder is necessary to be able to use the Tessa effectively.

A second limitation concerns the clinical situation of the clients. Tessa's supplier indicated that it is important that Tessa's users have a good short-term memory and good hearing. The clients who participated in this experiment were at an advanced stage of dementia. As a result, one could argue that the clients with a similar clinical situation are not part of the target group for the Tessa robot. This was discussed with the HW's, that indicate that there are also clients who are in a less advanced stage of dementia. This limitation grounds further research using clients with a different clinical situation.

A third limitation concerns the size of the experiment. Three clients participated in this experiment. While one could argue that the qualitative explorartory approach allows for such sample size, the outcome of the current approach are difficult to generalize outside of this specific context.

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

543

7   Conclusion

The goal of this research was to identify how the acceptance and effectiveness of the Tessa robot could be improved. To achieve this, the following research question was answered, '*How can the acceptance and effectiveness of the Tessa robot for both HW's and people with dementia be improved?*'

Our study shows that both acceptance and effectiveness are influenced by the HW's and the clients. Several interviews revealed prejudices of HW's about the use of (social) robots. By making both HW's and clients more aware of the robot, not only the acceptance but possibly also its effectiveness of the robot could increase. This could potentially best be achieved by means of a learning program for the HW's. The HW's could attend a kick-off session and a simple, visual manual, that helps them to understand Tessa's purpose and functionalities. It is presumed that the effectiveness increases because the HW's can take the announcements into account and thus know when and how to rely on Tessa. Additionally, the findings show that, to increase the use of the Tessa, a hardware and software update is needed, e.g. improvement of the microphone capabilities as well as the addition of more interaction capabilities on top of the current response options. The functionalities currently are limited, with the consequence that for the target group described in this paper the robot could be effective as a companion, without achieving the goal of relieving announcement tasks of HW's. Furthermore, it can be concluded that the stage of dementia has a great impact on the experience with Tessa. For clients in an intramural care unit, the disease is often more advanced than for people in home care (often extramural). This has an effect on how people respond to Tessa's announcements. The robot failed to respond to a majority of answers from the participants. Most importantly, for the robot to be accepted and effective, it must be implemented properly in the existing healthcare processes, otherwise it might serve as a companion, but will not relieve the workload of healthcare workers.

8   Future Research

The findings provide multiple opportunities for future research. The first opportunity is to investigate how to increase the trust of HW's in social robots. Those results might allow social robots to be used more effectively in healthcare. The second opportunity is to investigate the Tessa in home care (extramural). Clients

who receive home care are often classified in a lesser advanced stage of dementia and are therefore more independent. Another opportunity emerged from the interviews and observations, where HW's believe that the robot needs a humanoid shape that can stimulate the clients better (e.g., arms and legs). Future research could therefore focus on whether the Tessa robot can be modified or whether another robot is more suitable for activating people with dementia. The last opportunity for future research is the speech recognition of the Tessa robot. As indicated earlier, Tessa is now limited in its communication by its constrained interpretation abilities. The Tessa can only understand "yes" or "no" when a music moment has been scheduled. It would be interesting to see whether a robot with speech recognition stimulates clients more, as they get a response to the things they say to Tessa. Moreover, during the observations, it happened several times that the clients already answered before the microphone of the robot was turned on. Such delayed interactions significantly hampered the effectiveness of Tessa. Future research should focus on improving the speech recognition so that the microphone is switched on, taking into account the privacy and safety aspects of the client.

Acknowledgements

References

Alnanih, R., & Ormandjieva, O. (2016). Mapping HCI Principles to Design Quality of Mobile User Interfaces in Healthcare Applications. Procedia Computer Science, 94, 75–82. https://doi.org/10.1016/j.procs.2016.08.014

Andrés, A., Pardo, D. E., Díaz, M., & Angulo, C. (2015). New instrumentation for human robot interaction assessment based on observational methods. Journal of Ambient Intelligence and Smart Environments, 7(4), 397–413. https://doi.org/10.3233/AIS-150331

Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The Place of Inter-Rater Reliability in Qualitative Research: An Empirical Study. Sociology, 31(3), 597–606.

Bouwhuis, D. G. (2016). Current use and possibilities of robots in care. Gerontechnology, 15(4), 198–208.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding In-depth Semistructured Interviews: Problems of Unitization and Intercoder Reliability and Agreement. Sociological Methods & Research, 42(3), 294–320.

Casaccia, S., Revel, G. M., Scalise, L., Bevilacqua, R., Rossi, L., Paauwe, R. A., … Nap, H. H. (2019). Social Robot and Sensor Network in Support of Activity of Daily Living for People with Dementia. In R. Brankaert & W. IJsselsteijn (Eds.), Dementia Lab 2019. Making Design Work: Engaging with Dementia in Context (Vol. 1117, pp. 128–135). Cham: Springer.

K. Smit, M. Smakman, S. Bakker, J. Blokhuis, G. Evertzen &L. Polman:
Hello, is Someone There? A Case Study for Using a Social Robot in Dementia Care

545

Castillo-Montoya, M. (2016). Preparing for Interview Research: The Interview Protocol Refinement Framework. The Qualitative Report, 21(5), 811–831.

Cheng, S.-T. (2017). Dementia Caregiver Burden: a Research Update and Critical Analysis. Current Psychiatry Reports, 19(9), 64. https://doi.org/10.1007/s11920-017-0818-2

de Graaf, M. M. A., & Ben Allouch, S. (2013). Exploring influencing variables for the acceptance of social robots. Robotics and Autonomous Systems, 61(12), 1476–1486.

de Graaf, M., Somaya, B. A., & van Dijk, J. (2016). Long-Term Acceptance of Social Robots in Domestic Environments: Insights from a User's Perspective. In AAAI Spring Symposia.

Domínguez-Rué, E., & Nierling, L. (Eds.). (2016). Ageing and technology: perspectives from the social sciences. Bielefeld: Transcript.

Downes, M. J., Dean, R., & Bath‐Hextall, F. J. (2013). Animal‐assisted therapy for people with serious mental illness. Cochrane Database of Systematic Reviews, (12).

Duffy, B. R. (2003). Anthropomorphism and the social robot. Robotics and Autonomous Systems, 42(3), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

Edmondson, A. C., & Mcmanus, S. E. (2007). Methodological Fit in Management Field Research. Proceedings of the Academy of Management, 32(4), 1155–1179.

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value Sensitive Design and Information Systems. In The handbook of information and computer ethics (pp. 55–95).

Góngora Alonso, S., Hamrioui, S., de la Torre Díez, I., Motta Cruz, E., López-Coronado, M., & Franco, M. (2019a). Social Robots for People with Aging and Dementia: A Systematic Review of Literature. Telemedicine and E-Health, 25(7), 533–540.

Góngora Alonso, S., Hamrioui, S., de la Torre Díez, I., Motta Cruz, E., López-Coronado, M., & Franco, M. (2019b). Social Robots for People with Aging and Dementia: A Systematic Review of Literature. Telemedicine and E-Health, 25(7), 533–540. https://doi.org/10.1089/tmj.2018.0051

Johnson, M. J., Johnson, M. A., Sefcik, J. S., Cacchione, P. Z., Mucchiani, C., Lau, T., & Yim, M. (2020). Task and Design Requirements for an Affordable Mobile Service Robot for Elder Care in an All-Inclusive Care for Elders Assisted-Living Setting. International Journal of Social Robotics, 12(5), 989–1008. https://doi.org/10.1007/s12369-017-0436-5

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1), 159. https://doi.org/10.2307/2529310

Miguel, S., Alvira, M., Farré, M., Risco, E., Cabrera, E., & Zabalegui, A. (2016). Quality of life and associated factors in older people with dementia living in long-term institutional care and home care. European Geriatric Medicine, 7(4), 346–351. https://doi.org/10.1016/j.eurger.2016.01.012

Moharana, S., Panduro, A. E., Lee, H. R., & Riek, L. D. (2019). Robots for Joy, Robots for Sorrow: Community Based Robot Design for Dementia Caregivers. In 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 458–467). Daegu, Korea (South): IEEE. https://doi.org/10.1109/HRI.2019.8673206

Mordoch, E., Osterreicher, A., Guse, L., Roger, K., & Thompson, G. (2013). Use of social commitment robots in the care of elderly people with dementia: A literature review. Maturitas, 74(1), 14–20.

Murphy, J., Gray, C. M., Cox, S., & Joseph Rowntree Foundation. (2007). Communication and dementia: how Talking Mats can help people with dementia to express themselves. York: Joseph Rowntree Foundation.

Naneva, S., Sarda Gou, M., Webb, T. L., & Prescott, T. J. (2020). A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots. International Journal of Social Robotics, 12(6), 1179–1201. https://doi.org/10.1007/s12369-020-00659-4

Pervan, G., & Maimbo, M. (2005). Designing a case study protocol for application in IS research. In Proceedings of the Ninth Pacific Asia Conference on Information Systems (pp. 1281–1292). PACIS.

Qwiek. (2021). Belevingsgerichte zorg met zorgondersteuning. Retrieved March 23, 2021, from https://www.qwiek.eu/up

Robillard, J. M., Cleland, I., Hoey, J., & Nugent, C. (2018). Ethical adoption: A new imperative in the development of technology for dementia. Alzheimer's & Dementia, 14(9), 1104–1113.

Robotzorg. (2021). Tessa, sociale robot. Retrieved March 23, 2021, from https://www.robotzorg.nl/product/tessa-sociale-robot/

Roser, M., Ortiz-Ospina, E., & Ritchie, H. (2019). Life Expectancy. Retrieved from https://ourworldindata.org/life-expectancy#

Runeson, P., & Höst, M. (2008). Guidelines for conducting and reporting case study research in software engineering. Empirical Software Engineering, 14(2), 131.

Savela, N., Turja, T., & Oksanen, A. (2018). Social Acceptance of Robots in Different Occupational Fields: A Systematic Literature Review. International Journal of Social Robotics, 10(4), 493–502.

Toulmin, S. E. (2003). The Uses of Argument. The Press Syndicate of the University of Cambridge.

Turja, T., Van Aerschot, L., Särkikoski, T., & Oksanen, A. (2018). Finnish healthcare professionals' attitudes towards robots: Reflections on a population sample. Nursing Open, 5(3), 300–309.

Tyack, C., & Camic, P. M. (2017). Touchscreen interventions and the well-being of people with dementia and caregivers: a systematic review. International Psychogeriatrics, 29(8), 1261–1280.

Valentí Soler, M., Agüera-Ortiz, L., Olazarán Rodríguez, J., Mendoza Rebolledo, C., Pérez Muñoz, A., Rodríguez Pérez, I., ... Martínez Martín, P. (2015a). Social robots in advanced dementia. Frontiers in Aging Neuroscience, 7. https://doi.org/10.3389/fnagi.2015.00133

Valentí Soler, M., Agüera-Ortiz, L., Olazarán Rodríguez, J., Mendoza Rebolledo, C., Pérez Muñoz, A., Rodríguez Pérez, I., ... Martínez Martín, P. (2015b). Social robots in advanced dementia. Frontiers in Aging Neuroscience, 7(133). https://doi.org/10.3389/fnagi.2015.00133

van Beek, A. P. A., Frijters, D. H. M., Wagner, C., Groenewegen, P. P., & Ribbe, M. W. (2011). Social engagement and depressive symptoms of elderly residents with dementia: a cross-sectional study of 37 long-term care units. International Psychogeriatrics, 23(4), 625–633.

van der Aalst, M. (2019). Factsheet Zorg. UWV Afdeling Arbeidsmarktinformatie en –advies. Retrieved from https://www.uwv.nl/overuwv/kennis-cijfers-en onderzoek/arbeidsmarktinformatie/factsheet-zorg-2019.aspx

van Kemenade, M., Hoorn, J., & Konijn, E. (2018). Healthcare Students' Ethical Considerations of Care Robots in The Netherlands. Applied Sciences, 8(10), 1712.

Visser, G., & Vandemeulebroucke, T. (2018). Sociale robots voor thuiswonende mensen met dementie: plezier of bittere noodzaak? Geron, 20(4), 60–63. https://doi.org/10.1007/s40718-018-0179-x

Whelan, S., Murphy, K., Barrett, E., Krusche, C., Santorelli, A., & Casey, D. (2018). Factors Affecting the Acceptability of Social Robots by Older Adults Including People with Dementia or Cognitive Impairment: A Literature Review. International Journal of Social Robotics, 10(5), 643–668.

A Follow-Up on the Changes in the Use Intention of Digital Wellness Technologies and Its Antecedents Over Time: The Use of Physical Activity Logger Applications Among Young Elderly in Finland

Markus Makkonen,[1,2] Tuomas Kari[1,2] & Lauri Frank[3]

[1] Institute for Advanced Management Systems Research, Turku, Finland; e-mail: markus.v.makkonen@jyu.fi, tuomas.t.kari@jyu.fi
[2] University of Jyvaskyla, Jyvaskyla, Finland; e-mail: markus.v.makkonen@jyu.fi, tuomas.t.kari@jyu.fi
[3] University of Jyvaskyla, Faculty of Information Technology, Jyvaskyla, Finland; e-mail: lauri.frank@jyu.fi

Abstract This study aims to further promote the understanding of the antecedents of the acceptance and use of digital wellness technologies among elderly people through a follow-up to our two prior studies, one which examines the potential longer-term temporal changes in the use intention of digital wellness technologies and its antecedents in the case of the young elderly segment and physical activity logger applications. We base this examination theoretically on UTAUT2 and empirically on survey data that is collected from 92 Finnish young elderly users of a physical activity logger application in three subsequent time points and analysed with partial least squares structural equation modelling (PLS-SEM). We find that the initial strong decline in the scores of the antecedent constructs and use intention becomes weaker as the construct scores stabilise over time, whereas especially the effects of performance expectancy and effort expectancy on use intention remain relatively unstable.

1   Introduction

Physical inactivity has become an increasingly prevalent problem among elderly people (Sun, Norman & While, 2013), thus raising a call for new and innovative ways to promote their levels of physical activity. One potential way to do this are different types of digital wellness technologies, such as smartphone and smartwatch applications, which have been found very promising in terms of promoting the levels of physical activity not only among young but also among elderly people (e.g., McGarrigle & Todd, 2020). In addition to elderly people in general, their potential has been highlighted especially in the more specific segment of young elderly, which consists of people aged approximately 60–75 years (e.g., Carlsson & Walden, 2016). However, there is a lack of prior studies that have examined the antecedents of the acceptance and use of digital wellness technologies among elderly people, particularly from a longitudinal perspective of how their use evolves after the initial acceptance. These kinds of longitudinal studies can be considered highly important in the context of digital wellness technologies because, as it is suggested in theories like the lived informatics model of personal informatics (Epstein, Ping, Fogarty & Munson, 2015), the use of these technologies, especially those aimed at self-tracking, is often characterised by "lapses" in their use. This suggests that the intention to use the technologies and its antecedents do not remain constant but change over time. However, in prior information systems (IS) literature, such temporal changes have not been studied from the perspective of technology acceptance and use.

The objective of this study is to address this gap in prior research by studying *how the use intention of digital wellness technologies and its antecedents among elderly people potentially change over time*. We examine this research question in the case of the young elderly segment and one common type of digital wellness technology: physical activity logger applications. By *physical activity logger applications*, we refer to mobile applications that enable users to keep track of their physical activities in everyday life as well as view different types of reports about them. As the theoretical foundation for conceptualising the antecedents of the intention to use physical activity logger applications and formulating the research model for examining the potential temporal changes in use intention and its antecedents, we use UTAUT2 by Venkatesh, Thong, and Xu (2012), which is one of the most comprehensive and established IS theories for explaining technology acceptance and use in consumer contexts, such as the one of this study. In turn, as the empirical data for the

examination, we use survey data that is collected from 92 Finnish young elderly users of a physical activity logger application in three subsequent time points and analysed with partial least squares structural equation modelling (PLS-SEM). The study was conducted as part of our broader DigitalWells research program, which focuses on young elderly in Finland and in which the participants are provided for free both a physical activity logger application to keep track of their daily physical activities as well as the training and support for setting up and using it. The study is a follow-up to our two prior studies (Makkonen, Kari & Frank, 2020, 2021), in which we initially proposed and tested our research model for explaining the acceptance and use of digital wellness technologies in the case of young elderly and physical activity logger applications as well as examined the potential changes in use intention and its antecedents between about four months and about 12 months of use. Here, this time span is extended to about 18 months, thus enabling the examination of even longer-term changes.

After this introductory section, we describe in more detail the research model and the research methodology of the study in Sections 2 and 3. This is followed by reporting of the research results in Section 4. The results are discussed in more detail in Section 5 before concluding the paper with a brief discussion about the limitations of the study and some potential paths of future research in Section 6.

2   Research Model

As already mentioned above, the research model of the study is based on UTAUT2 by Venkatesh et al. (2012), which is an extension of the unified theory of acceptance and use of technology (UTAUT) by Venkatesh, Morris, Davis, and Davis (2003) from organisational to consumer contexts. UTAUT2 has been applied to explain technology acceptance and use in numerous IS contexts, including also the context of mobile health and fitness applications and devices (e.g., Yuan, Ma, Kanthawala & Peng, 2015; Duarte & Pinho, 2019; Talukder, Chiong, Bao & Malik, 2019; Dhiman, Arora, Dogra & Gupta, 2020; Beh, Ganesan, Iranmanesh & Foroughi, 2021) and the context of elderly users (e.g., Macedo, 2017). However, none of these prior studies have combined the two contexts by examining, for example, the acceptance and use of physical activity logger applications among young elderly, as it is done in this study. In UTAUT2, the *behavioural intention* (BI) to use a particular technology is

hypothesised to be positively affected by seven antecedents (Venkatesh et al., 2012): *performance expectancy* (PE – i.e., the degree to which using a technology will provide benefits to consumers in performing certain activities), *effort expectancy* (EE – i.e., the degree of ease associated with consumers' use of technology), *social influence* (SI – i.e., the extent to which consumers perceive that important others believe they should use a particular technology), *facilitating conditions* (FC – i.e., consumers' perceptions of the resources and support available to perform a behaviour), *hedonic motivation* (HM – i.e., the fun or pleasure derived from using a technology), *price value* (PV – i.e., the consumers' cognitive trade-off between the perceived benefits of the technology and the monetary cost for using it), and *habit* (HT – i.e., the extent to which people tend to perform behaviours automatically because of learning). In addition, UTAUT2 hypothesises three moderators for the effects of these seven antecedents on use intention: age, gender, and experience. However, due to the limited sample size of this study, these moderators are omitted from the research model. In addition, we omit two of the seven antecedents: facilitating conditions and price value. These were considered irrelevant in the present study because the application was free for all the participants and they all had the same resource requirements for taking part in the research program (e.g., owning a smartphone on which the application can be installed) as well as were given the same training and support for setting up and using the application, thus assumably resulting in very low variance in their perceptions of these issues. Finally, as in many studies on technology acceptance and use, the research model also concentrates on explaining only use intention and not actual *use behaviour* (UB). The final research model of the study, with the omitted constructs and effects presented as dashed, is illustrated in Figure 1.

M. Makkonen, T. Kari & L. Frank:
A Follow-Up on the Changes in the Use Intention of Digital Wellness Technologies and Its Antecedents   551
Over Time: The Use of Physical Activity Logger Applications Among Young Elderly in Finland
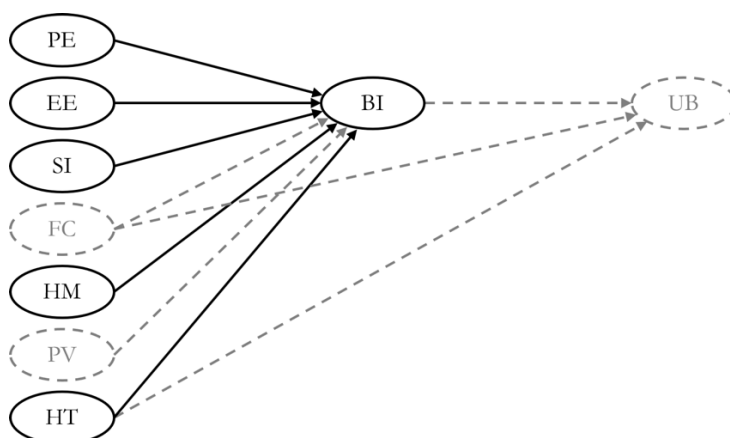
Figure 1: Research model (the dashed constructs and effects are omitted in this study)

3   Methodology

The data for the study was collected from the participants of our research program in three subsequent surveys. These were conducted in autumn 2019 after about four months of using the application, in summer 2020 after about 12 months of using the application, and in winter 2021 after about 18 months of using the application. In the remainder of this paper, these three time points, respectively, are referred to as T1, T2, and T3. The first survey was administered as a pen-and-paper survey in face-to-face group meetings with the participants, whereas the second and third survey were both administered as online surveys due to the ongoing COVID-19 pandemic. Because Finland has two official languages, the participants had the option to respond to the surveys in either Finnish or Swedish. In the surveys, each construct of the research model was measured reflectively by three indicators. All the indicators were adapted from the study by Venkatesh et al. (2012) and their wordings in English are reported in Table 1. The measurement scale was a seven-point Likert scale ranging from one (strongly disagree) to seven (strongly agree). Because we wanted to avoid forced responses, the participants also had the option not to respond to a particular item, which resulted in a missing value.

Table 1: Indicator wordings

Indicator	Wording
PE1	I find the app useful in achieving my daily exercise goals.
PE2	Using the app helps me achieve my exercise goals more quickly.
PE3	Using the app increases my efficiency in achieving my exercise goals.
EE1	Learning how to use the app to achieve my exercise goals is easy for me.
EE2	I find using the app to achieve my exercise goals easy.
EE3	It is easy for me to become skilful at using the app to achieve my exercise goals.
SI1	People who are important to me think that I should use the app to achieve my exercise goals.
SI2	People who influence my behaviour think that I should use the app to achieve my exercise goals.
SI3	People whose opinions I value prefer that I use the app to achieve my exercise goals.
HM1	Using the app to achieve my exercise goals is fun.
HM2	Using the app to achieve my exercise goals is enjoyable.
HM3	Using the app to achieve my exercise goals is entertaining.
HT1	The use of the app to achieve my exercise goals has become a habit for me.
HT2	I am addicted to using the app to achieve my exercise goals.
HT3	I must use the app to achieve my exercise goals.
BI1	I intend to continue using the app to achieve my exercise goals.
BI2	I will always try to use the app to achieve my exercise goals.
BI3	I plan to use the app regularly to achieve my exercise goals.

Due to the limited sample size of this study, the collected data was analysed with variance-based structural equation modelling (VB-SEM), more specifically partial least squares structural equation modelling (PLS-SEM). As a statistical software for PLS-SEM, we used SmartPLS 3.3.3 by Ringle, Wende, and Becker (2015). We also followed carefully the previously published guidelines for conducting PLS-SEM in IS research given by Hair, Hollingsworth, Randolph, and Chong (2017). For example, in accordance with the given guidelines, we used mode A as the indicator weighting mode of the constructs, path weighting as the weighting scheme, +1 as the initial weights, and $< 10^{-7}$ as the stop criterion in model estimation, whereas the statistical significance of the model estimates was tested by using bootstrapping with 5,000 subsamples. As the threshold for statistical significance, we used $p < 0.05$. The potential missing values were handled by using mean replacement.

The estimated model consisted of three submodels, which were otherwise identical and formulated based on the research model illustrated in Figure 1, but which were estimated by using the data collected at T1, T2, and T3, respectively. The three submodels were also connected by so-called carry-over effects (cf. Roemer, 2016), which were used to examine how the scores of a specific construct at a previous time point (i.e., T1 or T2) affect the scores of that same construct at a subsequent time point (i.e., T2 or T3). After estimating the model and evaluating the reliability and validity of its three submodels at both construct and indicator levels, the potential changes in the estimated construct scores and effect sizes from T1 to T2 and from T2 to T3 were examined. This examination followed the procedure proposed by Roemer (2016) for evolution models with panel data (also referred to as model type A.1 in her paper). First, the statistical significance of the changes in the means of the estimated unstandardised construct scores from T1 to T2 and from T2 to T3 were tested by using the parametric Student's paired samples t-test. Its results were additionally confirmed by using the nonparametric Wilcoxon (1945) signed-rank test if the compared means were not found to be normally distributed as suggested by the Shapiro-Wilk (1965) test. Second, the estimated size of each effect at a previous time point (i.e., T1 or T2) was compared against the 95% bias-corrected and accelerated (BCa) confidence interval (cf. Hair, Hult, Ringle & Sarstedt, 2017) of the estimated size of that same effect at a subsequent time point (i.e., T2 or T3). If the estimate at the previous time point did not fall within the confidence interval of the estimate at the subsequent time point, then the change in the effect size could be considered statistically significant.

4   Results

In total, 115 participants provided valid responses to the survey at T1, of which 99 participants did so also at T2 (a drop-out rate of 13.9%), and of which 92 participants did so also at T3 (a drop-out rate of 7.1%). The descriptive statistics of these three samples in terms of the gender, age, and response language of the participants as well as a subjective assessment of their level of physical activity are reported in Table 2. As can be seen, the drop-outs did not result in any considerable changes in the sample profiles. As the sample for this particular study, we used the last-mentioned sample of 92 participants who had provided valid responses to the survey at all the three time points. Of them, about two-thirds were women and over nine out of ten

assessed their level of physical activity as moderate or higher. Their age ranged from 49 to 79 years, with a mean of 69.1 years and a standard deviation of 4.7 years. Although some of the participants were slightly younger or older than our target young elderly segment consisting of people aged approximately 60–75 years, we decided not to omit these people from the study due to our limited sample size.

Table 2: Sample statistics

	T1 (N = 115)		T2 (N = 99)		T3 (N = 92)	
	N	%	N	%	N	%
Gender						
Man	43	37.4	34	34.3	32	34.8
Woman	72	62.6	65	65.7	60	65.2
Age						
Under 60 years	3	2.6	2	2.0	2	2.2
60–64 years	11	9.6	10	10.1	8	8.7
65–69 years	44	38.3	39	39.4	38	41.3
70–74 years	39	33.9	35	35.4	32	34.8
75 years or over	18	15.7	13	13.1	12	13.0
Language						
Finnish	69	60.0	63	63.6	60	65.2
Swedish	46	40.0	36	36.4	32	34.8
Level of physical activity						
Very high	1	0.9	1	1.0	1	1.1
High	18	15.7	16	16.2	15	16.3
Moderate	84	73.0	73	73.7	68	73.9
Low	4	3.5	3	3.0	3	3.3
Very low	8	7.0	6	6.1	5	5.4
Totally passive	0	0.0	0	0.0	0	0.0

4.1 Model Estimation

The estimation results of the three submodels in terms of the standardised size and statistical significance of the effects of the antecedent constructs on use intention at T1, T2, and T3 are reported in Table 3. As the reported values show, at all the three time points, the effects of hedonic motivation and habit were found to be positive and statistically significant, whereas the effect of social influence was found to be statistically not significant. In contrast, performance expectancy was found to have a positive and statistically significant effect at T1 and T3, but a statistically not significant effect at T2, whereas effort expectancy was found to have a positive and statistically significant effect at T2, but a statistically not significant effect at T1 and

M. Makkonen, T. Kari & L. Frank:
A Follow-Up on the Changes in the Use Intention of Digital Wellness Technologies and Its Antecedents
Over Time: The Use of Physical Activity Logger Applications Among Young Elderly in Finland   555

T3. In terms of explanatory power, the proportion of explained variance (R^2) in use intention was 72.7% at T1, 77.3% at T2 and 83.2% at T3.

Table 3: Effects on use intention (* = p < 0.001, ** = p < 0.01, * = p < 0.05)**

	T1		T2		T3	
	Size	**95% CI**	**Size**	**95% CI**	**Size**	**95% CI**
PE → BI	0.348**	[0.140, 0.567]	0.091	[-0.129, 0.324]	0.337**	[0.143, 0.565]
EE → BI	0.064	[-0.092, 0.225]	0.324***	[0.177, 0.493]	0.077	[-0.083, 0.278]
SI → BI	0.025	[-0.109, 0.146]	0.091	[-0.041, 0.207]	-0.047	[-0.202, 0.078]
HM → BI	0.247**	[0.069, 0.420]	0.243**	[0.085, 0.399]	0.267*	[0.019, 0.519]
HT → BI	0.308**	[0.102, 0.509]	0.272*	[0.069, 0.497]	0.227**	[0.060, 0.382]

Table 4: Carry-over effects (* = p < 0.001, ** = p < 0.01, * = p < 0.05)**

	T1 → T2		T2 → T3	
	Size	**R^2 by T1 at T2**	**Size**	**R^2 by T2 at T3**
PE → PE	0.453***	0.205	0.707***	0.500
EE → EE	0.295*	0.087	0.624***	0.389
SI → SI	0.534***	0.285	0.600***	0.360
HM → HM	0.466***	0.217	0.801***	0.641
HT → HT	0.466***	0.217	0.763***	0.582
BI → BI	0.107*	0.045	0.146	0.110

In turn, Table 4 reports the standardised size and statistical significance of the carry-over effects between the constructs of the three submodels as well as the proportion of explained variance (R^2) in the scores of a specific construct at a subsequent time point by the scores of that same construct at a previous time point. As the reported values show, all the carry-over effects except for the one concerning use intention between T2 and T3 were found to be statistically significant. All the carry-over effects also seemed to be considerably stronger between T2 and T3 in comparison to T1 and T2, meaning that the construct scores provided by the participants became more stable over time. The only slight exceptions to this were the carry-over effects concerning social influence and use intention, which remained approximately equally strong between T1 and T2 as well as T2 and T3.

4.2 Construct Reliability and Validity

Table 5: Construct statistics

		CR	AVE	PE	EE	SI	HM	HT	BI
PE	**T1**	0.908	0.766	0.875					
	T2	0.898	0.746	0.864					
	T3	0.904	0.759	0.871					
EE	**T1**	0.876	0.703	0.464	0.838				
	T2	0.879	0.707	0.586	0.841				
	T3	0.940	0.840	0.572	0.917				
SI	**T1**	0.930	0.816	0.485	0.297	0.903			
	T2	0.861	0.675	0.513	0.452	0.822			
	T3	0.902	0.754	0.635	0.369	0.868			
HM	**T1**	0.932	0.820	0.677	0.472	0.533	0.906		
	T2	0.903	0.756	0.742	0.520	0.391	0.869		
	T3	0.949	0.860	0.830	0.674	0.595	0.927		
HT	**T1**	0.815	0.597	0.698	0.474	0.478	0.667	0.773	
	T2	0.875	0.700	0.754	0.529	0.484	0.642	0.837	
	T3	0.870	0.691	0.793	0.564	0.510	0.729	0.831	
BI	**T1**	0.891	0.733	0.772	0.495	0.491	0.732	0.758	0.856
	T2	0.913	0.777	0.759	0.706	0.538	0.727	0.755	0.881
	T3	0.895	0.741	0.850	0.653	0.522	0.843	0.810	0.861

Construct reliabilities were evaluated by examining the composite reliability (CR) of each construct (Fornell & Larcker, 1981), which is commonly expected to be greater than or equal to 0.7 (Nunnally & Bernstein, 1994). The CR of each construct at T1, T2, and T3 is reported in the first column of Table 5, showing that all the constructs at all the three time points clearly met this criterion. In turn, construct validities were evaluated by examining the convergent and discriminant validities of the constructs by using two criteria based on the average variance extracted (AVE) of each construct (Fornell & Larcker, 1981). In order to exhibit satisfactory convergent validity, the first criterion requires that each construct should have an AVE of at least 0.5. The AVE of each construct at T1, T2, and T3 is reported in the second column of Table 5, showing that all the constructs at all the three time points met also this criterion. In turn, in order to exhibit satisfactory discriminant validity, the second criterion requires that each construct should have a square root of AVE greater than or equal to its absolute correlation with the other model constructs. The square root of AVE of each construct at T1, T2, and T3 (on-diagonal cells) and the correlations between the constructs at T1, T2, and T3 (off-diagonal cells) are

M. Makkonen, T. Kari & L. Frank:
A Follow-Up on the Changes in the Use Intention of Digital Wellness Technologies and Its Antecedents Over Time: The Use of Physical Activity Logger Applications Among Young Elderly in Finland

557

reported in the remaining columns of Table 5, showing that also this criterion was met by all the constructs at all the three time points.

4.3   Indicator Reliability and Validity

Table 6: Indicator statistics (*** = all loadings have p < 0.001)

	T1				T2				T3			
	Mean	SD	NA	λ***	Mean	SD	NA	λ***	Mean	SD	NA	λ***
PE1	5.682	1.474	4.3%	0.863	5.473	1.515	1.1%	0.855	5.045	1.685	3.3%	0.896
PE2	5.284	1.494	12.0%	0.892	4.722	1.696	2.2%	0.833	4.943	1.564	4.3%	0.824
PE3	5.216	1.572	4.3%	0.870	4.945	1.656	1.1%	0.902	4.878	1.648	2.2%	0.892
EE1	6.330	1.155	1.1%	0.832	5.811	1.564	2.2%	0.851	5.978	1.382	1.1%	0.934
EE2	6.143	1.179	1.1%	0.882	5.934	1.315	1.1%	0.809	5.826	1.573	0.0%	0.938
EE3	5.747	1.495	1.1%	0.799	5.167	1.691	2.2%	0.862	5.477	1.576	4.3%	0.876
SI1	4.264	2.130	21.7%	0.916	4.090	1.949	15.2%	0.860	3.975	2.124	14.1%	0.903
SI2	4.603	2.110	26.1%	0.925	3.939	1.990	10.9%	0.770	3.695	2.141	10.9%	0.880
SI3	5.219	1.797	20.7%	0.868	4.538	1.855	13.0%	0.832	4.679	2.042	15.2%	0.820
HM1	5.739	1.255	4.3%	0.932	5.352	1.456	1.1%	0.859	5.011	1.742	1.1%	0.945
HM2	5.906	1.076	7.6%	0.894	5.270	1.643	3.3%	0.894	5.185	1.630	0.0%	0.903
HM3	5.141	1.536	7.6%	0.891	4.571	1.634	1.1%	0.854	4.611	1.733	2.2%	0.934
HT1	6.135	1.333	3.3%	0.758	5.833	1.448	2.2%	0.841	5.231	1.820	1.1%	0.840
HT2	4.241	1.935	9.8%	0.712	3.956	1.914	1.1%	0.839	3.899	1.995	3.3%	0.819
HT3	5.136	1.717	4.3%	0.842	4.615	1.855	1.1%	0.831	4.333	1.818	2.2%	0.834
BI1	5.869	1.495	8.7%	0.858	5.831	1.487	3.3%	0.904	5.639	1.551	9.8%	0.841
BI2	5.595	1.262	8.7%	0.808	4.822	1.680	2.2%	0.873	4.890	1.722	1.1%	0.864
BI3	5.841	1.437	4.3%	0.899	5.523	1.470	4.3%	0.867	5.379	1.713	5.4%	0.877

Indicator reliabilities and validities were evaluated by using the standardised loading (λ) of each indicator, which are reported at T1, T2, and T3 in Table 6 together with the mean and standard deviation (SD) of the indicator scores as well as the percentage of missing values (NA). In the typical case where each indicator loads on only one construct, its standardised loading is commonly expected to be statistically significant and greater than or equal to 0.707 (Fornell & Larcker, 1981). As the reported values show, all the indicators at all the three time points met this criterion.

4.4 Changes in Construct Scores

Table 7: Construct scores and the changes in construct scores

	T1		T2		T3		ΔT1→T2		ΔT2→T3	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
PE	5.396	1.279	5.077	1.387	4.957	1.405	-0.320	1.398	-0.120	1.069
EE	6.082	1.056	5.670	1.260	5.770	1.367	-0.412	1.385	0.101	1.143
SI	4.726	1.584	4.219	1.475	4.120	1.698	-0.507	1.479	-0.099	1.434
HM	5.634	1.111	5.077	1.356	4.931	1.570	-0.557	1.292	-0.146	0.946
HT	5.265	1.229	4.925	1.417	4.531	1.538	-0.340	1.378	-0.394	1.025
BI	5.754	1.141	5.411	1.336	5.317	1.389	-0.343	1.346	-0.094	0.961

Table 8: Testing of the changes in construct scores

	ΔT1→T2					ΔT2→T3				
	Student's t-test			Wilcoxon test		Student's t-test			Wilcoxon test	
	t	df	p	z	p	t	df	p	z	p
PE	-2.192	91	0.031	-1.429	0.153	-1.078	91	0.284	-1.150	0.250
EE	-2.854	91	0.005	-2.746	0.006	0.844	91	0.401	-1.215	0.224
SI	-3.287	91	0.001	-3.754	< 0.001	-0.661	91	0.510	-0.289	0.773
HM	-4.132	91	< 0.001	-4.188	< 0.001	-1.479	91	0.143	-1.335	0.182
HT	-2.366	91	0.020	-2.481	0.013	-3.684	91	< 0.001	-3.559	< 0.001
BI	-2.446	91	0.016	-2.924	0.003	-0.940	91	0.350	-0.537	0.591

In terms of the changes in construct scores, Table 7 reports the means and standard deviations (SD) of the estimated unstandardised construct scores at T1, T2, and T3 as well as the means and standard deviations (SD) of the changes in them from T1 to T2 and from T2 to T3. As can be seen, the participants had relatively high scores in the case of all the constructs at all the three time points, but the scores seemed to decline over time, more drastically from T1 to T2 and less drastically from T2 to T3. The statistical significance of these changes was tested by using both parametric and nonparametric testing because most of the compared construct mean scores were not found to be normally distributed. The results of these tests are reported in Table 8, showing that from T1 to T2, the changes in the construct mean scores were found to be statistically significant in the case of all the antecedent constructs except potentially for performance expectancy, in the case of which the statistical significance of the change suggested by parametric testing could not be confirmed

by nonparametric testing. In contrast, from T2 to T3, the changes in the construct mean scores were found to be statistically significant only in the case of habit.

4.5 Changes in Effect Sizes

In terms of the changes in effect sizes, Table 3 additionally reports the 95% confidence interval (CI) of the estimated size of each effect at T1, T2, and T3. As can be seen, the estimated size of the effects of performance expectancy and effort expectancy at T1 and T2, respectively, did not fall within the 95% CI of the estimated size of the same effects at T2 and T3, respectively, thus suggesting that the changes in the size of these effects from T1 to T2 and from T2 to T3 were statistically significant. More specifically, the effect of performance expectancy seemed to become weaker from T1 to T2 and stronger from T2 to T3, whereas the effect of effort expectancy seemed to become stronger from T1 to T2 and weaker from T2 to T3. In addition, the estimated size of the effect of social influence at T2 did not fall within the 95% CI of the estimated size of the same effect at T3, but this effect remained statistically not significant at both these two time points.

5 Discussion and Conclusions

In this study, we examined the potential longer-term temporal changes in the use intention of digital wellness technologies and its antecedents in the case of the young elderly segment and physical activity logger applications. In comparison to our two prior studies (Makkonen et al., 2020, 2021), we made three main findings. First, we found that our research model continued to perform very well in explaining use intention also after about 18 months of using the application by being able to explain about 83% of its variance at T3 as well as having acceptable reliability and validity at both construct and indicator levels. Thus, it seems to be well suited also for longitudinal study settings in which the time span extends well beyond one year. Second, we found that the strong decline in the scores of the antecedent constructs and use intention from T1 to T2 became weaker in the case of most of the constructs from T2 to T3, which was also supported by the finding concerning the stabilisation of the construct scores over time. The only exception to this was habit, the scores of which continued to decline about as strongly between T2 and T3 as between T1 and T2, although its scores also became more stable over time. This initially strong but then increasingly weaker decline in the construct scores is most likely explained

by a novelty effect that causes a particular digital wellness technology to be first perceived very favourably by its potential users but these perceptions to converge towards realism as the use progresses, first more quickly, as also the hype concerning the technology is higher, but then more slowly. Third, we found that the effects of social influence, hedonic motivation, and habit on use intention continued to remain very stable between T2 and T3, as they did also between T1 and T2, whereas more instability could be observed in the effects of performance expectancy and effort expectancy. That is, whereas the effect of performance expectancy become weaker and the effect of effort expectancy stronger from T1 to T2, these changes were now reversed, with the effect of performance expectancy once again becoming stronger and the effect of effort expectancy weaker. What is actually causing this instability, as well as whether it is driven more by internal changes in the users themselves or external changes in their environment, requires more in-depth examinations. However, all in all, the aforementioned temporal changes in both the effects of the antecedent constructs on use intention and the scores of the antecedent constructs themselves would seem to provide some much-needed theoretical explanations for the "lapses" in the use of personal informatics or self-tracking technologies, such as physical activity logger applications, which have been suggested in theories like the lived informatics model of personal informatics (Epstein et al., 2015). In turn, from a more practical perspective, the findings of the study highlight the need for the providers of various digital wellness technology products and services to actively adapt their offerings to the aforementioned temporal changes as well as to continuously promote the novelty of their offerings through approaches like gamification (e.g., Kari, Piippo, Frank, Makkonen & Moilanen, 2016) and exergaming (e.g., Kari, 2014; Kari & Makkonen, 2014) in order to prevent the perceptions of the users from becoming less favourable as the initial novelty effect fades out.

6   Limitations and Future Research

Like our two prior studies, this study can be considered to have three main limitations. First, the study focused on the specific case of physical activity logger applications and the Finnish young elderly segment, which is why future studies are called for to examine the generalisability of its findings to other types of digital wellness technologies and to the elderly population in general. Second, the research setting of the study does not fully correspond to the real-life market environment in

M. Makkonen, T. Kari & L. Frank:
A Follow-Up on the Changes in the Use Intention of Digital Wellness Technologies and Its Antecedents   561
Over Time: The Use of Physical Activity Logger Applications Among Young Elderly in Finland

which consumers make decisions on technology acceptance and use. For example, the participants were provided for free both the application as well as the training and support for setting up and using it, without which factors like facilitating conditions and price value may also have played an important role as antecedents of use intention. Third, there were some participants who left the research program already before T1 or between T1 and T2 or T2 and T3, and, thus, had to be omitted from the study. Some of them may have been individuals who would have reported very low scores in terms of use intention and its antecedents and whose omission, consequently, may have introduced some bias in the data. In future studies, we aim to address these aforementioned limitations and to augment the preliminary results of this study by refining our research model as well as collecting data from more participants and over a longer period of time as our research program progresses.

References

Beh, P. K., Ganesan, Y., Iranmanesh, M., & Foroughi, B. (2021). Using smartwatches for fitness and health monitoring: The UTAUT2 combined with threat appraisal as moderators. Behaviour & Information Technology, 40(3), 282–299. doi:10.1080/0144929X.2019.1685597

Carlsson, C., & Walden, P. (2016). Digital wellness services for young elderly – A missed opportunity for mobile services. Journal of Theoretical and Applied Electronic Commerce Research, 11(3), 20–34. doi:10.4067/S0718-18762016000300003

Dhiman, N., Arora, N., Dogra, N., & Gupta, A. (2020). Consumer adoption of smartphone fitness apps: An extended UTAUT2 perspective. Journal of Indian Business Research, 12(3), 363–388. doi:10.1108/JIBR-05-2018-0158

Duarte, P., & Pinho, J. C. (2019). A mixed methods UTAUT2-based approach to assess mobile health adoption. Journal of Business Research, 102, 140–150. doi:10.1016/j.jbusres.2019.05.022

Epstein, D. A., Ping, A., Fogarty, J., & Munson, S. A. (2015). A lived informatics model of personal informatics. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 731–742). New York, NY: ACM. doi:10.1145/2750858.2804250

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50. doi:10.2307/3151312

Hair, J., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017). An updated and expanded assessment of PLS-SEM in information systems research. Industrial Management & Data Systems, 117(3), 442–458. doi:10.1108/IMDS-04-2016-0130

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (2nd ed.). Thousand Oaks, CA: Sage.

Kari, T. (2014). Can exergaming promote physical fitness and physical activity: A systematic review of systematic reviews. International Journal of Gaming and Computer-Mediated Simulations, 6(4), 59–77. doi:10.4018/ijgcms.2014100105

Kari, T., & Makkonen, M. (2014). Explaining the usage intentions of exergames. In B. Tan, E. Karahanna & A. Srinivasan (Eds.), Proceedings of the 35th International Conference on Information Systems. Atlanta, GA: Association for Information Systems.

Kari, T., Piippo, J., Frank, L., Makkonen, M., & Moilanen, P. (2016). To gamify or not to gamify? Gamification in exercise applications and its role in impacting exercise motivation. In J.

Versendaal, C. Kittl, A. Pucihar & M. Kljajić Borštnar (Eds.), Proceedings of the 29th Bled eConference (pp. 393–405). Kranj, Slovenia: Moderna organizacija.

Macedo, I. M. (2017). Predicting the acceptance and use of information and communication technology by older adults: An empirical examination of the revised UTAUT2. Computers in Human Behavior, 75, 935–948. doi:10.1016/j.chb.2017.06.013

Makkonen, M., Kari, T., & Frank, L. (2020). Applying UTAUT2 to explain the use of physical activity logger applications among young elderly. In A. Pucihar, M. Kljajić Borštnar, R. Bons, H. Cripps, A. Sheombar & D. Vidmar (Eds.), Proceedings of the 33rd Bled eConference (pp. 567–582). Maribor, Slovenia: University of Maribor Press.

Makkonen, M., Kari, T., & Frank, L. (2021). Changes in the use intention of digital wellness technologies and its antecedents over time: The use of physical activity logger applications among young elderly in Finland. In Proceedings of the 54th Hawaii International Conference on System Sciences (pp. 1202–1211). Honolulu, HI: University of Hawai'i at Mānoa. doi:10.24251/HICSS.2021.147

McGarrigle, L., & Todd, C. (2020). Promotion of physical activity in older people using mHealth and eHealth technologies: Rapid review of reviews. Journal of Medical Internet Research, 22(12), article e22201. doi:10.2196/22201

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric Theory (3rd ed.). New York, NY: McGraw-Hill.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3. Bönningstedt, Germany: SmartPLS GmbH. Retrieved from https://www.smartpls.com

Roemer, E. (2016). A tutorial on the use of PLS path modeling in longitudinal studies. Industrial Management & Data Systems, 116(9), 1901–1921. doi:10.1108/IMDS-07-2015-0317

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611. doi:10.2307/2333709

Sun, F., Norman, I. J., & While, A. E. (2013). Physical activity in older people: A systematic review. BMC Public Health, 13, article 449. doi:10.1186/1471-2458-13-449

Talukder, M. S., Chiong, R., Bao, Y., & Malik, B. H. (2019). Acceptance and use predictors of fitness wearable technology and intention to recommend: An empirical study. Industrial Management & Data Systems, 119(1), 170–188. doi:10.1108/IMDS-01-2018-0009

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. doi:10.2307/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly, 36(1), 157–178. doi:10.2307/41410412

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics Bulletin, 1(6), 80–83. doi:10.2307/3001968

Yuan, S., Ma, W., Kanthawala, S., & Peng, W. (2015). Keep using my health apps: Discover users' perception of health and fitness apps with the UTAUT2 model. Telemedicine and e-Health, 21(9), 735–741. doi:10.1089/tmj.2014.0148

THE EFFECTS OF CONSUMER DEMOGRAPHICS AND PAYMENT METHOD PREFERENCE ON PRODUCT RETURN FREQUENCY AND REASONS IN ONLINE SHOPPING

MARKUS MAKKONEN,[1] LAURI FRANK[1] &
TIINA KEMPPAINEN[2]

[1] University of Jyvaskyla, Faculty of Information Technology, Jyvaskyla, Finland; e-mail: markus.v.makkonen@jyu.fi, lauri.frank@jyu.fi
[2] University of Jyvaskyla, School of Business and Economics, Jyvaskyla, Finland; e-mail: tiina.j.kemppainen@jyu.fi

Abstract In online shopping, product returns are very common. In order to reduce them, one must first understand who are making them and why are they being made. In this study, we aim to address these questions by examining product return behaviour from a consumer-centric rather than the more traditional product-centric, retailer-centric, and order-centric perspectives. More specifically, we focus on the effects of four demographic characteristics of consumers (i.e., gender, age, education, and income) as well as their payment method preference on their product return frequency and product return reasons. As the data, we use the responses from 560 Finnish online consumers, which were collected with an online survey and are analysed both quantitatively and qualitatively. We find gender, age, payment method preference, and average online shopping frequency to affect average product return frequency, whereas product return reasons were found to be affected by only gender and average product return frequency.

1 Introduction

In online shopping, returning a purchased product back to its seller is very common (Ofek, Katona & Sarvary, 2011). The product return rates in online shopping are about three times as high as in traditional brick-and-mortar shopping, and by 2022, about 13 billion products with a total worth of $573 billion have been forecasted to be returned annually in the United States alone (Deloitte, 2019). These high numbers can be considered problematic for both businesses and society at large. From a business perspective, product returns result in not only loss in sales but typically also in additional costs related to reverse logistics, return handling, and potentially throwing away the returned products if they cannot be resold (Ofek et al., 2011). In turn, from a social perspective, shipping products back and forth between sellers and buyers is far from being environmentally friendly or in line with the principles of sustainable electronic commerce (Oláh et al., 2019). Therefore, it makes sense to both business and society at large to aim at reducing product returns. However, to be able to do this, two fundamental questions must first be asked and answered: who are making the product returns and why are they being made?

In prior research, these questions have traditionally been approached from a very product-centric, retailer-centric, or order-centric perspective, whereas the studies adopting a more consumer-centric perspective have been rare (cf. Section 2). In this study, our objective is to address this gap in prior research by examining exploratively (without any *a priori* hypotheses) the effects of four demographic characteristics of consumers (i.e., gender, age, education, and income) as well as their payment method preference (i.e., how do they typically prefer to pay when shopping online) on their product return frequency (i.e., how often do they make product returns) and product return reasons (i.e., why do they make product returns). As the data for this, we use the responses from 560 Finnish online consumers, which were collected with an online survey in 2019 and are analysed quantitatively by using ordinal and binomial logistic regression as well as qualitatively by using content analysis.

After this introductory section, we will briefly discuss the theoretical foundation of the study in Section 2. This is followed by reporting of the research methodology and the research results in Sections 3 and 4. The results are discussed in more detail in Section 5 before concluding the paper with a brief discussion about the limitations of the study and some potential paths of future research in Section 6.

2   Theoretical Foundation

As already mentioned above, most of the prior studies on the antecedents of product return behaviour in the context of online shopping have focused on a very product-centric, retailer-centric, or order-centric perspective by examining how product return behaviour is affected by the factors related to the ordered product, the retailer from whom it is ordered, or the particular order transaction. Some examples of these factors are inventory availability, order delivery reliability, and expected order delivery timeliness (Rao, Rabinovich & Raju, 2014), assortment size and order size (Yan & Cao, 2017), retailer reputation (Walsh, Albrecht, Kunz & Hofacker, 2016), shipping and return fees (Lantz & Hjort, 2013; Lepthien & Clement, 2019; Shehu, Papies & Neslin, 2020), product reviews (Minnema, Bijmolt, Gensler & Wiesel, 2016; Sahoo, Dellarocas & Srinivasan, 2018; Wang, Ramachandran & Sheng, 2021), as well as package opening process (Zhou, Hinz & Benlian, 2018).

In contrast, far fewer studies have adopted a more consumer-centric perspective by focusing on the characteristics of the consumers who are ordering the products. Of them, in this study, we will focus on four demographic characteristics (i.e., gender, age, education, and income) as well as payment method preference. Our reason for selecting gender, age, education, and income as explanatory variables is based on the fact that although the effects of demographic characteristics on online shopping adoption have been examined in numerous prior studies (cf. Cheung, Zhu, Kwong, Chan & Limayem, 2003; Chang, Cheung & Lai, 2005; Cheung, Chan & Limayem, 2005; Zhou, Dai & Zhang, 2007), with the main focus being on the same four variables that we are focusing in this study (i.e., gender, age, education, and income), we are not aware of any prior studies that would have examined their effects on product return behaviour, although similar effects can be assumed to exist also in this context. In turn, our reason for selecting payment method preference as an explanatory variable is based on the fact that although we are not aware of any prior studies that would have examined the effects of payment method preference on product return behaviour, Yan and Cao (2017) have argued that the payment method used in a particular order (which once again is an order-centric rather than a consumer-centric factor) does affect product return behaviour. They also found support for this argument by observing that paying an order with a credit card results in a higher product return rate. Thus, rather than the payment method used in a

particular order, also payment method preference more generally can be assumed to have similar effects.

3   Methodology

The data for this study was collected in an online survey between February and March 2019. The respondents were recruited mainly by sharing the survey link through the internal communication channels of our university. In addition, because the respondents who completed the survey were able to take part in a price draw of ten cinema tickets, the survey link was posted to six websites promoting online competitions. The survey questionnaire was in Finnish and consisted of multiple items related to the demographics, personality, values, and online shopping behaviour of the respondents. This study utilises the responses to eight of those items. The first four items measured the gender, age, education, and income (as yearly personal taxable income) of the respondents. The fifth item measured payment method preference by using a closed-ended question with five answer options: a bank payment (a direct payment from a bank account), a card payment (a payment with either a bank or credit card), PayPal or MobilePay (the two most popular online payment services in Finland at the time, which charge the payment from one's linked bank account, bank card, or credit card), invoice (in one or multiple instalments), and cash on delivery (a payment when the order is delivered). In addition, the respondents had the option to state any other payment method if needed, but nobody used this option. The sixth and seventh items measured average online shopping frequency and average product return frequency by using closed-ended questions. The eighth and final item measured the most typical reasons for making product returns by using an open-ended question in which the respondents could state one or multiple reasons.

The collected data was analysed in three phases. In phase one, we used cumulative odds ordinal logistic regression to examine the effects of gender, age, education, income, and payment method preference on average product return frequency by using average online shopping frequency as a control variable. In phase two, we analysed the most typical reasons for making product returns by using content analysis in which we read each response, identified the reasons in them, and then tried to group them into more general categories based on common themes. In phase three, we used binomial logistic regression to examine the effects of gender,

age, education, income, and payment method preference on stating a specific reason for making a product return by using average online shopping frequency and average product return frequency as control variables. As the statistical software for conducting the logistic regression analyses, we used IBM SPSS Statistics 27.

4   Results

Table 1: Sample statistics

	N = 560		N = 462		N = 302	
	N	%	N	%	N	%
Gender						
Man	169	30.2	140	30.3	79	26.2
Woman	391	69.8	322	69.7	223	73.8
Age						
Under 30 years	262	46.8	230	49.8	149	49.3
30–49 years	201	35.9	156	33.8	113	37.4
50 years or over	97	17.3	76	16.5	40	13.2
Education						
Lower than tertiary education	222	39.6	185	40.0	109	36.1
Tertiary education or higher	338	60.4	277	60.0	193	63.9
Yearly personal taxable income						
Less than 15,000 €	223	39.8	216	46.8	130	43.0
15,000–29,999 €	102	18.2	101	21.9	67	22.2
30,000 € or more	145	25.9	145	31.4	105	34.8
Do not want to disclose	90	16.1	–	–	–	–
Payment method preference						
Bank payment	275	49.1	227	49.1	138	45.7
Card payment	96	17.1	83	18.0	54	17.9
PayPal or MobilePay	106	18.9	94	20.3	67	22.2
Invoice	74	13.2	58	12.6	43	14.2
Cash on delivery	9	1.6	–	–	–	–
Average online shopping frequency						
Yearly or less frequently	136	24.3	109	23.6	59	19.5
Monthly	361	64.5	299	64.7	205	67.9
Weekly	63	11.3	54	11.7	38	12.6
Average product return frequency						
Less frequently than yearly	338	60.4	277	60.0	143	47.4
Yearly	167	29.8	140	30.3	122	40.4
Monthly	55	9.8	45	9.7	37	12.3

In total, we received 580 responses to our survey. However, of them, we had to drop 20 responses due to missing or invalid data, thus resulting in a sample size of 560 responses to be used in this study. The descriptive statistics of this sample in terms

of gender, age, education, income, payment method preference, average online shopping frequency, and average product return frequency are reported in Table 1. In addition, for the analyses of phase one, we had to drop an additional 98 respondents who had not wanted to disclose their income or had preferred cash on delivery as a payment method, which was too small a category, thus resulting in a sample size of 462 respondents. In turn, for the analyses of phases two and three, we had to drop an additional 160 respondents who had not stated any reasons for making product returns, thus resulting in a sample size of 302 respondents. As can be seen from Table 1, these drops did not considerably change the sample profile.

4.1   Effects on Product Return Frequency

Before examining more closely the effects on product return frequency, we first checked the non-multicollinearity and proportional odds assumptions of cumulative odds ordinal logistic regression. The non-multicollinearity assumption was checked by using the variance inflation factor (VIF) values from basic multiple linear regression. These were all below two, thus suggesting no multicollinearity (Hair, Black, Babin & Anderson, 2018). In turn, the proportional odds assumption was checked by comparing the fit of a model with the proportional odds constrain to a model without the proportional odds constrain with a likelihood-ratio test. Its result ($\chi^2(10) = 17.417$, p = 0.066) supported the proportional odds assumption.

The estimated effects are reported in Table 2. All in all, the model was able to explain from 9.6% to 19.0% (McFadden (1973) $R^2 = 0.096$, Cox-Snell (1989) $R^2 = 0.158$, Nagelkerke (1991) $R^2 = 0.190$) of the variance in average product return frequency, fitted the data better than the baseline model with no explanatory variables (as suggested by the likelihood-ratio test), and had an overall good fit with the data (as suggested by the deviance goodness-of-fit test). The statistical significance of the effects was tested with the Wald (1943) χ^2 test, whereas the effect sizes are reported as odds ratios (OR) and their 95% confidence intervals (CI). For categorical variables, the effects are reported for a specific category in comparison to a reference category (in parenthesis). Additionally, if a variable has more than two categories, the result of an omnibus test is reported (on the same row as the name of the variable). As can be seen, gender, age, payment method preference, and average online shopping frequency were all found to have a statistically significant effect on average product return frequency, whereas the effects of education and income were

found to be statistically not significant. More specifically, women had 2.134 times greater odds than men of being more frequent returners, whereas the odds of being a more frequent returner decreased with age by an odds ratio of 0.975 per year. In terms of payment method preference, those who preferred paying by invoice had 2.999 times greater odds of being more frequent returners than those who preferred bank payments. Finally, as expected, in terms of average online shopping frequency, more frequent shoppers also seemed to be more frequent returners. That is, those who shopped monthly had 3.743 times greater odds of being more frequent returners than those who shopped yearly or less frequently, whereas those who shopped weekly had 5.932 times greater odds of being more frequent returners than those who shopped yearly or less frequently.

Table 2: Effects on average product return frequency

	Wald χ^2			Odds ratio	
	χ^2	df	p	OR	95% CI
Gender	–	–	–	–	–
Woman (vs. man)	10.972	1	< 0.001	2.134	[1.363, 3.341]
Age	7.033	1	0.008	0.975	[0.957, 0.993]
Education	–	–	–	–	–
Tertiary or higher (vs. lower than tertiary)	1.917	1	0.166	1.365	[0.879, 2.120]
Yearly personal taxable income	2.778	2	0.249	–	–
15,000–29,999 € (vs. less than 15,000 €)	2.070	1	0.150	1.477	[0.868, 2.513]
30,000 € or more (vs. less than 15,000 €)	2.033	1	0.154	1.485	[0.862, 2.557]
Payment method preference	13.579	3	0.004	–	–
Card payment (vs. bank payment)	0.911	1	0.340	1.303	[0.756, 2.246]
PayPal or MobilePay (vs. bank payment)	2.339	1	0.126	1.487	[0.894, 2.474]
Invoice (vs. bank payment)	13.179	1	< 0.001	2.999	[1.657, 5.425]
Average online shopping frequency	23.559	2	< 0.001	–	–
Monthly (vs. yearly or less frequently)	19.462	1	< 0.001	3.743	[2.082, 6.728]
Weekly (vs. yearly or less frequently)	19.913	1	< 0.001	5.932	[2.714, 12.967]

Likelihood-ratio $\chi^2(10)$ = 79.527, p < 0.001, deviance goodness-of-fit $\chi^2(760)$ = 618.853, p = 0.814
McFadden R^2 = 0.096, Cox-Snell R^2 = 0.158, Nagelkerke R^2 = 0.190

4.2 Product Return Reasons

When analysing the stated reasons for typically making product returns, we were able to identify four main reasons. These are listed and described in more detail below. The list also includes the number and the proportion of the 302 respondents who stated a specific reason. Note that one respondent could state multiple reasons.

- **Wrong size or bad fit** (stated by 193 or 63.9%): The most frequently stated reason was the wrong size or bad fit of the ordered product. This typically concerned products that are worn, such as clothes or shoes.
- **Mismatch with product information** (stated by 71 or 23.5%): The second most frequently stated reason was the mismatch of the ordered product with the product information provided by the retailer. For example, the product did not match the product description or product pictures in terms of colour, material, and quality, had some other deviances, or was an entirely wrong product.
- **Faulty or damaged product** (stated by 70 or 23.2%): The third most frequently stated reason was that the ordered product was faulty or damaged during delivery. In other words, there was some extreme quality issue in the product, which went well beyond the product not merely matching the product information.
- **Mismatch with needs, wants, or expectations** (stated by 58 or 19.2%): The most infrequently stated reason was the mismatch of the ordered product with one's needs, wants, or expectations. In other words, there was no obvious mismatch with the product information or other issues in the product, but one just did not like it, found it useless, or experienced buyer's remorse.
- **Other reasons** (stated by 6 or 2.0%): Finally, there were also a few respondents who stated some other reasons, such as suspicion of fraud, ordering the product just to meet some order limit, or returning the product just to spend time.

4.3 Effects on Product Return Reasons

Before examining more closely the effects on product return reasons, we once again first checked the non-multicollinearity assumption of binomial logistic regression by using the VIF values from basic multiple linear regression. These were all below two, thus suggesting no multicollinearity (Hair et al., 2018). When examining the

estimated effects, only two of them were found to be statistically significant. First, gender was found to have a statistically significant effect on stating wrong size or bad fit as a reason ($\chi^2(1) = 21.573$, $p < 0.001$) as well as stating a faulty or damaged product as a reason ($\chi^2(1) = 29.228$, $p < 0.001$), whereas average product return frequency was found to have a statistically significant effect on stating a faulty or damaged product as a reason ($\chi^2(2) = 21.872$, $p < 0.001$) as well as stating a mismatch with needs, wants, or expectations as a reason ($\chi^2(2) = 18.285$, $p < 0.001$). More specifically, women had 3.939 times greater odds than men of stating wrong size or bad fit as a reason, whereas men had 6.173 times greater odds than women of stating a faulty or damaged product as a reason. In turn, those who returned less frequently than yearly had 4.202 times greater odds than those who returned yearly and 12.987 times greater odds than those who returned monthly of stating a faulty or damaged product as a reason. In contrast, those who returned yearly had 2.086 times greater odds than those who returned less frequently than yearly of stating a mismatch with needs, wants, or expectations as a reason, whereas those who returned monthly had 8.034 times greater odds than those who returned less frequently than yearly of stating a mismatch with needs, wants, or expectations as a reason.

5   Discussion and Conclusions

In this study, we examined the effects of gender, age, education, income, and payment method preference on product return frequency and product return reasons. In terms of the effects on product return frequency, we found women to have greater odds than men of being more frequent returners and the odds of being a more frequent returner to also decrease with age. In addition, those who preferred paying by invoice were found to have greater odds of being more frequent returners than those who preferred a bank payment. Of these, the finding concerning the effect of payment method preference is largely in line with the study by Yan and Cao (2017), who found that paying an order with a credit card results in a higher product return rate. They explain this finding with the "buy-now-pay-later" mentality associated with credit cards, which is likely to result in more impulsive consumption behaviour and lower the threshold of making a product return because no exchange of money has yet occurred. A similar mentality is likely to also explain why preferring to pay by invoice results in a higher product return rate. In turn, the findings concerning the effects of gender and age are most likely explained by the different

online shopping habits of men versus women and younger versus older consumers. For example, women and younger consumers may be more likely to order products with higher return rates, such as clothes and shoes (Deloitte, 2019), whereas men and older consumers may be more likely to order products with lower return rates, such as consumer electronics (Deloitte, 2019). In addition, women may be more likely than men to order products for not just themselves but also others in their family, such as their children. In terms of age, there may also exist a generational gap. That is, older consumers, who are typically less experienced in shopping online, may make product returns more conservatively, resorting to them only when there is something severely wrong with the ordered product. In contrast, younger consumers, who are typically more experienced in shopping online, may make product returns more liberally, sometimes returning the ordered product even when there is actually nothing wrong with it. Or they may even practice bracketing, which means ordering multiple similar products with the intention of keeping only some of them and returning the rest. There were 12 respondents in our sample who explicitly mentioned doing this, and most of them were young consumers in their 20s.

In terms of the effects on product return reasons, we first identified four reasons why consumers typically make product returns: (1) wrong size or bad fit, (2) a mismatch with product information, (3) a faulty or damaged product, and (4) a mismatch with needs, wants, or expectations. These are largely in line with the reasons that have been identified in prior studies. For example, a study by Deloitte (2019) found the top five reasons to be (1) a too small or large size, (2) changing one's mind, (3) style not as expected, (4) not as described, and (5) a defective product. After this, we examined the effects on stating each of the four reasons, finding that women had greater odds than men of stating wrong size or bad fit as a reason, whereas men had greater odds than women of stating a faulty or damaged product as a reason. In addition, we also found that those who made product returns more frequently had greater odds of stating a mismatch with needs, wants, or preferences as a reason, whereas those who made product returns less frequently had greater odds of stating a faulty or damaged product as a reason. One explanation for the finding concerning the gender effect may be the fact that women more often shop online for products like clothes and shoes, in which wrong size or bad fit is likely to be an issue, whereas men more often shop online for products like consumer electronics, which are more prone to faults and more likely damaged during delivery. In turn, one explanation for the findings concerning the effects of

product return frequency may be the fact that if one has the tendency of making returns very rarely, then the reasons for those rare returns are likely to relate to some severe issue in the ordered product, such as it being faulty or damaged during delivery. In contrast, if one has the tendency of making returns relatively often, then it becomes less likely that the reasons for them only relate to actual issues in the product and more likely that they relate to things like the mismatch of the product with one's needs, wants, or expectations.

From a theoretical perspective, the main contribution of the study is the finding that consumer-centric factors like gender, age, and payment method preference – in addition to the product-centric, retailer-centric, and order-centric factors that have been more traditionally examined in prior research – can act as antecedents of product return behaviour by being able to explain a considerable amount of the variance in both product return frequency and product return reasons. In turn, from a practical perspective, the main contribution of the study is the implication that if the aforementioned factors indeed affect product return frequency and product return reasons, then online retailers can try to utilise these factors and effects in lowering their product return rates. For example, as those who prefer to pay by invoice were found to have greater odds of making product returns more frequently, some retailers may see it preferable not to offer invoicing as a payment method. Similarly, some retailers may see it preferable to target only the gender and age segments in which the product return frequencies are known to be relatively low.

6   Limitations and Future Research

This study can be seen to have three main limitations. First, because we focused only on Finnish online consumers, we cannot make claims on the generalisability of our findings to other countries. Second, because our sample was not entirely representative of the Finnish online consumer population especially in terms of gender and age, we also cannot make claims on how common or rare the identified product return reasons actually are. For example, wrong size or bad fit may have been found to be the most common reason simply because of the gender and age biases in our sample. However, we do not see these biases affecting our other findings concerning the effects on product return frequency and product return reasons because, by examining the effects simultaneously in one model, we

essentially controlled the effects of the other variables when examining the effect of a specific variable. Third, because most of the identified product return reasons are related to the root cause of there being something wrong with the ordered product, there is some conceptual overlap between them. However, we still consider them to give a good overview of the motivational aspects for why consumers make product returns. In future research, some interesting and important paths to follow would be to examine more closely the underlying mechanisms that cause the effects that we observed in this study as well as how the ongoing COVID-19 pandemic has potentially affected our findings.

References

Chang, M. K., Cheung, W., & Lai, V. S. (2005). Literature derived reference models for the adoption of online shopping. Information & Management, 42(4), 543–559. doi:10.1016/j.im.2004.02.006

Cheung, C. M. K., Chan, G. W. W., & Limayem, M. (2005). A critical review of online consumer behavior: Empirical research. Journal of Electronic Commerce in Organizations, 3(4), 1–19. doi:10.4018/jeco.2005100101

Cheung, C. M. K., Zhu, L., Kwong, T., Chan, G. W. W., & Limayem, M. (2003). Online consumer behavior: A review and agenda for future research. In R. T. Wigand, Y.-H. Tan, J. Gričar, A. Pucihar & T. Lunar (Eds.), Proceedings of the 16th Bled eCommerce Conference (pp. 194–218). Kranj, Slovenia: Moderna organizacija.

Cox, D. R., & Snell, E. J. (1989). Analysis of Binary Data (2nd ed). Boca Raton, FL: Chapman & Hall.

Deloitte (2019). Bringing it back: Retailers need a synchronized reverse logistics strategy. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/us/Documents/process-and operations/us-bringing-it-back.pdf

Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). Multivariate Data Analysis (8th ed.). Andover, United Kingdom: Cengage.

Lantz, B., & Hjort, K. (2013). Real e-customer behavioural responses to free delivery and free returns. Electronic Commerce Research, 13(2), 183–198. doi:10.1007/s10660-013-9125-0

Lepthien, A., & Clement, M. (2019). Shipping fee schedules and return behavior. Marketing Letters, 30(2), 151–165. doi:10.1007/s11002-019-09486-8

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), Frontiers in Econometrics (pp. 105–142). New York, NY: Academic Press.

Minnema, A., Bijmolt, T. H. A., Gensler, S., & Wiesel, T. (2016). To keep or not to keep: Effects of online customer reviews on product returns. Journal of Retailing, 92(3), 253–267. doi:10.1016/j.jretai.2016.03.001

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. Biometrika, 78(3), 691–692. doi:10.1093/biomet/78.3.691

Ofek, E., Katona, Z., & Sarvary, M. (2011). "Bricks and clicks": The impact of product returns on the strategies of multichannel retailers. Marketing Science, 30(1), 42–60. doi:10.1287/mksc.1100.0588

Oláh, J., Kitukutha, N., Haddad, H., Pakurár, M., Máté, D., & Popp, J. (2019). Achieving sustainable e-commerce in environmental, social and economic dimensions by taking possible trade-offs. Sustainability, 11(1), article 89. doi:10.3390/su11010089

Rao, S., Rabinovich, E., & Raju, D. (2014). The role of physical distribution services as determinants of product returns in Internet retailing. Journal of Operations Management, 32(6), 295–312. doi:10.1016/j.jom.2014.06.005

Sahoo, N., Dellarocas, C., & Srinivasan, S. (2018). The impact of online product reviews on product returns. Information Systems Research, 29(3), 723–738. doi:10.1287/isre.2017.0736

Shehu, E., Papies, D., & Neslin, S. A. (2020). Free shipping promotions and product returns. Journal of Marketing Research, 57(4), 640–658. doi:10.1177/0022243720921812

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society, 54(3), 426–482. doi:10.2307/1990256

Walsh, G., Albrecht, A. K., Kunz, W., & Hofacker, C. F. (2016). Relationship between online retailers' reputation and product returns. British Journal of Management, 27(1), 3–20. doi:10.1111/1467-8551.12120

Wang, Y., Ramachandran, V., & Sheng, O. R. L. (2021). Do fit opinions matter? The impact of fit context on online product returns. Information Systems Research, 32(1), 268–289. doi:10.1287/isre.2020.0965

Yan, R., & Cao, Z. (2017). Product returns, asymmetric information, and firm performance. International Journal of Production Economics, 185, 211–222. doi:10.1016/j.ijpe.2017.01.001

Zhou, L., Dai, L., & Zhang, D. (2007). Online shopping acceptance model – A critical survey of consumer factors in online shopping. Journal of Electronic Commerce Research, 8(1), 41–62.

Zhou, W., Hinz, O., & Benlian, A. (2018). The impact of the package opening process on product returns. Business Research, 11(2), 279–308. doi:10.1007/s40685-017-0055-x

COPYRIGHT ENFORCEMENT IN THE DUTCH DIGITAL MUSIC INDUSTRY

NERKO HADŽIARAPOVIĆ, MARLIES VAN STEENBERGEN & PASCAL RAVESTEIJN

HU University of Applied Science, Utrecht, The Netherlands; e-mail: Nerko.Hadziarapovic@hu.nl, marlies.vansteenbergen@hu.nl, Pascal.Ravesteijn@hu.nl

Abstract There is a lack of interest and empirical analysis in the existing literature on composers' relations with their publishers and the role of Collective Management Organizations (CMOs) within the system of music copyright. The purpose of this paper is to explore and understand the influence of digitization within the music industry on the copyright enforcement in the Netherlands and on rights holders and the CMOs. Also to explore and understand how their mutual relationships are affected by digitization of the music industry. A qualitative analysis was done by reviewing scientific literature, performing a documents analysis and doing open interviews. In the existing economics of copyright literature, the main focus is set on transaction costs, efficiency and welfare topics. The findings can be used to understand and model how rights holders and CMOs cope with the digitization and contribute to the policy makers and economic actor's discussion about future improvement of the copyright enforcement system.

1   Introduction

The advent of technologies, such as music streaming, poses a significant challenge to repertoire management and has led to failures regarding the compensation of rights holders in the music industry (Handke, 2010). According to Silver (2013) and Towse (2017) copyright law is becoming more complex in the attempt to keep up with each technological advance, especially where consumers and markets are in the lead. The impact of technological innovation on the music industry has stimulated research in economics of copyright and sparked the interest of policy makers (Belleflamme, 2016). The music industry is considered a forerunner in technological change and there are many lessons that can be learned from the music industry for the benefit of the entire Creative Industry (Lyons, Sun, Collopy, Curran & Ohagan, 2019). However, the focus of the economic copyright analysis has been on broader structures, leaving a need for structured knowledge building on the economic rationales and consequences at a micro level (The Allan Consulting Group, 2003). Bargfredde & Panay's (2015) make clear that one of the problems on micro level is that a significant part of the copyright fees are improperly distributed by the Collecting Management Organizations (CMOs)[1]. The unjust distribution of copyright money harms creators, is costly to the economy and has a negative impact on our society (Mahoney, 2015).

Recent discussions (Department for Digital, Culture, Media & Sport, 2017; Music Business Worldwide, 2018) on rates paid by Big Tech companies, such as Spotify, Apple, Amazon, Google and Facebook, to the Collective Management Organizations (CMOs) suggest that issues such as accountability and transparency regarding music use have not been completely resolved. For example, the CMOs collect the money but do not receive the usage data and thus cannot distribute the money to the rightful rights holders. While music is increasingly being consumed through digital channels (Williamson & Cloonan, 2012; Wikström, 2013; Samuel, 2014; Ingham, 2015) the number of empirical studies, particularly in the field of music copyright, is limited (Schlesinger & Waelde, 2012; Williamson & Cloonan, 2012; Phillips & Street, 2015; Towse, 2017), especially for the research on the impact of digitization on the rights holders of popular music. Hitherto, there are little empirical studies available that involve rights holders and their mutual formal

---

[1] Collective management organizations, such as collecting societies, typically represent groups of copyright and related rights owners, such as authors, composers, publishers, writers, photographers, musicians and performers.

and informal relationships. This research aims to fill that gap. Hence, the objective of this paper is twofold, first to explore and understand the influence of digitization within the music industry on the copyright enforcement in the Netherlands and on rights holders and the CMOs. Also to explore and understand how their mutual relationships are affected by digitization of the music industry.

Within the scope of this research, which focuses on the Netherlands, rights holders refer to music composers, lyricists and music publishers of popular music. CMOs in the Dutch context refers only to Buma/Stemra, the Dutch CMO appointed by the Dutch Government to collect the money for use of music and distribute the collected money to the rights holders. Buma/Stemra also has the responsibility to negotiate the tariffs for the use of music with different parties (users of music).

2   Theoretical Foundation

The economics of copyright literature beholds copyright as a theoretical economic stage where all the players are homogenous and rational. Also a stage where enforcement of copyright is perfect and where the relationships between practitioners are well defined and rational (Atkinson, 2012; Handke, 2012; Towse, 2017). Copyright research can be examined from multiple perspectives and includes law, technology, philosophy and economics (Handke, 2010; Wu, 2018; Lyons et al., 2019). Since much of copyright policy is about economics, it is important to understand the differences among different economic perspectives (Atkinson, 2012). Atkinson (2012) and Handke (2012) summarized key results in the empirical literature on copyright, put them into context and highlighted noteworthy gaps and contradictions in the literature. According to Atkinson (2012) the focus on transaction costs, efficiency and society welfare topics revolves around three 'classic' economic doctrines: conservative neoclassical; liberal neoclassical and neo-Keynesian. In the recent two decades a new economic doctrine has emerged, Innovation Economics, also referred to as neo-Schumpeterian or evolutionary economics. Innovation economics postulates that innovation (the development and adoption of new products, processes, and/or business models) drives growth (Atkinson, 2012). For studies of technological change in existing markets the neo-Schumpeterian or evolutionary economic literature provides a coherent and evidence-based foundation (Handtke, 2010). Technological change causes the spread of new products and production processes. Disruptive innovation is an

innovation that creates a new market and value network and thereby ultimately disrupts existing markets and value network (Ab Rahman et al., 2017). The products or services perceived as disruptive innovations tend to skip stages in the traditional product design and development process to quickly gain market traction and competitive advantage (Reyes-Mercado & Rajagopal, 2017). The actors are generally perceived as being different, for example with regard to their access to information, their ability to handle information, their capital and knowledge base (asymmetric information) or their routines (Nelson & Winter, 1982; Lipsey et al., 2005). These differences also apply to institutions designed to remain stable over time (Lundvall & Archibugi, 2001), but as the speed of technological change varies and is not always predictable, formal and informal institutions, technology and markets are 'out of sync'.

The music industry has rapidly digitized over the past 20 years. Legislation, institutions and CMOs are lagging behind these developments as there is a non-synchronous situation within the music industry (Lyons et al., 2019). Mostly national institutions (such as CMOs) deal with international 'Big Tech'-organizations from a skewed balance of power position. This is caused by the information asymmetry as Big Tech companies do not share available data with the CMOs and therefore have a much stronger negotiating position vis-à-vis CMOs and rights holders (Spoerri, 2019). Furthermore, the CMOs are not equipped to deal with the large amounts of data and the systems to convert this data into reliable information (Roberts, 2021). There is little empirical analysis on composers' relations with their publishers (contracting) (Towse, 2017) and the role of CMOs within the system of music copyright (Philips & Street, 20015; Watt, 2015), for example how the collected copyright revenue has been distributed amongst creators and other intermediaries (Towse, 2006).

3   Methodology

We believe that a better understanding of the phenomenon of digitization of music industry would allow the stakeholders to proceed from a more informed perspective in terms of designing, implementing and applying the future copyright enforcement system.

Qualitative research is grounded in an essentially constructivist philosophical position and its intent is to examine a social situation or interaction by allowing, us, the researchers, to enter the world of others and attempt to achieve a holistic understanding (Bogdan & Biklen, 2007; Locke et al., 2013; Maxwell, 2012; Merriam et al., 2015). In our view, considering the complex nature of the economics of copyright and the different economic doctrines outlined in section two, these grounds of qualitative research fit well with this study because its objective is to achieve a holistic and better understanding about the contemporary effects of digitization on copyright enforcement but also on the interactions between the stakeholders.

In order to select the sample for this study, a purposeful sampling procedure was used. Since one of us has been working in the Dutch music industry for over two decades, we started within our own network of possible participants. Also, a snowball sampling strategy was employed (Patton, 2015). The participants were selected using the following selection criteria: 1) composers and lyricists have had at least five songs released in the last 4 years, 2) they are registered members of Buma/Stemra and 3) either own their own publishing company or are represented by an official registered publisher in The Netherlands or elsewhere. Criteria in selecting publishers are that 1) they have a relevant repertoire of professional authors they represent, 2) they are professionally active in the copyright music industry for at least ten years. Finally, regarding CMO, the individual participants should have a management position within their organization with at least 5 years of relevant working experience. The delimiting time frames of 4, 10 and 5 years were decided to insure adequate working experience in the music industry. The research sample consists of six individuals included: two composer/lyricist with a broad repertoire of internationally successful songs who now own their own publishing companies (first one Grand Mono and the second one The Unexpected); a formal member of the Council of Rights Owners of Buma/Stemra (The Dutch CMO); the Dutch CEO of one of the biggest Global Independent Publishing companies (wishes to stay anonymous), a Buma/Stemra manager responsible for Business Development and a lobbyist of Buma/Stemra who operates on national and EU-level.

The following steps were used to carry out this research: 1) available literature and peer reviewed articles were searched, selected and analyzed, 2) collection and analysis of copyright-related documents (e.g. law and regulations) and 3) interviews with

participants. In relation to step 1, ongoing and selective review of literature was conducted. The main focus of the review was to acquire knowledge and gain understanding of the legal framework of copyright law, what the economics of copyrights are, how the enforcement system of copyright is designed and implemented and who the key stakeholders are within this system. In step 2 the associated activities were to name, collect, categorize and systematically analyze the relevant and available documents regarding the enforcement of copyright in The Netherlands. The collected documents were categorized in public and non-public documents. Besides literature this concerns at least the following documents: "income statements" from CMOs to rights holders; annual reports of CMOs (all public); the reports of the Supervisory Board for Collective Management Organizations for Copyright and Related Rights and available agreements (contracts) between publishers and composers and lyricists (non-public). Also, copyright law documents were considered. Although the legal framework of copyright lays outside the scope of this research, it can still provide important insights in the rationale and justification of copyright law from the legal perspective. The main focus of the document analysis was to gain a deeper understanding of the enforcement system of copyright law on meso and micro economical level. In step 3 six in-depth open interviews were conducted with participants who work in the Dutch copyright industry. This was the primary data collection method in this research because of its potential to elicit thick descriptions and enable us to search for additional information. A major benefit of individual in-depth interviews is that it also offers the potential to capture a person's perspective of an event or experience (Marshall & Rossman, 2014). In the case of this research our reason for choosing this method was that it is a good way to generate data through interaction with people and capture the meaning of their experience in their own words (Bloomberg & Volpe, 2019). Regarding the process of the interviews, we send emails and/or LinkedIn direct messages to prospective participants describing the purpose of the research with a request for a convenient date and time for an online interview. The interviews were conducted between December 2020 and February 2021. All the interviews were audio recorded and afterwards manually transcribed verbatim and with full permission of the participants. The interviews lasted between 45 minutes and 2.5 hours and covered different themes depending on the role and interests of the participants. Appendix 1 provides an overview of theme's and questions asked dependent on the type of interviewee. At the end of each interview the participants were asked if they could recommend a next potential participant. The data analysis

and data collection activities were done simultaneously in order to avoid the risk of repetitious, unfocused and overwhelming data (Merriam et al., 2015). The documents and transcripts were first coded with open coding for identifying and naming the data and developing major categories of information (Bloomberg & Volpe, 2019). In the next phase the categories were connected and we searched for relationships among them (Birks & Mills, 2015; Corbin & Strauss, 2014; Holton & Walsh, 2016), where we compared threads and patterns within categories. In the last phase of the synthesizing process, we situated the current work to prior research and compered and contrasted it with issues found in the broader literature (Bloomberg & Volpe, 2019). Credibility, dependability and confirmability of the research are ensured by triangulating sources (Patton, 2015) and member checks (Bloomberg & Volpe, 2019); transferability by purposeful sampling and thick descriptions (Gay et al., 2019; Merriam et al., 2015; Patton, 2015). For this process Atlas.ti software is used.

Although generalizability was not a goal of this study, through detailed description of the background and context, this study could be assessed for its applicability in other similar contexts. The findings are discussed below in section 4 of this article.

4   Findings

The major findings of this research are:

1. The literature study and document analyses contributed to the understanding on the practical application of the enforcement of copyright in The Netherlands;
2. All the participants indicated that the digitization affected the mutual relationships amongst creators, creators and publishers (rights holders) and CMOs. The relationships are now more complex and dynamic which results in different types of possible contracts between creators and publishers;
3. All participants acknowledged the effects of digitization on music copyright, complexity of contemporary system and existence of the 'old' legacy software used for the enforcement of copyright in The Netherlands;
4. All participants indicated that digitization of the music industry contributed to the existence of black boxes in the copyright processes and expressed the need for an appropriate solution;
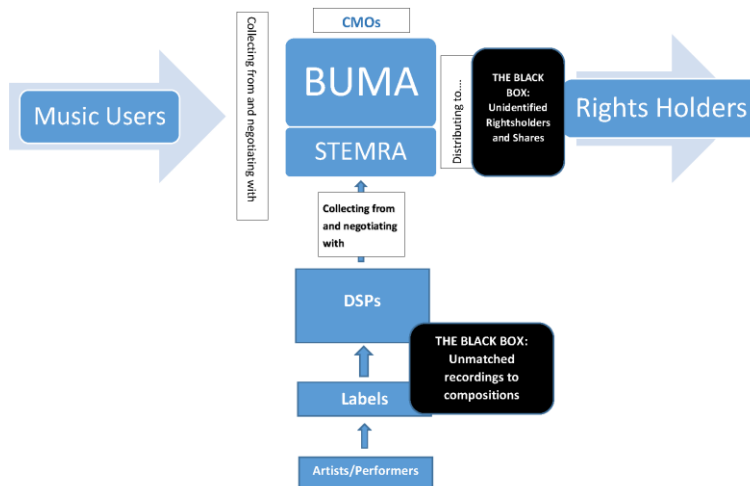
Figure 1: The black Box of the Music Copyright

Based on the document analysis the enforcement system of music copyright in the Netherlands is modeled (figure 1), including the stakeholders (players), their mandates and their relationships as formally described. According to literature and the analyzed documents, the 'users of music' pay for the use of music by annual or monthly contribution to the CMOs. The Dutch CMO, Buma/Stemra, is appointed by the Dutch Government to collect money from users of music and distribute the collected money to the rights holders. Buma/Stemra is also responsible for and given the mandate to negotiate the tariffs for use of music with different parties. The rights holders in the Netherlands are the composers, lyricists and the music publishers. The split of the copyright is divided equally by those three, each owns 33,33% of the copyright. In case of a composition without lyrics, this split is equal to 50%. The publishers are, depending on the signed agreements with the creators of music, responsible for the exploitation and administration of created musical works. There are different kinds of agreements between publishers and creators and the publishing share of 33,33% can (partly) flow back to the creators, depending on the type of contract (see table 1). When a musical work is created, the role of the creators is to

register their work with the CMO, in order to receive the revenue they are entitled to for the use of their work. Buma/Stemra is responsible for collection and distribution of performance rights and of mechanical reproductions rights. The latter is only relevant when a song or a composition is recorded by performers or artists and released (distributed) by, for example, a record label and reproduced on content carriers or digitally on for example Spotify or comparable online services. Registering a composition or lyrics for the rights holders is not experienced as convenient. As one of the interviewees stated:

"Imagine you write a song, you don't have a recording and someone else is performing it. What then happens is that you have to trust that there is always someone sitting there who writes down the title and the authors neatly and that that is copied well at Buma/Stemra, so that will be a bit of manual work. Nowadays there is also a lot of automation in it, but there is more margin of error in it." [Participant 1]

All the participants indicated that the mutual relationships amongst creators, creators and publishers (rights holders) and rights holders and CMOs are affected by digitization of the music and that these relationships are complex and dynamic. This results in different sort of agreements between creators and publishers. According to the participant who now owns his own publishing company:

"I worked with a publisher. I worked with them from 2013 to 2018. I felt that they were not doing enough and that they were not active enough with my music to justify getting such a share in my music." [Participant 5]

Based on the interviews we found that there are four possible contracts (table 1) between the creators of music (composers and lyricists) and publishers.

Table 1: Different contract types between creators and Publishers

Song / Title Agreement	This type of music publishing contract is an agreement between the writer and the music publisher in which the writer grants certain rights to a publisher for one or more songs. In a single song publishing contracts, the writer is sometimes paid a one-time recoupable advance.
Exclusive Songwriter Agreement ("ESWA")	Under the ESWA or "staff writer" contract, the songwriter generally grants all of the publishers share of the income to the music publisher. The writers' services are exclusive to the music publishers for a specified period of time. Thus, any compositions written within that period belong to the music publisher. These publishing contracts are usually offered to writers with some degree of commercial success.
Co-publishing Agreement ("Co-pub")	Under this deal, the songwriter and the music publisher are "co-owners" of the copyrights in the musical compositions. The writer becomes the "co-publisher" (i.e. co-owner) with the music publisher, based on an agreed split of the royalties (or kickback).
Administration Agreement ("Admin / Sub Publishing"):	Under this music publishing contract, the music publisher simply administers the copyrights for another publisher/copyright owner[2]. Under this coveted arrangement, ownership of the copyright is usually not transferred to the administrator. Instead, the music publisher usually gets 10-20% of the gross royalties received from administering the songs for a certain period of time and for a certain territory.

All participants acknowledged the effects of digitization on music copyright, complexity of the current system and existence of 'old' legacy software used for the enforcement of copyright in The Netherlands. During the times that music publishing was only based on exploitation of sheet music, the implementation of the system was uncluttered and relatively controllable. The contemporary and digitized music industry of today has become much more complex and intricate and there are now many more stakeholders in the music "ecosystem" than ever before.

"Enforcement and legislation lag behind technological developments, so once a law has been passed, after three years or so, the technology has already been developed in such a way that you can actually start working on a new law right away." [Participant 5]

---

[2] Publishers can only register their part of the copyright with the CMO, which has a maximum of 33,33% and cannot legally register the other two parts (composition and lyrics).

It has become almost impossible for the CMOs to collect and process all of the available data in order to collect and distribute the copyright money in the most fair and just way (figure 1). According to an interviewee who is both a composer and publisher:

"Buma / Stemra has to deal with hundreds of thousands of parties. That can often go wrong so in itself that is inherent to the system and there is nothing wrong with that. If your song is played on many thousands of TV and internet channels you cannot expect that everything will go smoothly. For authors, if you want to get what you are entitled to, you have to be on top of it."
[Participant 1]

And according to the interviewed manager of the Dutch CMO Buma/Stemra, there are more problems:

"We are still working with what is then called a monolithic system, so one large system that contains everything and that will at some point have reached the end of its life. Then you have to look for something new and a project has now started, which will of course take a few years before it is finished and rolled out, a new IT environment is developed and rolled out."[Participant 2]

The Netherlands is a relatively 'small player' compared to countries like Japan, USA, Germany, UK and France. Collecting and analyzing music using data from these countries (and many others) is almost impossible and very complicated.

"Of course we live in a digital age but a lot of that software is written by people so there are a lot of mistakes in it. That's just year after year, you know how it works, uh, IT is terribly difficult to get right year after year, patch after patch. Such a software system does not always improve…"
[Participant 1]

And according to the interviewed international publisher:

The fact is that you do not know what happens to your copyright and that the person who uses your copyright is actually not in breach at all. [Participant 3]

Another phenomenon of the music copyright industry has been discussed frequently in the recent global media: the black box of copyright (figure 1) (Bargfredde & Panay, 2015; Music Business Worldwide, 2018). All the participants indicated the effects of

digitization on existence of such black box of copyright and expressed the need for an appropriate solution. The black box is an 'umbrella' term used with different meanings. The most used definition is that these are unclaimed royalties collected by the CMOs. Basically, CMOs have collected the money but do not know who to give the collected money to. The reasons for the existence of such black boxes vary; from makers and publishers not registering their work, to labels releasing and reproducing the songs digitally without reporting the rightful owners and to unmatched databases or music users not correctly reporting the use of music (Music Business Worldwide, 2018). Also the digital data exchange between CMOs in different countries is a major reason for their existence. In words of the board member of Buma/Stemra:

"The black box within the copyright world means the following: money comes in and it is not clear how it is distributed. The black box is actually more of a collective name for various problems within the music copyright industry." [Participant 4]

"That black box is of course glued to everything they don't see…" [Participant 2]

5 Discussion, Conclusions and Recommendations

The purpose of this study was to explore the influence of digitization within the music industry on the copyright enforcement in the Netherlands and on creators of music, their publishers and the CMOs. Also to explore how their mutual relationships are affected by digitization. Following is the discussion of the findings and the conclusions drawn from this research.

5.1 The practical application of the enforcement of copyright in The Netherlands and the effects of digitization on music copyright

The first major finding of this research is that the design of the copyright enforcement system is well documented, transparent and institutionalized in the Dutch and European legal framework. The mandates and responsibilities are well defined and experienced as such by all the participants. However, there is a difference between the design of the system 'on paper' and practical application of the system. A conclusion to be drawn from this finding is that the designed system and the legal framework are rigid and not agile to adjust to the fast exogenous innovation. The digitization of the music industry started a tsunami of Big Data and

the key actors of the copyright enforcement, the CMOs, with the mandate to collect and distribute money from user to the rights holders are not ready to cope with the fast changing environment, are not equipped with the right software tools and their bargaining power towards the 'Big Tech' companies and the new major users of music, like Spotify, has diminished due to asymmetrical information. A further conclusion that can be drawn is that, although the justification of copyright in a broader sense is well-argued by scholars and policy makers, the implementation and the policy are not perfectly aligned, as one would expect from the findings in the literature covering the economics of copyright.

5.2   The mutual relationships are affected by the digitization

The second major finding is that all the participants have emphasized the existence of rather complex relationships between creators (composers and lyricist) and their publishers. For the legislation, the rights holders, creators and publishers, are homogeneous and enjoy the same rights. However, these two groups have different interests and their views on the distribution of income differ: "Artist versus the businessmen". In practice, these different views have led to the emergence of different forms of collaborations and different types of contracts between the two. One example is that on one hit **song,** there are sometimes more than 10 creators and more than 10 (sub)publishers involved, thus many contracts and splits between all parties involved exist. A related conclusion is that the digitization of the music industry enlarged the gap between the enforcement of copyright and the legal framework.

5.3   The existence of the black box of copyright

The last finding of this study are the effects of digitization of the music industry on the black box of copyright. All the participants were aware of the existence of the black box and indicated that it is a term used for not one, but many problems of the copyright enforcement. The overarching view of the participants is that the black box is an "umbrella term" used to describe the inability of the CMOs to distribute the collected funds to the correct rights holders. As stated before, the reasons for its existence vary, from outdated legacy software to data exchange problems between countries and the big tech companies withholding the data about the use of music

but also the efficiency reasons related to the transaction costs of the distribution to the somewhat smaller rights holders.

One of the limitations of this study is potential bias and subjectivity regarding one of the researchers own participation as a professional in the Dutch music industry and his personal experience with the enforcement of copyright in The Netherlands. The second limitation is that the research sample was restricted to six individuals, which could limit the knowledge produced by this study to be applied in other countries and similar contexts. We took the following measures once the possible limitations were recognized. First, a broad literature review and document analysis were inducted in order to recognize the research agenda and state the assumptions prior to the interviews. Secondly, the collection of data, analysis and findings were reviewed by faculty colleagues and advisors of this research. Although generalizability was not a goal of this study, through detailed description of the background and context, this study could be assessed for its applicability in other similar context.

Based on this research we find that further research should be conducted to gain more understanding about the current system of copyright enforcement and its complexities. As the number of participants to this research is limited, interviewing a larger number of active composers, lyricists, publishers, CMO-representatives and others involved, would contribute to the following objectives: 1) create more insights, 2) assess the extent to which the same or comparable findings can be found but also to uncover the similarities and differences in perspectives of the participants based on their role and position, 3) understand and model how creators, publishers and CMOs cope with the exogenous technological innovation in the music industry and 4) contribute to the policy makers and economic actors discussion about future improvement of the copyright enforcement system.

References

Ab Rahman, A., Hamid, U. Z. A., & Chin, T. A. (2017). Emerging technologies with disruptive effects: A review. Perintis E-Journal, 7(2), 111-128.

Atkinson, R. D. (2012). Copyright policy and economic doctrines. Information Technology and Innovation Foundation (ITIF). Retrieved from https://itif. org/publications/2012/11/26/copyright-policy-and-economic-doctrines.

Bargfredde, A., & Panay, P. (2015). Fair music: Transparency and payment flows in the music industry. Boston, Massachusets: BERKLEE ICE. Retrieved from

https://www.berklee.edu/sites/default/files/Fair%20Music%20-
%20Transparency%20and%20Payment%20Flows%20in%20the%20Music%20Industry.pdf

Belleflamme, P. (2016). The economics of digital goods: A progress report. Review of Economic Research on Copyright Issues, 13(2), 1-24.

Birks, M., & Mills, J. (2015). Grounded theory: A practical guide. Sage.

Bloomberg, L. D., & Volpe, M. (2019). Completing Your Qualitative Dissertation (Fourth Edi).

Bogdan, R. C., & Biklen, S. K. (2007). Research for education: An introduction to theories and methods. (5th edition). Boston, MA: Pearson Education

Corbin, J., & Strauss, A. (2014). Basics of qualitative research: Techniques and procedures for developing grounded theory. Sage publications.

Department for Digital, Culture, Media & Sport. (2017). Independent review of the creative industries. London: GOV.uk. Retrieved from https://www.gov.uk/government/publications/independent-review-of-the-creative-industries

Handke, C. W. (2010) The creative destruction of copyright: Innovation in the record industry and digital copying, EUR –Faculty of History and Arts, Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1630343

Handke, C. (2012). A Taxonomy of Empirical Research on Copyright-How Do We Inform Policy?. Review of Economic Research on Copyright Issues, 9(1), 47-92.

Holton, J. A., & Walsh, I. (2016). Classic grounded theory: Applications with qualitative and quantitative data. Sage Publications.

Ingham, T. (2015). Global record industry income drops below $15 bn for first time in decades. Music Business Worldwide,

Lipsey, R. G., Carlaw, K. I., & Bekar, C. T. (2005). Economic transformations: General purpose technologies and long-term economic growth OUP Oxford.

Locke, L. F., Spirduso, W. W., & Silverman, S. J. (2013). Proposals that work: A guide for planning dissertations and grant proposals. Sage Publications.

Lundvall, B., & Archibugi, D. (2001). The globalizing learning economy Oxford University Press New York

Lyons, F., Sun, H., Collopy, D., Curran, K., & Ohagan, P. (2019). The music data dilemma: Issues facing the music industry in improving data management: Music 2025. Retrieved from SSRN Electronic Journal, doi:10.2139/ssrn.3437670

Mahoney, L. (2015). Why equality and fairness matters . Retrieved from https://cubegroup.com.au/why-equality-and-fairness-matters/

Marshall, C., & Rossman, G. B. (2014). Designing qualitative research. Sage publications.

Maxwell, J. A. (2012). Qualitative research design: An interactive approach (Vol. 41). Sage publications.

Merriam, S. B., & Tisdell, E. J. (2015). Qualitative research: A guide to design and implementation. John Wiley & Sons.

Music Business Worldwide. (2018). How the global music copyright business grew by $1.5bn in 2016 (and why that's amazing news today). Retrieved from https://www.musicbusinessworldwide.com/how-the-global-music-copyright-business-grew-by-1-5bn-in-2016/

Nelson, R. R., & Winter, S. G. (1982). An evolutionary theory of economic change. Belknap, Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. Health Services Research, 34(5 Pt 2), 1189.

Patton, M. Q. (2015). Qualitative research & evaluation methods: Integrating theory and practice.

Phillips, T., & Street, J. (2015). Copyright and musicians at the digital margins. Media, Culture & Society, 37(3), 342-358.

Reyes-Mercado, P., & Rajagopal. (2017). Dynamics of disruptive innovations in outperforming global brands: A study in Mexico. International Journal of Business Excellence, 11(1), 1-1

Roberts Dave. (2021, March 10). 'streaming data files are getting bigger and bigger – they are not manageable manually anymore.'. Music Business Worldwide, Retrieved

from https://www.musicbusinessworldwide.com/streaming-data-files-are-getting-bigger-and-bigger-they-are-not-manageable-manually-anymore/

Samuel, M. (2014). Winds of change. Journey of UK music from the old world to the new world. Review of Economic Research on Copyright Issues, 11(2), 27-59.

Schlesinger, P., & Waelde, C. (2012). Copyright and cultural work: An exploration. Innovation: The European Journal of Social Science Research, 25(1), 11-28.

Silver, J. (2013). Digital Medieval, Xtorical Publications Media. Available from Amazon.

Spoerri, T. (2019). On upload-filters and other competitive advantages for Big Tech Companies under Article 17 of the directive on copyright in the digital single market. J. Intell. Prop. Info. Tech. & Elec. Com. L., 10, 173.

The Allen Consulting Group. (2003). Economic perspectives on copyright law. Centre for Copyright Studies Ltd. Retrieved from https://static-copyright-com au.s3.amazonaws.com/uploads/2015/08/CentreCopyrightStudies_AllenGroup-EconomicPerspectivesonCopyrightLaw.pdf

Towse, R. (2006). Copyright and artists: a view from cultural economics. Journal of economic surveys, 20(4), 567-585.

Towse, R. (2017). Economics of music publishing: copyright and the market. Journal of Cultural Economics, 41(4), 403-420.

Watt, R. (2015). The efficiencies of aggregation: an economic theory perspective on collective management of copyright. Review of Economic Research on Copyright Issues, 12(1), 2.

Wikström, P. (2013). The music industry (2nd Ed.). Cambridge: Polity Press.

Williamson, J., & Cloonan, M. (2012). Contextualizing the contemporary recording industry. The international recording industries (pp. 29-48) Routledge.

Wu, Z. (2018). No title. A Legal Framework for Global Joint Copyright Management in Musical Works: Based on Rawls's Theory of Justice

Appendix 1: Discussion Topics Interviews

	Participants Role		
	Composer/lyricist	**Publisher**	**CMO**
	Discussion topics	Discussion topics	Discussion topics
Theme: the practical application of the enforcement of copyright in The Netherlands	- How is the copyright system set up in the Netherlands? - How is the distribution of music rights organized in the Netherlands? - How do systems for registering works at Buma/Stemra work? - Which meta data is required to register a work at Buma/Stemra? - How does CMO distribute the collected funds? - To what extent are the creators aware of their rights and obligations with regard to copyright enforcement?	- How is the copyright system set up in the Netherlands? - How is the distribution of music rights organized in the Netherlands? - What is publishing and what roles does a publisher fulfill? - Which meta data is required to register a work at Buma/Stemra? - How does CMO distribute the collected funds? - To what extent are the creators aware of their rights and obligations with regard to copyright enforcement?	- How is the copyright system set up in the Netherlands? - How is the distribution of music rights organized in the Netherlands? - How do systems for registering works at Buma/Stemra work? - Which meta data is required to register a work at Buma/Stemra? - What is the role of CMO? - Which parties are the music users? - How does CMO distribute the collected funds? - How does CMO know where the money should go? - To what extent are the creators aware of their rights and obligations with regard to copyright enforcement?
Theme: relationships amongst creators, creators and publishers	- What contracts are possible between creators and publishers? - Why do you have a publisher or why do you not have a publisher?	- What contracts are possible between creators and publishers? - How does CMO know where the money should go?	

(rights holders) and CMOs	- Do all makers have the same interests or are there differences amongst them? - How do rapid changes affect relationships between creators and publishers?	- How do rapid changes affect relationships between creators and publishers?	
Theme: the effects of digitization on music copyright	- What about software systems at CMO? - What is the influence of technology on the copyright system? - To what extent are the users involved by Buma/Stemra, for example, in the development of such a registration portal?	- What about software systems at CMO? - What about alignment between EU legislation and technological developments? - What is the influence of technology on the copyright system? - To what extent are the users involved by Buma/Stemra, for example, in the development of such a registration portal?	- What about software systems at CMO? - How is the interconnectivity between different IT systems arranged? - How does CMO collect money from music users? - What about alignment between EU legislation and technological developments? - What is the influence of technology on the copyright system? - To what extent are the users involved by Buma/Stemra, for example, in the development of such a registration portal?
Theme: Black box	- To what extent is Buma/Stemra doing well? - What can they do better? - What is the role of CMO? - To what extent is the copyright system transparent?	- To what extent is Buma/Stemra doing well? - What can they do better? - What is the role of CMO? - To what extent is the copyright	- To what extent is the copyright system transparent? - Do things ever go wrong with regard to the collection and / or distribution of funds by CMO? - What is the copyright black box?

	- Do things ever go wrong with regard to the collection and / or distribution of funds by CMO? - What is the copyright black box?	system transparent? - Do things ever go wrong with regard to the collection and / or distribution of funds by CMO? - What is the copyright black box?	

INFORMATION REQUIREMENT IN THE TRANSITION TOWARDS A CIRCULAR FASHION INDUSTRY

MARJA EXALTO-SIJBRANDS & PASCAL RAVESTEIJN

HU University of applied sciences, Utrecht, Research Center of Digital Business and Media, The Netherlands; e-mail: marja.exalto-sijbrands@hu.nl, pascal.ravesteijn@hu.nl

Abstract Management of fashion (related) companies need to become convinced that circularity delivers positive financial results and incentives. This research aims to provide the first information requirement insights needed to enable the transition to a circular fashion industry. Due to easy access and abundant information 'Jeans' were selected as example item. Using the Design Science research approach the required information within in a closed loop supply chain (CLSC) in fashion was derived. Semi-structured interviews validated the CLSC information requirements derived from literature. Next, observations and additional literature findings supported the interview results. The outcomes show that information to support integration and collaboration of both: supply and recycle chain is necessary. Independently operating recycle organizations miss 'central loop management', 'information integration' and 'a chain-common objective' to successfully adopt circularity. The main bottlenecks found in relation to circularity are: 'overlooking the customer as stakeholder' and 'a lack of chain integration', this applies not only to jeans items. Therefore, the indicative study outcomes contribute to the body of knowledge of circular fashion value chain information requirements in general.

1   Introduction

The international fashion industry flourishes on large volumes of newly produced garments. With the expected growth in global population and average global wealth, the volume of discarded apparel is predicted to grow progressively. This, combined with fast fashion as business model will within 25 years lead to scarcity of clothing fibres. Therefore, the need grows for the fashion industry to addapt circularity (Ellen MacArthur Foundation (hence EMF), 2017). The business model of fast fashion stimulates unreasonably cheap and poorly produced items and blocks achieving affordable sustainable clothing for all global citizens in 2050 and beyond (Wicker, 2016). Sustainable development is often related to environmental improvements (Seuring, S., Müller, M., 2008), which in the fashion industry is adopted as: supplier's management limitation strategy, performance and/or image risk strategy (Seuring, S., Müller, M., 2008). These improvements are labelled as green washing and mask eagerness for short-term revenues, regardless its' effect (EMF, 2017). In 2016 this 'take-make-dispose-of' fashion business model delivered around 1.2 billion tons of greenhouse gasses; health threatening fertilizers and colouring chemicals; half a million ton of micro plastics in our oceans; and child labour practices (EMF, 2017). With the expectation of the indirect effects such as drought, heavy rains, and longer growing times, the future global production of new fibres is under pressure. When new fibre production falls short the need for reborn fibres will emerge. Also a growing market for second-hand items is expected (from 1% of the total fashion sales currently, up to 10% in 2030).

Since 2000 the first fashion entrepreneurs focused on circularity via innovative solutions of mechanical and chemical 'fiberization' of fashion items into reborn fibres. Unfortunately, not enough of these renewed fibres end up as new fashion items. Where the majority of Brands and Retailers in the Netherlands state that the perception of consumers repels from wearing reused fibres (MODINT, 2019), research from Morgen and Birtwistle (2009) and WRAP (2017) both show that consumer behaviour is able to adopt sustainability. Furthermore, currently the price for reborn fibres cannot compete with newly produced ones. The integrated supply chain (SC) with its strong price negotiation results in prices for new fibres that lay far below their true cost (BCI, 2018). Therefore, reborn fibres are less attractive as raw material. Promotion of Green Logistics and Circular Economy (CE) at education and entrepreneurs is of key importance (Seroka-Stolka, O., Ociepa-

Kubicka, A., 2019). Unfortunately, insights miss in how a closed loop fashion supply chain becomes an attractive and profitable business model. Therefore, the Dutch Industry Association for Re-winning textile fibres (VHT) aims to (1) reduce textile waste and (2) increase the potential of fibre reusability. They understand that the reusability capacity of fibres (in number and quality) requires measurements and a financial leverage of the circularity business case. In this lies the motivation for this research: delivering 'information requirements' within the circular fashion business model from fashion items to fibres as first step towards profitable circular fashion business cases.

2   Theoretical background

For the literature review the university's search engine HUGO and Google scholar were used with keywords related to linear supply chain fashion successes. The included standard terms were 'fashion' or 'textiles', and 'supply' or 'recycle chain'. These were combined with more specific terms such as: 'circular(ity)', 'performance indicators' and 'information management'. Compared to thousands of hits with the standard terms, the results of combinations with specific terms were limited. The most recent and best fitting articles (within the context of logistics) were selected. Prior to defining the key terms to this study, the research team discussed the relevance of each term based on an analysis on its value chain importance. Below we elaborate on the most important concepts related to this study.

Circular Economy

A circular economy regenerates its resources and commodities in order to utilize, maintain and recover products, their components and raw materials. The aim is to keep all products, components and materials at the highest value possible. EMF (2018) states that this viewpoint is needed in order to let commercial and non-commercial organizations adopt circularity across their supply chains, delivering new and different jobs and sustainability. Ashby (2018) sees the core of CE in recovering value from tangible commodities to narrow the definition of closed-loop to reuse and restoration. Battini, et al see (2017) 'Circularity' of an industry as option to reduce the environmental impact of the whole supply chain. In this research we combined the above into the following definition of CE: *an economy where companies close the supply chain as a loop with the aim to keep all products, components and materials at the*

highest value possible and to recover value from tangible commodities, extending job opportunities with new and different ones whilst delivering sustainability.

Closed Loop Supply Chain

As the definition substantiates, CE is connected to closed-loop supply chains (CLSC). CLSC maximizes value creation over the entire product life-cycle, starting at product design (McDonough and Braungart, 2002; EMF, 2013). It requires a control and operation system that also includes a dynamic recovery of the values used. This makes returns retrieved over the product life-time indispensable (Van Wassenhove and Guide, 2009). Where companies move from sustainability towards circularity, sustainable production systems increasingly become based on resource reuse and remanufacturing (Svensson, 2007; Angelis-Dimakis et al., 2016).

Recycle Chain

The first 'design for reuse' requirements were presented in regards to electronics (McDonough and Braungart, 2002). Accordingly, recycling became a Reverse Supply Chain (RSC) with a waste challenge perspective. RSC of electrical products followed a true 'reverse flow', returning the item to its manufacturer. This product return flow required additional logistics and performance measurement solutions that compared to those of the SC. Management information requirements were extended with a different kind of business economics combined with environmental metrics (Ahi, et al. 2015). However, recycled products do not necessarily follow a 'reverse' route. Recycling as such has developed as an autonomous operational activity that helps to limit waste and the product footprint. Accordingly, recycling should be recognized as a SC Comparable Activity that is also linear in behaviour. With the expectation that the Recycle Chain (RC) requires SC-comparable performance measurements, this study recognizes 'Recycling' *as activity within the RC, that requires SC comparable integration based on process information.*

The R-Ladder

Recycling has been acknowledged for its different product-life-stages presented as steps on a ladder (Netherlands Environmental Assessment Agency, 2019). The purpose of each step depends on the particular product-life-stage and product

characteristics. Each step is expressed with an 'R': Refuse, Rethink, Reduce, Reuse, Repair, Refurbish, Remanufacture, Repurpose, Recycle, and Recover. Each R-ladder step presents a different strategy towards the materials a product is made of. The strategy of the last step 'Recover' presents the recovery of all energy used during production of the initial product.

Circular fashion industry

'Circular fashion' is defined as: "Clothes, shoes or accessories that are designed, sourced, produced and provided with the intention to be used and to be circulated responsibly and effectively in society for as long as possible in their most valuable form, and hereafter return safely to the biosphere when no longer of human use" (Muthu, 2019). This definition is adopted because of the sustainability principles included, which form the bases of the information requirements this study searches for, being e.g. usage of biodegradable materials, toxic chemicals and pesticides, volume of re-used sustainable materials, and product quality (EMF, 2015).

Supply chain integration information

Logistics information and planning technologies are indispensable for SC integration, and SC process improvements subsequently result in higher business performance for chain partners (Rai et al., 2006). An integrated SC should be able to quickly react to market changes (Zailani & Rajagopal, 2005). Leuschner et al. (2013) stated the importance of information as part of supply chain integration. Furthermore, SC Integration improves financial performance when supported by top management (Zhao, Feng, Wang, 2015). Whereas SC integration is well known, RC integration is undescribed and integration between the SC and RC in order to create a CLSC is almost non-existent regardless the fact that a true CLSC requires sharing of information to support the total circular system. Therefore, it seems that the transition towards a circular fashion industry misses the holistic role of logistics.

Key terms commonality

This literature review emphasis that: *successful circularity* for fashion requires *an attractive circular business model* that concentrates on *minimizing waste, resource extraction and environmental impact* whilst keeping *focus on economic growth potential* (EMF, 2015;

McKinsey & Company 2014; Morgan, and Birtwistle, 2015) by *adopting logistics integration principles* known from the supply chain domain and *that goes beyond money* (Jonker, 2014).

3 Methodology

This research studies the life-cycle of Jeans, entailing: being bought and worn; re-used through second hand sales; recycled into a set of fibres to produce a new Jeans-item and the subsequent product life-cycle. Fibres missing the required quality level for reuse, end up at the R-ladder step of 'Recover'. To limit complexity, this research omits the R-ladder steps: Repair, Refurbish, Remanufacture, and Repurpose.

3.1 Research question

Based on interviews with VHT and insights from the literature review the research question adopted is: *What information is required to create an integrated Closed Loop Jeans (Fashion) Supply Chain?*

This descriptive and evaluating research delivers insight in information required to determine a Jeans CLSC business proposition. The theoretical perspective is checked in practice at three Jeans companies, fashion retailers, and additional companies such as three recyclers. Within these organizations the supply and recycle processes and supporting systems are studied.

3.2 Sub-questions and approach

The research was performed by answering four sub-questions.

1. *What is the current (circular) closed loop supply chain for Jeans within the fashion industry?*
   A circular supply chain fashion model by Wageningen University Research (2019) formed the bases to develop a circular fashion process model. Next, at two different conferences: "Logistics in the Circular Economy" and "Fashion and Design for Sustainability" we selected at random fashion professionals for participation in our research. They were asked to perform reliability and validation checks. This resulted into the concept Generic Circular Fashion Textiles Loop, which was discussed with VHT. Twenty

three companies who are involved in some manner with the fashion loop participated in a research project presentation in which the concept CLSC was discussed. During this research presentation (January 15th 2020) all twenty three professionals present approved the circular fashion model. With input from Jeans brands (desk research and interviews) the Jeans Closed Loop Supply Chain (Jeans-CLSC) was derived.

2. *What information is currently used per echelon to support jeans-item circularity?* Through desk research the logistics viewpoint was added to the basic Jeans-CLSC. This resulted in finding the required information drivers to enable circular chain control. To understand what information stimulates the business and the circular economy, four organisations (Jeans, Fashion retailer, Recycler and Branch organisation) were interviewed. Two researchers validated the findings that formed the 'IST" Jeans-CLSC.

3. *What information creates a state-of-the art circular Business model for jeans?* Desk research was performed on information requirements from documents of an existing state-of-the-art full CLSC Jeans brand. Additional, high performing circular loops within recycling such as paper, glass, plastic and disposal fee obliged equipment were studied to determine best-practices. Next, the defined State-of-The-Art Circular Jeans fashion Business model was discussed as a potential SOLL-situation with the circular Jeans manufacturer.

4. *What adoption is needed to extend the current Jeans fashion business into a circular one?* The difference between the 'IST' and 'SOLL' models was analysed with focus on business and circular economy (including sustainability) drivers and the information objects needed.

3.3 Design Science Research

Because insights in information requirements was the main goal, the research followed the Design Science Research approach by Hevner et al. (2004). Exploring the jeans supply and recycle chains and the relationship between these in regards to information requirements is the first step in Hevner's model, part of the 'Problem's Environment'. Determining information requirements to achieve a circular business model is done using the 'Existing Knowledge Base" of the SC (Hevner's model second step and theory based). Mapping the problem's environment with the existing knowledge base results in 'Designing a conceptual solution' (third step). The

designed solution needs testing prior to the real 'Solution Design'. This last step of Hevner's model is left for a next research project.

4   Results

The Jeans-CSCL and the related information requirements are presented below.

4.1   Circular Jeans Fashion

The described Jeans-CLSC (see Figure 1) compares to the generic CLSC processes, echelon types and order (see sub-question 1). Three circular loops are recognized: 1) the fibre recycle loop; 2) the B2C second hand market recycle loop; 3) the C2C loop where consumers sell to consumers. The large fibre recycle loop shows 13 echelons. The differences between Jeans-CLSC and general fashion-CLSC are found within the content of the processes. For example, due to the heterogeneous fibre mix used in most general fashion-items Step 6. – quality selection, is far more complex at fashion-CLSC than at Jeans. The circular loops described, show the most important transition points in achieving a true circular industry: 1) Consumer is currently excluded as CLSC stakeholder; 2) Value proposition between the SC (left) and RC (right) parts of the loop misses; 3) Fibre management misses (high value sorted fibres result into low value products loosing good fibres to non-fashion loops); 4) Circular pricing optimization misses. Reborn fibres should be able to compete with newly produced fashion fibres, which currently is impossible due to missing true price and true cost from sustainability factors.
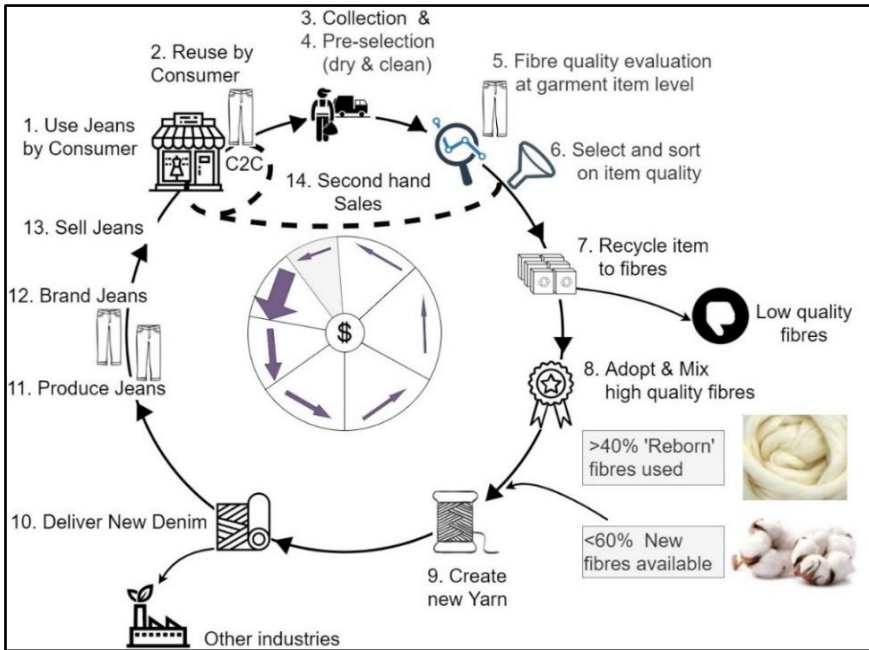
Figure 1: Circular Jeans Fashion

4.2   CLSC Information

Based on desk research of the typical current fashion SC and RC, the information requirements found and confirmed by the companies participating in this study are:

Customer demand (orders in numbers and volumes); SC integration information on demand, supply and performance to prevent inventory risks; Cost reduction factors resulting from integrated planning (e.g. on transport, warehousing and distribution); Image building result information on communication about adoption of bio and green aspects; and Sustainability performance indicators. Remarkable is that specific circularity information factors are uncommon.

Desk research and the five interviews at RC-companies show that these companies perform a single or limited number of echelon activities of the total RC. The need for integration within the linear RC has not yet been acknowledged, although this would result in comparable advantages as recognized for the integrated linear SC. Research revealed that high quality reborn fibres are downgraded to low-end

product markets (e.g. isolation materials, or cushion filings), where utilization as raw new yarn material was an option (Frankenhuis, 2019). Here circular economy is restricted to an open loop and the reborn fibres extend their product life outside the fashion industry. Companies within the RC operating autonomously limit their options to optimise their performance within a linear optimised RC as well as their paths for a transition to circularity (CLSC). This result leads to the hypothesis that the RC will also optimise the individual company performance by adopting integration as the SC has proven. Like the SC this requires information sharing. Without integration the RC companies are unable to reach a higher level of performance and thereby also influences the option for the fashion industry to become circular.

4.3   Finding state-of-the-art integration

Two circular jeans manufacturers (using 40% reborn fibres) acknowledged that only centrally orchestrated CLSCs successfully integrate the SC and the RC. Two generic fashion brands selling Circular Jeans (using 20% reborn fibres) underwrite this approach (Koppert, 2017). Therefore, integration of processes as common practise is required in both SC and RC as well as between these two chains. Their circular business model is based on: sales, information on sustainability, and adopting more manufacturing transparency at item level (Olugu, et al. 2010). Also, sustainable fibre information is required and volume and quality measurements based on items and fibres need to be shared for transparency within the circular value chain.

Next, the collector's decision to enable an attractive circular business model should be based on an Economic and Sustainability trade off (ESTO). This helps to appoint the collected item to the small recycle loop (second hand) or the large one (recycle into reborn fibre), as shown in figure 1. Which choice after collection is advised, depends on insight in fibre quality, actual costs, footprint calculations and sustainability cost effects. This is a complex trade-off.

Besides quality information at item selection, fibre adoption at yarn production also requires quality insights. For this, batch quality registration at 'fiberization' is required. Although it is expected that the number of times a fibre can be reused, lies between four to six times actual system tracking to deliver this information misses.

4.4 Transition towards circular Jeans business

Transition of the Jeans SC into a Jeans CLSC requires adoption of the RC activities and integration of all loop echelons including the customer as loop partner. Available information must stimulate the customer to return its item into the loop.

In practise collection of Jeans is done through the Jeans shops or via generic fashion item collection within multiple channels across different type of organisations. Research at two sustainable Jeans shops showed a significant loss of their annual item returns in spite of rewarding systems and lease constructions to stimulate consumers' return behaviour. This makes the item-offering process by the customer more important to consider rather than the collection process effectiveness and efficiency of the collector.

With the loop of Figure 1 as reference the information required to achieve a circular business model is presented in Table 1.

Table 1: transition purpose and information objects

Purpose	Information object
Total transparency in Jeans sustainability for customers on item level (from harvesting cotton to fibre recycling energy used).	• Denim cotton fibre mixing norm in volumes and types (batch wise information); • Sustainability performance on Denim yarn, Denim, and item production and specifics on e.g. labour and colouring processes (SER, 2016; EO, 2019); • Identification tracking information;
Total transparency in CLSC	• Company information of all stakeholders involved (name, website, location); • Full production information (production date, location, factory, material, labour used, quality check, human resources, etc.); • All cost and risk information of the integrated chain;
Sustainability image information per echelon	• Footprint (all negative effects on nature, plus restauration); • Positive effects on nature; • Labour conditions
USP to transform the consumer into a stakeholder	• Fibre information including fibre history details; • Return rate Jeans at point of collection ; • Item history;
Branding 'awareness'	• Measurement of 'honest production'

	• Earth preservation guarantee (rate)
Decision information recycle loop (small or long loop)	• Item quality norm; • Fibre quality norm; • Footprint; • Nature and labour effect measurements
Sustainability guarantee at item collection	• Intake quality measures: 'dry', 'clean', and 'reusable'
Increase of Consumer reuse rate	• Pre-selection business information on volumes, quantity, price and cost information as basic information; • Re-use norm set by sales in volume, quantity, price and cost;
Item attractiveness	• All item information travels with the item creating item history;

5   Conclusions, Recommendations, Limitations

This study shows that the acquisition process of materials in a CLSC forms a risk under uncertain quantity and quality of recycled products. Such uncertainties are revealed in the RC and subsequently effect the entire CLSC. These risks emphasize the importance of sharing fibre and item quality and quantity information. Currently 'Circularization' depends on the collector's decision how to collect and whether to appoint a collected item to the reuse or recycle loop, whilst this fundamental decision should be based on fibre quality information and should be in line with fibre demand information from the SC. Next, the business of mixing reborn with new fibres must become more attractive and transparent to the consumer. Mixing fibres at yarn production will only develop when the fibre market adopts a normal supply and demand system on true pricing. This means adopting a mechanism where farmers with transparent sustainable cotton production sell against fair prices, and where reborn fibres offer the fashion manufacturer a cost reduction. Due to fair cost sharing, only the integration between SC and RC supports this development.

For circularity the major challenge lies in attracting the autonomously operating companies of the RC to become performance oriented across the entire SC. An orchestrated (jeans) loop fails without integrated RC. Therefore, a Circular Business model becomes economically attractive when performance integration supports the pie growing and pie sharing capability of all stakeholders involved, including the customer. Also, the value flow needs to become circular, and the transition must be case focused rather than generic.

As with any study this research has limitations. First of all, the study focused on one type of fashion item (Jeans) which oversimplifies the challenges faced in the entire industry. Furthermore, a limited amount of organizations participated in this research and these could be considered front runners thereby providing a specific perspective. Finally, the main limitation to the outcome of this study is that it's outcome is a conceptual solution that still requires testing and follow up research.

Acknowledgements

References

Ahi, P. and Searcy, C. (2015). An analysis of metrics used to measure performance in green and sustainable supply chains. Journal of Cleaner Production. Vol 85, p.360-377.

Angelis-Dimakis, A., Alexandratou, A., Balzarini, A., (2016). Value chain upgrading in a textile dyeing industry. J. Clean. Prod., Vol.138, p.237-247

Ashby, A. (2018). Developing Closed Loop Supply Chains for Environmental sustainability: Insights from a UK clothing case study. Emerald, Journal of Manufacturing Technology Management. http://hdl.handle.net/10026.1/10898

Battini, D., Bogataj, M., Choudhary, A. (2017) Closed Loop Supply Chain (CLSC): Economics, Modelling, Management and Control. International Journal Production Economics. Vol.183 p.319–321 Retrieved 18-05-2020 via journal homepage: www.elsevier.com

Braungart, M., McDonough, W., Bollinger, A. (2007). Journal of Cleaner Production Vol. 15 p.1337-1348

BCI. (2018). Better Cotton Initiative. Retrieved from http://stories.bettercotton.com/2018-AnnualReport/index.html

Ellen MacArthur Foundation (EMF). (2018). Circular economy. Retrieved from www.ellenmacarthurfoundation.org/circular-economy

Ellen MacArthur Foundation (EMF). (2017). A new textiles economy: Redesiging fashion's future. Retrieved from http://www.ellenmacarthurfoundation.org/publications

Ellen McArthur Foundation (EMF). (2013). Towards the circular economy. Ellen MacArthur Foundation. Retrieved from https://www.ellenmacarthurfoundation.org/assets/downloads/publications/Ellen-MacArthur-Foundation-Towards-the-Circular-Economy-vol.1.pdf

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS quarterly, 75-105.

Jia, F., Yin, S. Chen, L., Chen, X. (2020). The circular economy in the textile and apparel industry: A systematic literature review. Journal of Cleaner Production Vol. 259 p.120728

Jonker, J. (2014). Nieuwe Business Modellen. Retrieved from https://www.nieuwebusinessmodellen.nl

Koppert, P., Brouw, H. o., & Adel, E. v. (2017). Roadmap Circulair Textiel.

Leuschner, R., Rogers, D.S., Charvet, F.F., (2013). A Meta-Analysis of Supply Chain Integration and Firm Performance. Journal of Supply Chain Management. Vol.49, No.2, p.34-57

McDonough, W. and Braungart, M. (2002). Cradle to cradle: remaking the way we make things. North Point Press.

McKinsey & Company. (2016). Style that's sustainable: A new fastfashion.

MODINT. (2019). Dossier MVO – Circulairiteit. Retrieved from https://modint.nl/dossier-imvo/

Morgan, L., & Birtwistle, G., (2009). An investigation of young fashion consumers disposal habits. International Journal of Consumer Studies. Vol. 33, p.190–198

Muthu, S.S., (2019). Circular Economy in Textiles and Apparel : Processing, Manufacturing, and Design. Elsevier Hongkong, e-book

Netherlands Environmental Assessment Agency. (2019). Circulaire economie in kaart. Den Haag: PBL Planbureau voor de Leefomgeving.

Olugu, E.U., Wong, K.Y., Shaharoun, A.M. (2010). A Comprehensive Approach in Assessing the Performance of an Automobile Closed-Loop Supply Chain. Sustainability. Vol. 2, p.871-889 Retrieved via doi:10.3390/su2040871

Rai, A., Patnayakuni, R., Seth, N., (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. MIS Q. 30 (2), 225–246.

Seroka-Stolka,O., Ociepa-Kubicka, A. (2019). Green logistics and circular economy. Transportation Research Procedia. Elsevier B.V. Vol. 39 p.471–479

Seuring, S., Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. Journal of Cleaner Production Vol. 16, p.1699–1710

Svensson, G. (2007). Aspects of sustainable SCM (SSCM): conceptual framework and empirical example. SCM: Int. J., Vol.12, Iss.4, p.262-266

Van Wassenhove, L.N. and Guide, . (2009). The Evolution of Closed-Loop Supply Chain Research. Operations Research. Vol. 57, No. 1, pp. 10–18

Wicker, A. (2016, september 1). Fast fashion is creating an environmental crisis. Newsweek.

WRAP. (2017). Clothing durability report.

Zailani, S., & Rajagopal, P., (2005) Supply chain integration and performance: US versus East Asian companies. Emerald Insights - Supply chain integration and performance. Retrieved from DOI:

https://doi-org.hu.idm.oclc.org/10.1108/JIMA-11-2015-0082.

RESPONSIBLE AI AND POWER: INVESTIGATING THE SYSTEM LEVEL BUREAUCRAT IN THE LEGAL PLANNING PROCESS

ROB PETERS,[1] KOEN SMIT[2] & JOHAN VERSENDAAL[2]

[1] Province of Utrecht, the Netherlands; e-mail: rob.peters@provincie-utrecht.nu
[2] HU University of Applied Sciences Utrecht, Digital Ethics, the Netherlands; e-mail: koen.smit@hu.nl, johan.versendaal@hu.nl

Abstract Numerous statements and pamphlets indicate that governments should increase the transparency of ICT-implementations and algorithms in eGovernment services and should encourage democratic control. This paper presents research among civil servants, suppliers and experts who play a role in the automation of spatial policymaking and planning (e.g. environment, building, sound and CO_2 regulation, mobility). The case is a major digitalisation programme of that spatial planning in the Netherlands. In this digital transition, the research assumption is that public and political values such as transparency, legitimacy and (perceived) fairness are difficult to validate in the practice of the design process; policy makers tend to lose sight of the algorithms and decision trees designed during the ICT-implementation of eGovernment services. This situation would implicate a power shift towards the system level bureaucrat. i.e., the digitized execution of laws and regulations, thereby threatening democratic control. This also sets the stage for anxiety towards ICT projects and digital bureaucracies. We have investigated perceptions about 'validation dark spots' in the design process of the national planning platform that create unintended shifts in decision power in the context of the legal planning process. To identify these validation dark spots, 22 stakeholders were interviewed. The results partially confirm the assumption. Based on the collected data, nine validation dark spots are identified that require more attention and research.

1   Introduction

The danger of unattainable algorithms in eGovernment services has been highlighted by popular and scientific heralds alike (Kool, Timmer, Royakkers, & Est, 2017; van den Hoven, Miller, & Pogge, 2017; Vereniging van Nederlandse Gemeenten, 2019). For example, the UK-based TV show Little Britain portrayed the citizen as a powerless victim of digitalised procedures already in 2004 with a series titled "computer says no". For this research, we focus on the Dutch situation in which the problem domain seems well recognisable. The Raad van State, the highest Dutch National advisory council has been addressing the problem of 'dehumanization' of government services in many reports, e.g. (van den Hoven et al., 2017). This growing anxiety with what computers can do to us in a bureaucracy grown beyond our control sparked a wave of pamphlets on digital ethics, computer ethics and general principles for government services (European Commission, 2019, 2020; gemeente Eindhoven, 2019; van den Hoven et al., 2017). The Dutch National society of municipalities and cities, the 'Vereniging van Nederlandse Gemeenten' (VNG) announced the agreement on digital principles and values (Vereniging van Nederlandse Gemeenten, 2019). The Dutch National Digital agenda 2020 included a whole chapter on ethical values and principles for digitalisation (Digitaleoverheid, 2020). The European Commission recently published a "white paper on Artificial intelligence and administrations" that lists several requirements (European Commission, 2020). This pamphlet promotes the use of AI in administrations, but points at issues of trust at the same time.

Yet, there still seems to be a gap between such general ethical ideas on digital public values and the operationalisation in current government processes. The Dutch Government Review board recently published a report including an auditing framework for algorithms that at least provides an 'auditing framework for algorithms' (Algemene Rekenkamer, 2021). We also observe, from the practitioner's side, that vendors claim transparency and open standards in their offerings without much substance, so one needs to dig deeper to determine the fairness of algorithms.

From a scientific perspective, the landscape seems partly covered. There is a body of knowledge in Business Rules Management (BRM) that is already touching on eGovernment services (Schlosser, et.al 2014). There has been twenty years of research on AI and Law addressing the automation of legal bureaucracies, see for

R. Peters, K. Smit & J. Versendaal:
Responsible AI and Power: Investigating the System Level Bureaucrat in the Legal Planning Process

613

example the Jurix, ICAIL, and IFIP conferences that started in the eighties. A growing population of researchers addresses the field of digital ethics (van den Hoven et al., 2017), (Vetzo, Gerards, & Nehmelman, 2018), and (van Engers & de Vries, 2019). Janssen et al., (2020) recently argued for a distinction between rule-based AI algorithms and Machine Learning-based algorithms in government applications, because of the demands on transparency. His focus was on the nature of the AI. Machine learning could be used to improve the design of the rule-based systems, however (Janssen et al., 2020). We stated earlier that there is a gap between high level ethical norms regarding transparency, fairness, discrimination regarding algorithms and the applicability of these norms in practice by civil servants. The theory according to Lipsky (1980) and Bovens & Zouridis (2002) states that the power of the system level bureaucrat will increase because of the instalment of algorithms in digital systems or platforms that automate the decision space of the street level civil servant, the street level bureaucrat. On the one hand, we adhere to the definition of Lipsky (1980) of a street level bureaucrat: "*a public employee that deals directly with citizens in the course of their jobs, and who have substantial discreation in the execution of their work.*" On the other hand, we adhere to the definition of Zouridis, van Eck & Bovens (2020) of a system level bureaucracy/bureaucrat: "*the discretionary powers of the street-level professionals have been disciplined by digital systems, and the locus of administrative discretion has shifted to those responsible for programming the decision-making process and translating the legislation into software*".

The transition of street level bureaucracy towards system level bureaucracy complicates the execution of legislation, because a translation has to be done in order to implement the legislation in the information systems replacing the system level bureaucrats. This translation offers room for a power shift towards the information systems as well as that such a translation and its output need to be validated thoroughly. If the latter fails, organizations could face severe consequences, such as lawsuits, high fines, negative publicity as well as political outfall (Smit, Versendaal, & Zoet, 2017; Smit, Zoet, & Berkhout, 2017). We, therefore, investigate the potential 'validation dark spots' in the design process where legislation is translated into information systems. With validation dark spots we mean those areas in the design process where participants of that design process identify potential unintended shifts in power between these two levels, i.e., street-level bureaucracy versus system level bureaucracy. Bajec & Krisper (2005) describe BRM research as follows: "*we presume that the ultimate goal of business rules research is to find a way and facilities that support automatic propagation of changes to business policies, respectively the business environment, to information*

systems used within organizations. The term validation is used to describe the integrity of the translation of law and regulations as well as internal business policies into information systems. Formally, it is referred to as the 'verification' and 'validation' capabilities (Smit & Zoet, 2018a; Zoet & Versendaal, 2013). Smit, Zoet, and Berkhout (2017) applied verification and validation capabilities on governmental legislation in search of levels of compliance of the actors involved. They came up with 28 verification capabilities (Smit, Zoet, et al., 2017). Other, similar studies and results, which focused on the identification and classification of verification capabilities for BRM are detailed in the work of (Corea, 2021). However, such research has only been partially conducted with regards to the validation capability. Therefore, we specifically seek validation dark spots in this study where the translation process from political/legal norms and values into information systems affect the position of the street level bureaucrat. An adequate case is found in the Dutch Omgevingswet, which is further detailed in the following sections. To investigate this case in search for validation dark spots, the following research question is addressed in this paper: '*What are the validation dark spots in the decision power shift from street-level bureaucracy towards system-level bureaucracy caused by digitization in the context of the Dutch Omgevingswet?*'

2   Background and Related Work

Within our object of study, we aim at the area of services where government affects the life events of citizens by means of permits and urban design decisions. A life event is defined as a "*A social experience or change with a specific onset and course that has a psychological impact on the individual.*" Examples are starting a new business, parental divorce, house relocation or school changes (Goodyer, 1991). The business rule-based algorithms are used to balance interests concerning economy, safety, mobility, housing, and ecology. Broad policy intentions are set in regulations and those are translated into an information system that deals with permit information and permit processes. The planner provides contextual information for the citizen when this person is asking for a permit to build or develop an object in the region. This information addresses both the permit requirements, the situational context and the process. The citizen will start providing relevant information digitally (forms) in a permit process. The translation of policy intentions to regulations that are in turn being translated into information systems is not without problems (Smit, Versendaal, et al., 2017; Smit & Zoet, 2018a, 2018b; Smit, Zoet, & Versendaal, 2018). The

R. Peters, K. Smit & J. Versendaal:
Responsible AI and Power: Investigating the System Level Bureaucrat in the Legal Planning Process

615

traditional policy maker offered policy goals up for decision making on the political platform and the resulting consensus was used and translated into juridical articles and norms. These norms were then interpreted, assessed, and adapted to the current situation by the street level bureaucrat for each individual case in a paper or email procedure. The system level bureaucrat, on the other hand, regards the current state of the data itself, as aggregated in a spatial data infrastructure, and applies business rules and standards in digital services to achieve the consensus policy goals and adapts the outcome of case decisions to the dynamic status of the data for each individual case. The assessment process is increasingly being automated by standard business rules in the balancing algorithms (Zouridis et al., 2020). This would be fine if all the translation steps were verifiable and could be validated by policy makers. To investigate this phenomenon, we want to determine to what extent and on which aspects decision power is being transferred from the street level civil servants towards the system designers when spatial planning legislation is being translated into information systems and algorithms. To do so, we need to define what decision-power in the context of spatial planning comprises.

2.1 Definitions of (decision) power in spatial planning

Bovens and Zouridis (2002) define the 'street level bureaucrat' as the executionary arm of government. The terms used for the level of freedom of the street level bureaucrat is the discretionary power to apply regulations on specific cases with autonomous space for interpretation. The European legislative level may set the framework for Natura2000 biodiversity, for example, and the Provinces may set the protected contours of that biosphere, but the individual street level civil servant decides on the legitimacy of the compensation actions offered against a building permit in that area.

The Omgevingswet programme (Koninkrijksrelaties, 2018) was introduced in politics and towards citizens as a simplification of the environmental legal arena to decrease complexity and increase user friendliness. Twenty-two regulations regarding, for example, soil, air quality, Natura2000, biodiversity, and water quality would be reduced to one all-encompassing legal structure as was applied in New Zealand. The digital platform Omgevingswet would take over some of that executionary role by introducing forms and decision trees based on business rules that guide the citizen through permit processes and pre-calculated levels of compensation required. The level of freedom of the street level bureaucrat in our

example of compensation for a building permit would be replaced by the team that designs the forms and the business rules behind them. Discretionary power can be based on two different sources: 1) The translation from general rules as set at a higher level of jurisdiction into the specific application of a case of that rule, and 2) the semantic room for interpretation. Together these add up to the effective decision power that is left after digitalisation of the process. The distinction between these two is not always straightforward and the legislator is not precise enough in its instructions (Peters, 2016; Teuben, 2004). It is also clear that when the business rules driving the relevant forms for obtaining the permit are translated into information systems, the translation is carried out by other specialists than the civil servants who represent the government agency.

3 Research Method

Bovens and Zouridis (2002) define the system engineer as a central role in the information system design process, but the question is what this role actually represents. In the case of the Omgevingswet digital platform, there are system engineers, legal knowledge engineers, legal planning experts, business rule specialists, domain specialists, programme managers, umbrella organisations of cities and regions, consultants and supplier-side developers who all have influence in the design process. The introduction of the Omgevingswet digital platform is chosen as the case for this explorative research because it represents more than just another case. Unlike many other studies about BRM in social security, immigration services, or tax returns, this platform will digitalise the main government service process of all Dutch cities. It is therefore a true situation of 'street level bureaucrats' and much less controllable by a small group of super experts hired by, for example, the National Tax office in a centralized approach. This legal planning platform is complex and thousands of civil servants, lawyers, developers, consultants, and project managers have been involved. The Platform is to be launched, after two failed deadlines, in 2021 when all 350 Dutch cities will have to transfer permit procedures to this digital platform.

To identify the relevant validation dark spots in the context of the development of the Omgevingswet digital platform, we interviewed 22 representatives from various stakeholder groups. The interviews took place between January 2020 and February 2021. The first eleven interviews were carried out online due to COVID-19

R. Peters, K. Smit & J. Versendaal:
Responsible AI and Power: Investigating the System Level Bureaucrat in the Legal Planning Process

617

restrictions. The respondents were all given the same description of the 'system level bureaucracy' based on Lipsky (1980) and van Bovens and Zouridis (2002) in advance. The first eleven interviews were conducted using a semi-structured approach. This works best when discussing certain topics while maintaining proper space for participants to digress on similar relevant topics or arguments (Pervan & Maimbo, 2005). This approach is in line with the explorative nature of this study. All interviewees have experience with the design process of the Omgevingswet digital platform. In total, three IT experts, three policy advisors, two BRM system suppliers, two GIS system suppliers, and one BRM scientist were interviewed. An interview protocol was standardized and utilized across the eleven semi-structured interviews, featuring 5 themes being: 1) role of the system designer, 2) room for interpretation of policies in the design process by the system designer, 3) room for decision power to make changes to the design of the policies that must be implemented in order to digitize them, 4) measures to control the quality of the system design (validation), and 5) the Omgevingswet. Each theme featured multiple questions to guide the interview and enable comparability of the results across all interviews to get a holistic view of the phenomenon of validation dark spots. In total, the protocol featured twenty-six questions. Additionally, the focus of the interview was scoped towards the Omgevingswet digital platform creation process over the last three years. The interviews were transcribed and thematically coded by one researcher and reviewed by another researcher. Furthermore, 11 non-structured open-ended interviews were conducted a few months after the first 11 interviews to gather more contextual information about the Omgevingswet platform, which were also conducted online. These interviews were only guided by the concept of the Dutch Omgevingswet and the Omgevingswet platform in development. In total, four National programme managers, two platform architects, one regional project manager, one expert on water management policies, one legal expert, one GIS expert and one BRM-system supplier were interviewed using this technique. The interviewees for the latter 11 interviews were selected based on their involvement from different perspectives as well as that they did not yet participate in the first round of 11 interviews.

4   Research Findings: the Validation Dark Spots

Based on the interview data and thematic coding, nine validation dark spots could be identified. These are further detailed in the subsections below.

4.1   The perception of power shifts

We first needed to establish whether the interviewees indeed perceive and recognize the coming of a system level bureaucracy as described by Lipsky (1980) and Bovens & Zouridis (2002). Most of the respondents agreed that a larger part of the environmental law shall transform towards a system that resembles a system level bureaucracy because of the characteristics as described by Bovens and Zouridis (2002). However, there appears to be a difference in opinions between the more technical oriented designers and the policy designers with respect to the effect of that transformation. The policy designers seem to think that the decision process will be automated entirely, including the balancing and prioritization of variables. The technically oriented designers think that the processes around permits are being digitalised without touching on the balancing and prioritising itself. During the intake of a case, the initiator of the permit request is confronted with a decision tree of choices and variables that then feed a workflow of steps along relevant authorities and governmental experts. The case is then processed by these experts, depending on its complexity. The confusion about the level of automation of the decision has several origins:

- The design of the decision tree is not without choices that affect the decision itself;
- The technical people are more aware of the limitations of the technology whilst the non-technical interviewees experience a 'feeling of being taken over' by the technology;
- The case-handling is supported by templates created by central government. Smaller cities copy these templates due to a lack of financial and knowledge resources. By doing so, these smaller organizations standardize their decisions unintentionally, which in turn shifts more influence towards system level bureaucrats designing the templates.
- The technically oriented system designers tend to make a distinction between simple and repetitive cases that could be automated and complex and more unique cases for which human intervention will always be required. They see the value of

R. Peters, K. Smit & J. Versendaal:
Responsible AI and Power: Investigating the System Level Bureaucrat in the Legal Planning Process

619

automation on top of digitalisation for reasons of efficiency. Policy experts do not always see where automation may be applicable to save resources.

Upcoming paragraphs describe and discuss the identified dark spots.

4.2 The mapping problem between the case and the platform characteristics

The first validation dark spot concerns the mapping between the use case and balancing rulings. The policy making respondents voiced their concerns about the ability of the platform to map 'their' reality into the correct and relevant decision trees. The 'permit checker module' of the platform that enables initiators to test if the activity can be performed free of permit restrictions, is a good example of this phenomenon. Legal experts take the position that once the computer says "no', you do not require a permit based on the decision tree, norms and values', the initiator is now allowed to proceed, even when the decision (or decision tree) was based on faulty assumptions or the wrong application of norms. This means that the capturing of the relevant data through fill-in forms beforehand (knowledge acquisition and elicitation in terms of BRM) is vital in the eyes of legal experts in order to maintain constitutional and legitimate.

4.3 Contextual information and complexity

The second validation dark spot concerns the mapping between the activity, its consequences for the environment, the specific context of circumstances and the relevant algorithms in the forms. As it turns out, the more complex the case is, the more the relevancy of contextual information and contextual factual data has to be taken into account. This sounds logical, but it puts a strain on the ability of pre-designed forms and reference regulation models to capture the relevant contextual information in the appropriate manner. The desired separation between a platform that facilitates the exchange of case related information by using information capturing forms and the street level bureaucrat handling the case and making decisions based on the fair balancing of the desired environmental values gets blurred. This is because it becomes increasingly hard to separate the context from the balancing act. As a consequence, the design of the facilitating platform pre-includes more elements of decisions and balancing at system level if it is to handle more complex cases, thereby shifting power towards the system bureaucrat. This

finding is confirmed by Smit et al., (2018a; 2018) where he described the elicitation process in BRM research.

4.4   Validation of standardization

The third validation dark spot as mentioned by the respondents concerns the influence of standardisation. There is a structural tension between governments trying to maintain open standards and commercial suppliers that require a timespan for profit. Larger organizations have a legacy of already installed systems and dislike quick versioning of standards. Innovative, and often smaller, organizations tend to adapt new sophisticated standards faster. The design of software that is capable of handling geospatial and textual objects around spatial planning is relatively new and the standards that can handle the level of detail are rapidly evolving and changing. This is not unusual as a pattern in digitalisation, but it is relatively new to the domain of business rules and norms in legal planning systems. We observed a heated debate around the semantic standard for activities and announcements in the environment (STAM) (Interprovinciaal Overleg, 2020), the template for official governmental publications (STOP) (Interprovinciaal Overleg, 2020), and template for environmental planning documents (TPOD) (Interprovinciaal Overleg, 2020) in The Netherlands. One of the issues concerned the notion of standardisation of the annotation field related to a permit for an activity regarding an object. Some parties in the debate defended the position that the annotation of the decision ground should always be retrievable afterwards, thereby requiring further standardisation, while others argued that the annotation is exactly the level of freedom and decision power that should be left with the street level bureaucrat without any restrictions due to standardisation. Another interesting part of the standardisation concerns the business rules themselves, standardized in the standard for business rules applications in environmental planning (STTR) (Interprovinciaal Overleg, 2020). Some respondents complained that the intended reduction of environmental regulations was in fact replaced by an increase of business rules, which are used to determine the exact nature of the case and the exact values and norms that would apply on that case. Unfortunately, the design of the business rules as part of the algorithm is often done later in time, after the city has written its environmental policy for that election period. The design of these business rules in sequence after the national laws and regional or city policies are created often creates a translation problem. The technical modellers are often confronted with ambiguities and

semantic problems and require assistance from the lawyers and domain specialists to avoid faulty interpretations. The technical designers would strongly advise to include their modelling expertise earlier at the design of the policies, especially when these are written in texts. All respondents confirmed that the standardisation discussion was very technical and very dynamic and required effort to follow, especially for non-experts.

4.5 Understanding Between lawyers and programmers

The fourth validation dark spot identified by the respondents concerns the difference in reference frameworks between the text-oriented legal experts and the object-oriented spatial GIS experts. Earlier research in business rules and AI & Law hinted at misunderstandings between the culture of law and the culture of ICT (Boer, van Engers, Peters, & Winkels, 2007). This notion could have effects on the validation process and power shifts. Two contributions in the body of knowledge came to the same conclusion about this gap from completely different perspectives, see: "*In addition to challenge 2B, the current value of the ability to validate the cohesion between business decisions and business logic by legal subject-matter experts is low* (Smit et al., 2018) and the work of (Boer et al., 2007). The problem was again confirmed in this research by both the key suppliers of BRM software for the Omgevingswet platform and the national programme manager of the entire platform, for example, one of the interviewees stated: "*the legal guys do not understand the notion of an object infrastructure*". It was also observed that, for example, legal experts tend to judge each case in its own right, whereas IT specialist are trained to think and act in the paradigm of platforms, object classes, attribute values of those classes and exceptions. The respondents were very concerned for the future viability of the Omgevingswet platform because of the separation of the two 'tribes' (legal versus IT-oriented professionals) in the National programme. Legal professionals come from the old text-based legislation publication process and the notion of the GIS platform version control for all legal values is very hard to explain to them. Every object has attribute values for that particular day and the next day they can and will change, much unlike a published law in the early days. This, potentially dangerous, validation gap was also confirmed by the National Council for the Digital agenda of provinces (Interprovinciaal Overleg, 2020). Based on this, a programme with the goal to establish an open public registry and more knowledge sharing about business rules was initiated as a result.

4.6   'Rich' policy or policy contours only

The fifth validation dark spot concerns the level of detail of policy making. Some interviewees argued for environmental planning in a very broad sense, setting the contours for norms and values, but leaving much space open for local and use-case based discussions. The idea is that there should be more room for manoeuvring for flexible solutions, e.g., such as windmill not in my backyard-situations and problems. Others argue for a 'rich' environmental planning, with more strict and more detailed norms to protect scarce green space and maintaining an absolute minimum biodiversity. The provincial environmental policy act is therefore different for each of the 12 Dutch provinces and reflects the political debate in each region. The outcome has a differentiating effect for the shifts in power and the need for validation in each province. The richer and stricter the act, the more mapping effort and the more validation effort is required.

4.7   Technical platform neutrality

The sixth validation dark spot concerns the neutrality of the (technical) platform in being only a carrier of environmental policy decisions versus the extent to which elements out of the domain policies such as norms and values are mixed with the standardisation of the business rules in the knowledge acquisition forms. As explained earlier, the translation of legal texts into digital algorithms is not yet without problems. Some interviewees make a clear distinction between templates of fill-in forms, model-regulations, and 'clean' business rules. Others do not. The arguments against mixing these levels of preparation or automation are transparency and scalability and the autonomy of cities to carry out their own policies. Nationally provided templates and reference model regulations should help civil servants in cities to run their own implementation of their own policies. However, the domain knowledge with environmental norms and values and the business rules tend to get mixed up.

4.8   The issue of knowledge resources

The seventh validation dark spot concerns the scarcity of knowledge in combination with that of human resources. We apply cameras to reduce the amount of police officers required to measure compliance to speed limits, but the business rules applied there are relatively simple. In spatial planning there are many variables and

R. Peters, K. Smit & J. Versendaal:
Responsible AI and Power: Investigating the System Level Bureaucrat in the Legal Planning Process

623

issues, and it requires both domain knowledge and ICT knowledge when legislation is mapped into systems, these competencies are also required when the systems are being used operationally. Digitalisation is often introduced to increase efficiency, and the introduction of information systems always includes the influence of system designers into the process. But the effect in legislative and enforcement processes is yet little understood. Bovens et al., (2018) states that even simple design decisions, such as inspection method and inspection timing and placing could already be defined as a form of discretionary power. The interviewees, from the supplier side, explicitly stated that Dutch cities often hired their skills, because they simply could not afford to have this level expertise in-house.

4.9 The outsourcing of knowledge

The eight validation dark spot concerns an issue stated by the interviewees, that is related to the knowledge scarcity as described earlier. It is about the fact that, for the last twenty years, the cities in the Netherlands have increasingly outsourced their knowledge in spatial and environmental planning to commercial organizations. This is a known problem and questions are raised by city councils about the democratic control already. The introduction of digital algorithms in the spatial planning process is increasing this anxiety of the actors in the field.

4.10 Timing and dynamics of release management

The ninth validation dark spot concerns the timing of legal 'releases' of the Omgevingswet platform. Interviewees have argued that many components are still in the early development stage and should not be released to the public yet. Many external reports have declared the platform as too complex and argued for downscaling ambition and complexity. The National government is accused of clumsiness and fragmentation and overambition. Others argue that one must start somewhere, and that this platform is just the first of many such operations in eGovernment. It is unclear what influence this dynamic release debate has on the power and validation discussion, but it should be mentioned as an important 'background noise' affecting all variables in some way.

5   Conclusion

The overall conclusion of the investigation has been that there is a noticeable shift of power in view of the interviewees. But, in this arena of spatial planning at least, it is less clearly cut and more complex than could be expected. There is a strong notion to maintain platform neutrality and maintaining autonomy at the level of the city planners, who are the main street level bureaucrats in our case. The shift of power is often indeed caused by unintended side-effects, such as ambiguity of legal texts, scarcity of knowledge, mapping problems of contextual data and case knowledge acquisition/elicitation. The issue of standardisation between the rule of law and the business rules behind the balancing algorithms is of special research interest since it seems that it is here where system level decision power resides the most. Interviewees acknowledged that it takes great effort for non-experts to maintain an overview of the consequences of the design decisions made about the service platform at that level of abstraction. The power shift seems to increase if the automation is applied on cases that are more complex. Further research is supported by the National council of provinces, who acknowledge the gap between the culture of text orientation and the culture of object orientation and the risk for ethical values and norms regarding digitalisation. We aim for a validation framework that is understandable by policymakers and regional politicians. To do so, bridging BRM research and eGovernment research seems productive.

6   Discussion and Future Research Directions

Based on this study and the resulting conclusions we can identify points for discussion and point out future research directions. The sample size of 22 interviewees representing different groups of stakeholders is rather small, therefore we argue that future research should incorporate a larger sample size as well as research methods to do so in order to be able to generalize the findings from this study towards other similar digitalization projects. Overall, future research should provide more factual and objective means of measuring the influence of digitalization of intelligence in administrative eGovernment processes because we see this is a growing practice. The Dutch board of regional councilors have agreed on the notion that the gap between the culture of Law and the culture of object-based platform design should be bridged. This notion would justify future research on similarities between legislation and business rules. In addition, the validation capabilities and the dark spots should provide criteria for a more objective validation

R. Peters, K. Smit & J. Versendaal:
Responsible AI and Power: Investigating the System Level Bureaucrat in the Legal Planning Process

625

framework that should be in place at the start of the implementation of the Omgevingswet platform. It is interesting to observe that the digital ethics and algorithms discussion is now evolving from a GIS platform validation problem towards a Digital twin validation. These parallels are opening new grounds for investigation in future studies. Lastly, another question is what will happen to the client or citizen of these services when they are faced with the computer as substitute for a desk. We have investigated the 'supplier side', but this research did not involve the effect on the client side as well, which should be taken into account in future research.

Acknowledgements

References

Algemene Rekenkamer. (2021). Aandacht voor algoritmes. Den Haag. Retrieved from https://www.rekenkamer.nl/binaries/rekenkamer/documenten/rapporten/2021/01/26/aandacht-voor-algoritmes/Aandacht+voor+algoritmes.pdf

Bajec, M., & Krisper, M. (2005). A methodology and tool support for managing business rules in organisations. Information Systems, 30(6), 423–443. https://doi.org/10.1016/j.is.2004.05.003

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

Bovens, M. A. P., Zouridis, S., & Van Eck, M. (2018). Algoritmische rechtstoepassing in de democratische rechtsstaat. Nederlands Juristenblad, 93(40), 3008–3017.

Bovens, M., & Zouridis, S. (2002). From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. Public Administration Review, 62(2), 174–184. https://doi.org/10.1111/0033-3352.00168

Corea, C. (2021). Handling Inconsistency in Business Rule Bases. Universitat Koblenz-Landau.

Digitaleoverheid. (2020). NL DIGIbeter 2020: Agenda Digitale Overheid. Retrieved from https://www.digitaleoverheid.nl/wp-content/uploads/sites/8/2020/07/nl-digibeter-2020.pdf

European Commission. (2019). Ethics guidelines for trustworthy AI. Brussels. Retrieved from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

European Commission. (2020). On Artificial Intelligence - A European approach to excellence and trust. Brussels. Retrieved from https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

gemeente Eindhoven. (2019). Digitalisering van de stad. Eindhoven. Retrieved from
    https://eindhoven.parlaeus.nl/user/bdocument/action=showannex/gdb=687/Bijlage_1_-
    _Nota_Digitalisering_van_de_stad.pdf

Goodyer, I. (1991). Life events, development and childhood psychopathology. Chichester: Wiley.

Interprovinciaal Overleg. (2020). Verslag 6e Bestuurlijke Kopgroep Digitalisering.

Janssen, M., Hartog, M., Matheus, R., Yi Ding, A., & Kuk, G. (2020). Will Algorithms Blind People?
    The Effect of Explainable AI and Decision-Makers' Experience on AI-supported Decision-
    Making in Government. Social Science Computer Review, 089443932098011.
    https://doi.org/10.1177/0894439320980118

Koninkrijksrelaties, M. van B. Z. en. (2018). Waarom een stelselherziening? Retrieved March 25, 2021,
    from https://www.omgevingswetportaal.nl/wet-en-regelgeving/waarom-een
    stelselherziening

Kool, L., Timmer, J., Royakkers, L., & Est, R. (2017). Opwaarderen: Borgen van publieke waarden in
    de digitale samenleving. Den Haag. Retrieved from
    https://www.rathenau.nl/sites/default/files/2018-02/Opwaarderen_FINAL.pdf

Lipsky, M. (1980). Street-level bureaucracy: dilemmas of the individual in public services. New York:
    Russell Sage Foundation.

Pervan, G., & Maimbo, M. (2005). Designing a case study protocol for application in {IS} research. In
    Proceedings of the {Ninth} {Pacific} {Asia} {Conference} on {Information} {Systems} (pp.
    1281–1292). PACIS.

Peters, R. M. (2016). The Law, the Map and the Citizen: Designing a legal service infrastructure where
    rules make sense again. Universiteit van Amsterdam.

Smit, K., Versendaal, J., & Zoet, M. (2017). Identifying Challenges in BRM Implementations Regarding
    the Verification and Validation Capabilities at Governmental Institutions. In PACIS 2017
    Proceedings. Langkawi, Malaysia: Association for Information Systems (AISeL).

Smit, K., & Zoet, M. (2018a). An Organizational Capability and Resource Based Perspective on
    Business Rules Management. In Proceedings of the International Conference On Information
    Systems (ICIS) (pp. 1–17). San Fransisco: AIS Electronic Library (AISeL).

Smit, K., & Zoet, M. (2018b). Identifying Challenges in Business Rules Management Implementations
    Regarding the Governance Capability at Governmental Institutions. In Proceedings of the 51st
    Hawaii International Conference on System Sciences (pp. 2383–2392). Waikoloa (HI), United
    States: IEEE.

Smit, K., Zoet, M., & Berkhout, M. (2017). Verification capabilities for business rules management in
    the Dutch governmental context. In International Conference on Research and Innovation in
    Information Systems, ICRIIS. https://doi.org/10.1109/ICRIIS.2017.8002499

Smit, K., Zoet, M., & Versendaal, J. (2018). Challenges in the Implementation of the Elicitation, Design
    and Specification BRM capabilities. Journal of Information Technology Theory and
    Application (JITTA), 18(4).

Teuben, K. (2004). Rechtersregelingen in het burgerlijk (proces)recht. Alphen aan de Rijn: Wolters
    Kluwer.

van den Hoven, J., Miller, S., & Pogge, T. (2017). Designing in Ethics. (J. van den Hoven, S. Miller, &
    T. Pogge, Eds.). Cambridge University Press. https://doi.org/10.1017/9780511844317

van Engers, T. M., & de Vries, D. M. (2019). Governmental Transparency in the Era of Artificial
    Intelligence. In JURIX (pp. 33–42).

R. Peters, K. Smit & J. Versendaal:
Responsible AI and Power: Investigating the System Level Bureaucrat in the Legal Planning Process

627

Vereniging van Nederlandse Gemeenten. (2019). Principes voor de digitale samenleving. Den Haag. Retrieved from https://vng.nl/sites/default/files/2019-11/09a-bijlage-principes-voor-de digitale-samenleving.pdf

Vetzo, M., Gerards, J., & Nehmelman, R. (2018). Algoritms en grondrechten. Den Haag: Boom Juridisch. Retrieved from https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2018/03/01/a lgoritmes-en-grondrechten/Algoritmes+en+grondrechten+opgemaakte+versie.pdf

Zoet, M., & Versendaal, J. (2013). Business Rules Management Solutions Problem Space: Situational Factors. In Proceedings of the 2013 Pacific Asia Conference on Information Systems (paper 247). Jeju: AIS.

Zouridis, S., Van Eck, M., & Bovens, M. (2020). Automated discretion. In Discretion and the quest for controlled freedom (pp. 313–329). Cham: Palgrave Macmillan.

Adapting to the Enforced Remote Work in the Covid 19 Pandemic

Liana Razmerita,[1] Armin Peroznejad,[1]
Niki Panteli[2] & Dan Kärreman[1,2]

[1] Copenhagen Business School, Denmark; e-mail: lra.msc@cbs.dk,
Armin.Peroznejad@gmail.com, dk.msc@cbs.dk
[2] Royal Holloway, United Kingdom; e-mail: Niki.Panteli@rhul.ac.uk

Abstract Remote work provides an alternative method of working for organisations, which in turn became a norm during the Covid-19 pandemic. In this, paper, we study adaptation practices introduced by both individuals and organisations as a way for managing the enforced remote work. The study draws upon 33 interviews collected over a year during different phases of the Covid 19 pandemic. We apply adaptation theory lenses to examine the adaptation process over time and new digital working practices. In our study, we extend technological practices by including organisational and behavioral practices. We approach adaptation as a way of coping with a radical change or dynamic situation and building resilience. Based on the data analysis, we expend the adaptation theory in relation to different forms of adaptation to new remote work practices (e.g. at technological, organisational, and behavioural level).

1   Introduction

Remote work has long been recognised as alternative work arrangements for organisations (Staples, Hulland & Higins, 1999), which in turn became a norm during the Covid-19 pandemic. National and regional lockdowns and travel restrictions caused by the spread of Covid-19 have triggered an astonishing and forced transition into remote work among knowledge workers and professionals, which are likely to secure an avenue for the future of work and lead to a permanent transformation of the workplace. Organisations which had no digital capabilities, have struggled to adapt to the dire set of locked down circumstances. As a result of the pandemic, remote work has become essential among both employees and organisations (Waizenegger et al., 2020). Organisations regardless of size and sector have transitioned their operations to allow remote working, also referred as "locked-down digital work" (Richter, 2020), meaning employees of different professional backgrounds are now working exclusively through digital technologies (Leonardi, 2020). Even before the pandemic, part of their digital transformation, organizations have deployed social-collaborative platforms (e.g. enterprise social media) to streamline knowledge processes (in particular communication and collaboration) and support more efficiencient ways of working. Beyond efficiency such platforms were also deployed to support more transparent forms of knowledge work, innovation, retention of employees and their knowledge (Kirchner & Razmerita, 2019). Such social platforms facilitate formal and informal communication, and the articulation of personal into collective knowledge in a synergistic approach (Razmerita et al., 2014).

Remote work has been defined as a flexible work arrangement where workers have no personal contact with coworkers, but they are able to communicate using technology (Wang, Liu, Qian, & Parker, 2020). Remote work doesn't always have to specifically be at home. Concurrently, there is an essential need to analyse and understand the effects of the Covid-19 pandemic. In particular, in this paper, we study adaptation practices introduced by both individuals and organisations as a way for managing the enforced remote work and the perceived effects (e.g. on their productivity, work-life balance). Further, in this paper we seek to assess the Covid-19 pandemic effects on remote working attitudes according to different professional backgrounds and possible implications for the changing nature of work. We aim to

L. Razmerita, A. Peroznejad, N. Pantelli & D. Kärreman:
Adapting to the Enforced Remote Work in the Covid 19 Pandemic

631

address the following research question: How individual knowledge workers and organisations are adapting to enforced remote work?

Our study findings suggest that on one hand such transition advanced the digitalisation transformation and the associated individual digital work practices, on the other hand it shapes not just the workspace landscape but also the role and meaning of the organisation as an entity. In what follows, we present a brief review of the extant literature on telework and virtual work pre-Covid-19, and then we compare this to the Covid-19 related remote work arrangement. We then introduce the concept of adaptability as the theoretical lens of the study before we present the methodological approach and findings of our study.

2 Remote Work

2.1 Remote Work in the Pre-Covid-19 work Context

Researchers have been studying remote work, including home-work, telework and virtual work arrangements for a period of more than 20 years (among the more recent ones Makarius and Larson, 2017; Raghuram et al. 2019). Literature on remote work has focused on the opportunities that this form of work provides for flexible working and achieving work-life balance (Felstead, 2002). A distinct feature is that this form of work has been presented as a choice driven by the organisation in its efforts to reduce overheads associated with office facilities, or a choice driven by individuals due to their preference for flexible working. Furthermore, it allowed organisations to hire across borders, to attract talents that are difficult to find locally or to offer family-friendly employment contracts to those who do not want to relocate. For this, remote work has been defined as work that is 'technologically feasible, flexible and autonomous, desirable and perhaps even inevitable, family- and community-friendly' (Bryant, 2000: 22).

Remote work or teleworking is most suitable and beneficial for employees who mainly perform knowledge-based tasks with limited face to face contact (Rupietta and Beckmann, 2017). A study on 273 knowledge workers from different professional backgrounds including: engineers, accountants, sales & marketing found that they perform better in virtual work, so long as their role doesn't require social collaboration (Golden and Gajendran, 2018). On one hand working from

home increases employee autonomy, on the other hand it lowers the firm's ability to monitor their employees. This all too familiar, all-encompassing scenario describes a typical principal-agent dilemma, wherein employees have incentive to abuse the power afforded to them through their autonomy, by reducing their individual work effort thus resulting in shirking behaviour (Rupietta and Beckmann, 2017).

Makarius and Larson (2017) suggest four key competences which lead to successful remote working and performance outcomes which improve effectiveness and satisfaction such as: establishing behavioural guidelines, developing trust, coordinating information, using media. Setting boundaries through behavioural guidelines is almost akin to a code of conduct policy style which sets boundaries for employees in virtual work. Coordinating information in order to have at disposal to be able to accomplish tasks to best possible and most accurate standard. Finally using media for communication putting it to best use such as via phone, email or now with the many conferencing applications currently at our disposal (Makarius and Larson, 2017).

2.2   Remote Work in the Covid-19 context

Regardless of the aforementioned literature, existing studies do not provide a thorough investigation of the enforced work from home as emerged due to Covid-19 locked-down (Waizenegger et al, 2020, Richter, 2020). Indeed, there are some distinct differences between the pre-Covid-19 and the Covid-19 remote work arrangement noted to the fact that although remote work is not new (e.g. Sayah, 2013), the Covid-19 context is. As a result of the pandemic, the switch to remote work took place suddenly with many organisations, traditionally collocated, being ill-prepared under the circumstances. Henceforth, many employees felt increasing pressure to 'make digital work, work' (Richter, 2020) which have led to a change of mindset but also to misconceptions or potential negative associations. Among these are: 1. digital work is more stressful and tiring, 2. adaptation and appropriation to digital work technology, 3. online (digital conferences) versus physical meetings, 4. autonomy and visibility.

L. Razmerita, A. Peroznejad,N. Pantelli & D. Kärreman:
Adapting to the Enforced Remote Work in the Covid 19 Pandemic

633

Recent study on remote work identified the main benefits and challenges for remote work during Covid 19. On one hand among the identified benefits are: it frees knowledge workers from office distraction and commuting, it helps concentration on individual tasks, and it allows knowledge workers to have flexibility in their schedule due to increased work autonomy. On the other hand, the overall use and wide range of available ICT it creates an "always on" culture. Furthermore individuals working remotely may feel alienated, isolated or worried (Waizenegger et al., 2020). Further, remote work has brought some additional stress and impact on well-being of employees bringing imbalances to work and home dichotomy especially for women (Amis and Greenwood, 2020). Another study by Dubey and Tripathi's (2020) aimed at analysing the sentiments and emotions of workers towards working from home during the Covid-19 pandemic. Their study revealed that more than 73% of people had positive sentiments towards working from home whilst 27% people had a negative perception towards working from home experience. Furthermore, over 60% of the people responded with emotions of trust, anticipation and joy for work-from-home culture while a handful replied with fear, sadness, anger and disgust. Their obtained results show that experiences of homeworking had a positive perception, globally.

3 Theoretical Lens of Adaptation

Adaptation is a concept that has been used in organisational and management studies in different ways to signify a response to a change situation. Within organisation studies, adaptation has been seen in terms of fidelity and extensiveness. The former implies practices that may be modified through localisation or re-invention (Ansari 2010). Extensiveness takes account of the degree of change and transformation of a given practice and this may vary from symbolic to substantive. Adaptive behaviour has been identified as a prerequisite for effective job performance and career development (Griffin & Hesketh, 2003; Cullen, et al, 2014). This body of literature has focused on individuals and their ability to adapt in changing situations. Park and Park (2019) have studied the antecedents of individual adaptive behaviour and found factors ranging from individual personality, skills and motivation, to job characteristics including the degree of autonomy and task inter-dependency to group and organisational characteristics entailing the support gained and learning environment, Further, work by (Makarius & Larson, 2017) focused on successful individual adaptation to virtual work. The study emphasized the need

to focus on individual cognition in virtual work and they propose the concept of virtual intelligence associated with behaviours and skills for the individual to adapt successfully to virtual work.

Studies have also examined adaptation practices both across organisations and within organisations. Ansari et al (2014) studied how a multinational organisation has managed the adaptation processes within the corporate and subsidiary levels. They argued that adaptation is an engineered practice that aims to vary from one location to another and therefore accommodating diversity in the local settings, increasing acceptance in this way. The literature also makes reference to team adaptation with its significance being linked to team performance (Woolley, 2009). Klein and Pierce (2001) defined adaptive teams as *"teams that are able to make the necessary modifications in order to meet new challenges"* (p. 4). The authors recognised that adaptation is a complex process, dependent on a number of dimensions which are specific to the team at hand. Further, adaptation theory has been used to examine adaptation practices within virtual teams (Thomas and Bostrom, 2010 a,b), notably switching, expanding, merging, modifying and creating new practices during the adaptation process. In a different study, the authors made reference to technology adaptation as a process that "involves the acquisition and usage of new ICTs or new features, and the modified usage of existing features in ICTs" (Thomas and Bostrom, 2008, p. 47). In their study on technology adaptation, Tyre and Orlikowski (1994) made reference to the temporal dimension of the adaptation process arguing that the process is not always linear nor incremental and continuous, but rather highly discontinuous with instances of bursts and stops. In this study, we extend adaptation theory applied previously in studies on virtual work by including organisational and behavioural practices in addition to technological practices. Adaptation within our study is defined as the acquisition of new practices, at both individual and organisational level, as a way for building resilience and coping with a changing and dynamic situation caused by the pandemic. Successful adaptation is a way of coping with the changing and dynamic situation over time.

L. Razmerita, A. Peroznejad,N. Pantelli & D. Kärreman:
Adapting to the Enforced Remote Work in the Covid 19 Pandemic

635

4 Research design

The study has been designed as a qualitative, longitudinal study in order to explore the remote work experiences and practices during the different pandemic phases. It applies an abductive interpretivist approach that help theorizing the remote work phenomenon (Alvesson and Kärreman 2007; Van Maanen, Sørensen, and Mitchell 2007).

The study is based on a series of semi-structured interviews with white collar professionals distributed over three phases across different sectors and organisations in hard-hit countries with severe lockdown such as the UK and not so hard-hit countries with less severe lockdowns (Sweden, Germany). In total 33 interviews were collected over a three-phase period. The first phase (P1) included 10 interviews and took place not long after the start of the intial lockdown (March-April 2020); the second phase (P2) included 14 interviews that took place between December 2020 to February 2021. In the third phase, which took place at the beginning of March 2021, we have re-interviewed 9 of the 10 interviewees from the first phase, in order to understand how their work from home practices and their attitudes have changed. Further insights from the authors' experiences have been integrated. The latter phases of interviews gave us the opportunity to expand on interviewees' remote work experiences following a period of adaptation for both individual employees and their employer organisations. In particular, we have expanded on aspects related to experience, learnings and adaptation to new remote work practices.

The interview guideline consisted of questions on the experience with remote work, how they have been managed, expectations, support in organising remote work but also preferences related to remote work, experience with work from home (WFH) prior to the pandemic, motivations and limitations. The interviews were conducted via telephone or video call, due to the pandemic circumstances, as well as being an international sample of interviewees. The interviews were structured in a 30-to-60-minute interview arranged with each participant. Interviews were recorded and the audio files transcribed. Our analytical approach was based on thematic analysis and was based on the core dimensions of adaptation theory in the enforced remote work context, notably technological, organisational and behavioural adaptation.

5   Findings

Most of the respondents had no prior experience of (paid) working from home, as part of their current or previous jobs up to this point. All seemed to be coping well with most the stresses coming from the situational side of the Covid-19 pandemic and lockdown regulations rather than the actual working from home. They generally agreed that the benefits outweigh the negatives and as of now their bigger concerns were regarding the effects of the Covid-19 pandemic on their respective businesses. Not surprisingly all interviewees shared the opinion that there was a general lack of preparation from themselves and their respective employers in the first phase (P1). Several of the interviewees noted that a strong sense of organisational culture and camaraderie was lacking during the abrupt transition to enforced remote work. The authors engaged in several rounds of data analysis and discussion of coding in relation with remote work experiences in a transparent, iterative way. We tried to identify more surprising or unexpected findings in relation with extant literature, during the second and third phase of data collection that could expand both remote work and the adaptation theory. We have agreed to focus on the digital knowledge work adaptation practices at different phases of the crisis.

Findings show that adaptation practices have been introduced by both individual employees and their respective organisations. Interviews emphasized the importance of self-management, setting goals and milestones "*I'm finding motivation by setting regular short-term goals for myself, such as weekly milestones, communicating frequently with colleagues and reflecting on how successful we have been in these challenging times as a collective outfit.*"(P2) Routine was found to ultimately be beneficial from P2 onwards: "*I've just been splitting up my day so that I wake up and have my breakfast as usual, then go through my emails and then I'll just work on through however long it takes me to get through the work once I've done it then I'll do lighter duties.*" Finding in turn that "*this has helped me take a much more structured approach to my life which I never really had before the pandemic.*"

Resilience is another element that has been discovered during the process "*For me personally, it has shown how resilient and flexible I can be in terms of learning new things quickly, so I've learned a lot a lot of new skills in a short time span.*"(P2) Interviewees also spoke candidly of the importance of the digital technologies deployed in P2 "*getting used to interface technology such as zoom has been a big part of the process and getting used to the lack of normal human interaction has been challenging.*" Those with young children, particurlarly

L. Razmerita, A. Peroznejad,N. Pantelli & D. Kärreman:
Adapting to the Enforced Remote Work in the Covid 19 Pandemic

637

mothers, seemed to have extra hurdles "*adapting to work from home where I'm also a mother has been difficult to juggle at times but now I just plan to do most of my work during the hours when the kids are in school.*"(P2) which also reflects on the intrusive nature of work in the family/home setting.

From the perspectives gained, at the individual level, motivation in the 1st phase simply came from job security "*meaning that if the work is not done then I lose the job, which is only fair and is to be expected as the same rule applies to any normal form of working.*" which can become perpetuate by: "*the comfort of home, as it makes me feel more productive as the setting and planning of the day is in my control so I can arrange it to fit me best*."(P1) Yet noticeably, distractions didn't appear as anything too dissimilar from the traditional office "*I haven't got too many distractions at home and feel they sort of even themselves out with the office anyway because you may have electronic devices and so on at home but it's a lot quieter and more peaceful as opposed to a buzzing and noisy office which can be very distracting in my line of work.*"(P2).

Individual adaptation does include behavioural adaptations, but also work environment adaptation such as: the adaptation of private space and general lifestyle adaptation. One respondent would switch on the TV to simulate the background noise present in any busy office. "*Whilst the TV can be a distraction for some, I actually enjoy having it on in the background, mostly just for some background noise which I feel replicates the background noise of an office somewhat and this makes me more productive because I prefer not to work in silence.*"(P1).

With it "*mainly just having been sort of limiting myself and being very strict*"(P2) as remote workers became familiar with "*out of sight, out of mind*"(P1) mentalities. Many identified; a lack of trust, feelings of being an outsider and lack of social support. Some respondents had "*a tight nit group*"(P2) so had been spurring each other on and helping each other out as much as possible. This camaraderie seemed to really help, providing harmony and unity thus increasing the overall performance of the team.

Whilst at the organisation level, such adaptation practices include providing the necessary education and continued support to help members adapt to remote work. The organisation "*understands the importance of the pandemic and is investing for the future and not just looking for a quick fix*" by "*investing £100,000 in a new telecoms system during the pandemic, to make life easier for the staff working from home*"(P2) such organisational

adaptation takes place through various layers which include management support, if assistance is required, as well as elements that pair with technology adaptation. For example, contact lines were made accessible for urgent assistance or adoption of new collaborative technology to support communication: "*The contact lines are open, and we have been made aware of this, so we can contact superiors or colleagues at any time if assistance is required. We have also created WhatsApp groups between different departments in order to communicate easily with other colleagues, this has been a great addition both for technical and morale support and is an addition I imagine will remain once we return to normal.*"(P1) However, employees were missing "*the social interaction of having people around and working directly with colleagues… part of the organisation has been doing monitoring just as a precaution to measure if they want more funding for various departments to take home more members of staff as our business books continue to grow*"(P2).

The second and third round of interviews were much more favourable to partial remote working schedules in the future compared to varying degrees of preference in the first round with less scepticism around it as exampled here "*I'm very much of the mindset that whatever I do from here on out, it must be majority remote. I have become very accustomed to it.*"(P3). Overall, out of the 33 interviews, all the respondents had successfully adapted to digital technology, despite some inital difficulties. What the data indicates is that regardless of preferability, sentiments and attitudes towards remote work have not impacted knowledge workers' percieved efficiency. A great majority of respondents (18 out of 24) have actually had a positive association towards remote work despite the certain inherent difficulties to adapt to this new way of working. Although it is worthwhile noting that the negative sentiments were more pronounced, sentiments such as "*I'm still mostly against it. I feel it's not good for me or my mental health.. Collaborating with people is the big one. Working from home you can't work on problems together and learn from each other. If I have a question, in the office there is always someone around that can answer, discuss, and look at the computer together to work it out together which I find better for learning.*"(P3) particurlarly stand out, even if fewer instances of such negative sentiments occurred. Employees have to find digital ways to learn to replace traditional knowledge sharing and transfer which take place in a traditional organization. Those thriving the most, over time, had set schedules, stayed active, took responsibility and accountability for their time and learning, finances and health through daily planning and routine setting.

6   Discussion and Conclusions

Adaptation is important in any organisational change (Ansari, 2010). This study was driven by an interest to investigate the enforced adaptation to remote work and thus to contribute to research on knowledge work, changing nature of work, and new digital ways of organising through technology.

Remote working has led to greater employee autonomy over how, when & where they work and how knowledge work is organised. This has inadvertently further blurred the boundaries between personal and working lives. Adaption has proved to be a critical factor in determining the success of enforced remote work for both individuals & organisations. Individual adaptation refers to new work practices (communication & collaboration via digital technologies), new routines, spatial adaptation, life-style adaptation, new-skill development & self-management. This study's findings present different levels of adaptation that took place during the Covid-19 pandemic: the first order adaptation takes place as a way to cope with the sudden changes due to the enforced lockdowns (e.g. learning new technologies, new collaboration tools) and the second order adaptation that involves more profound transformations at both individual and organisational levels. We identify a variety of factors that are intrinsic to individual experiences, those developed by individuals during the different phases of the Covid-19 pandemic, including work environment adaptation as well as management support and organisational adaptation.

The study indicates the feasibility and challenges of long-term remote work on a mass scale. Most of the respondents were able to succesfully adapt to the use of technology and have a positive sentiment towards remote work. This is in line with previous findings by Dubey and Tripathi (2020). However, the majority of respondents expressed a desire to go back to the office st some stage and find a hybrid mix between work from home and office work in the future. Remote work undoubtedly has benefits and challenges for both employer and employee; certain factors were found to moderate the effectiveness of remote work. These variables are twofold: at the individual level, an ability to work autonomously as well as an ability to self-manage through the adaptation processes. Increased autonomy led to employees being socially isolated, less connected with their colleagues and their organisations.

Whilst at the organisational level we find that the degree of digitalisation and type of work practices prior to the crisis could predict remote working success. Several barriers also emerged in accordance with remote work. On the individual level we found that employees with lower levels of trust found remote work to be more challenging due to limited contact with colleagues & supervisors, lack of knowledge exchange opportunities and social isolation. For individuals who had experienced remote work prior to the Covid-19 crisis, adaptations came easier. Whislt at the organisational level we observed that many firms lacked the proper tools and training systems thus requiring more time for adequate adaptation. Organizations also initiate new forms of distant management which consist of regular formal or informal meetings, updates by email, prerecorded videocasts or podcasts. Findings have shown that adaptation happened over time and it was on a continuum basis. As organisations were trying to respond to the needs of their remote employees, whilst individuals were learning to cope with changes to new technologies and a new work setting where they often had to negotiate their work space in a shared household.

Previous studies have pointed to the affordances and challenges of team collaboration (e.g. Waizenegger et al, 2020), opinions on potential effects of digital work (Leonardi, 2020) and digital work during the first phase of locked-down (e.g., Richter, 2020), while our study has investigated work processes adaptation based on a longitudinal study, over three critical periods of the Covid 19 crisis.

In conclusion, we examine work adaptation practices in the Covid-19 work context over three time periods and gained better insights of this enforced type of remote work. Our study contributes to the knowledge work literature by applying and extending adaptation theory; we do so by identifying technological, orgainsational and behavioural dimensions of adaptation in the enforced remote work context. A limitation of our study is that this did not sufficiently examined the organisational view; this should be in the agenda of future research. Further research is also needed to understand how increased use of remote work will impact organisational learning practices, organisational policies and cultures, but also how this experience may impact individuals' employment choices.

L. Razmerita, A. Peroznejad, N. Pantelli & D. Kärreman
Adapting to the Enforced Remote Work in the Covid 19 Pandemic

641

References

Alvesson, M., & Kärreman, D. (2007). Constructing mystery: Empirical matters in theory development. Academy of Management Review, 32(4), 1265–1281. https://doi.org/10.5465/AMR.2007.26586822

Amis, J., & Greenwood, R. (2020). Organisational Change in a (Post-) Pandemic World: Rediscovering Interests and Values. Journal Of Management Studies, 58(2), 582-586. https://doi.org/10.1111/joms.12663

Ansari, S., Fiss, P. C., & Zajac, E. J. (2010). Made to fit: How practices vary as they diffuse. *Academy of Management Journal*, *31*, 67–92.

Ansari, S., Reinecke, J., & Spaan, A. (2014). How are practices made to vary? Managing practice adaptation in a multinational corporation. *Organisation Studies*, *35*(9), 1313-1341.

Bryant, S. (2000), 'At Home on the Electronic Frontier: Work, Gender and the Information Highway', New Technology, Work and Employment 15, 1, 19–33.

Cullen, K. L., Edwards, B. D., Casper, W. C., & Gue, K. R. (2014). Employees' adaptability and perceptions of change-related uncertainty: Implications for perceived organisational support, job satisfaction, and performance. *Journal of Business and Psychology*, *29*(2), 269-280.

Dubey, A. and Tripathi, S., 2020. Analysing the sentiments towards work-from-home experience during COVID-19 Pandemic. Journal of Innovation Management, 8(1).

Felstead, A., Jewson, N., Phizacklea, A., & Walters, S. (2002). Opportunities to work at home in the context of work-life balance. Human resource management journal, 12(1), 54-76.

Griffin, B., & Hesketh, B. (2003). Adaptable behaviours for successful work and career adjustment. *Australian Journal of psychology*, *55*(2), 65-73.

Golden, T. D. (2009). Applying technology to work: toward a better understanding of telework. *Organisation Management Journal*, *6*(4), 241–250. https://doi.org/10.1057/omj.2009.33

Kirchner, K., & Razmerita, L. (2019). Managing the Digital Knowledge Work with the Social Media Business Value Compass. In 52nd Hawaii International Conference on System Sciences (pp. 6438–6447).

Klein, G., and L. G. Pierce. 2001. "Adaptive Teams." Proceedings of the Sixth ICCRTS Collaboration in the Information Age Track 4: C2 Decision Making and Cognitive Analysis.

Leonardi, P. M. (2020). COVID-19 and the New Technologies of Organising: Digital Exhaust, Digital Footprints, and Artificial Intelligence in the Wake of Remote Work. Journal of Management Studies. https://doi.org/10.1111/joms.12648

Makarius, E. E., & Larson, B. Z. (2017). Changing the perspective of virtual work: Building virtual intelligence at the individual level. Academy of Management Perspectives, 3(2), 159–178. https://doi.org/10.5465/amp.2014.0120

Park, S., & Park, S. (2019). Employee adaptive performance and its antecedents: review and synthesis. Human Resource Development Review, 18(3), 294-324.

Raghuram, S., Hill, N. S., Gibbs, J. L., & Maruping, L. M. (2019). Virtual work: Bridging research clusters. Academy of Management Annals, 13(1), 308–341. https://doi.org/10.5465/annals.2017.0020

Razmerita, L., Kirchner, K., & Nabeth, T. (2014). Social Media in Organizations: Leveraging Personal and Collective Knowledge Processes. Journal of Organizational Computing and Electronic Commerce, 24(1), 74–93. https://doi.org/10.1080/10919392.2014.866504

Richter, A. (2020). Locked-down digital work. International Journal of Information Management, 55. https://doi.org/10.1016/j.ijinfomgt.2020.102157

Rupietta, K. and Beckmann, M., 2017. Working from Home. Schmalenbach Business Review, 70(1), pp.25-55.

Thomas, D., & Bostrom, R. (2008). Building Trust and Cooperation through Technology Adaptation in Virtual Teams: Empirical Field Evidence. Information Systems Management, 25(1), 45-56. https://doi.org/10.1080/10580530701777149

Thomas, D., & Bostrom, R. (2010). Vital Signs for Virtual Teams: An Empirically Developed Trigger Model for Technology Adaptation Interventions. MIS Quarterly, 34(1), 115.

https://doi.org/10.2307/20721417

Tyre, M. J., & Orlikowski, W. J. (1994). Windows of opportunity: Temporal patterns of technological adaptation in organisations. *Organisation science*, *5*(1), 98-118.

Sayah, S. (2013). Managing work-life boundaries with information and communication technologies. *New Technology, Work and Employment*, 28(3): 179–196.

Staples, D., Hulland, J., & Higgins, C. (1999). A Self-Efficacy Theory Explanation for the Management of Remote Workers in Virtual Organisations. Organisation Science, 10(6), 758-776. https://doi.org/10.1287/orsc.10.6.758

Van Maanen, J., Sørensen, J. B., & Mitchell, T. R. (2007). The interplay between theory and method. Academy of Management Review, 32(4), 1145–1154.
https://doi.org/10.5465/AMR.2007.26586080

Waizenegger L., McKenna B., Cai W. & Bendz T. (2020): An affordance perspective of team collaboration and enforced working from home during COVID-19, European Journal of Information Systems, DOI: 10.1080/0960085X.2020.1800417

Wang, B., Liu, Y., Qian, J., & Parker, S. (2020). Achieving Effective Remote Working During the COVID‐19 Pandemic: A Work Design Perspective. Applied Psychology, 70(1), 16-59. https://doi.org/10.1111/apps.12290

Woolley, A. (2009). Putting first things first: Outcome and process focus in knowledge work teams. Journal Of Organisational Behavior, 30(3), 427-452. https://doi.org/10.1002/job.578

Exploring the Sustainability of Swiss Online Shops: Preliminary Evidence from a Clustering Approach

Thomas Wozniak, Guang Lu, Dominik Georgi, Anja Janoschka & Antonia Steigerwald

Lucerne School of Business, Institute of Communication and Marketing, Lucerne, Switzerland; e-mail: thomas.wozniak@hslu.ch, guang.lu@hslu.ch, dominik.georgi@hslu.ch, anja.janoschka@hslu.ch, antonia.steigerwald@hslu.ch

Abstract This paper proposes a framework for assessing the sustainability of online retailers. By adhering to the logic of the value chain, the framework captures all steps from procurement to shipping of products and handling product returns that typically occur in e-commerce. The framework is then applied to explore the sustainability of 227 Swiss online shops through a clustering approach. Core elements of the framework are dummified and condensed to 20 features using an autoencoder-based neural network. K-means clustering is used to identify three sustainability clusters that are visualised in the latent space. One-way ANOVA and chi-square analyses of the three-cluster solution indicate that the identified clusters are distinct across most elements of the proposed framework. These preliminary results hold a variety of future research avenues.

1   Introduction

Global e-commerce sales have continually grown over the past years and account for an increasing share of total retail sales. In 2019, emarketer predicted that global e-commerce sales would grow by 19 percent and account for 16 percent of total retail sales in 2020 (Lipsman, 2019). The recent COVID-19 pandemic has further increased the shift from offline to online retailing, especially due to social distancing requirements and lockdowns put in place to prevent spreading of the corona virus. For individual countries, e-commerce sales were thus predicted to grow more than 20 percent in 2020 (Cramer-Flood, 2020). It is expected that the accelerated trend towards online retailing will remain during the recovery from the COVID-19 pandemic (UNCTAD, 2021a).

E-commerce has environmental consequences (Fichter, 2003). Growth in e-commerce sales translates to, e.g., more intense use of ICT and logistics affecting the use of energy, material resources, and land, and the emission of greenhouse gas (Fichter, 2003). In light of climate change (Steffen et al., 2018) and even continents aiming to become climate-neutral (European Commission, 2020), analyzing the sustainability of e-commerce becomes more relevant than ever before.

This paper develops a framework for assessing the sustainability of online retailers (section 2.3), based on the concept of the value chain specifically in e-commerce (section 2.1) and considering extant research on sustainability in e-commerce (section 2.2). The framework serves as basis for developing a survey instrument for collecting empirical data from 227 online retailers in Switzerland (section 3.1). Using an autoencoder-based neural network, the data is condensed to 20 features for subsequent clustering (section 3.2.1). K-means clustering identifies three clusters of online shops which are visualized in the latent space (sections 3.2.2 and 3.2.3). The statistical analyses of the clusters indicate that they are distinct across most elements of the proposed framework (section 4). These preliminary results provide fertile ground for further and future research (section 5). Overall, this paper contributes to a better understanding of sustainability in e-commerce specifically by proposing an assessment framework, applying it to a unique dataset, and identifying three clusters of online shops according to their sustainability as assessed by our framework.

T. Wozniak, G. Lu, D. Georgi, A. Janoschka & A. Steigerwald
Exploring the Sustainability of Swiss Online Shops: Preliminary Evidence from a Clustering Approach

645

2   Conceptual background and framework development

2.1   The e-commerce value chain

The concept of the value chain, developed by (Porter, 1985), is widely used by academia and practice alike. Following a linear logic, the value chain entails five subsequent "generic categories of primary activities […] involved in the physical creation and delivery of the product to the customers" (Ricciotti, 2020, p. 193), namely inbound logistics, operations, outbound logistics, marketing and sales, and service (Porter, 1985). These sequential primary activities are sustained by four support activities such as procurement, technology development, human resource management, and firm infrastructure (Porter, 1985; Ricciotti, 2020). Several scholars have adapted the generic value chain concept to more specific contexts (Ricciotti, 2020). Some have departed from its sequential and linear logic and embraced a network logic to better reflect that firms are part of stakeholder networks that compete with each other, resulting in the concept of the value network (Ricciotti, 2020). While such network logic may more closely represent today's competitive environment, the simplicity of the value chain has not lost its appeal.

In the (online) retailing value chain, the most important players are manufacturers of products, institutional retailers, and consumers (Reinartz, Wiegand, & Imschloss, 2019). Such retailers may be traditional stationary retailers who also sell their products via online channels as part of multichannel strategies or so-called pure plays who sell online only (Reinartz et al., 2019). Manufacturers have also started running their own online operations aiming to engage directly with end customers, thereby circumventing retailers (Reinartz et al., 2019). In this paper, we focus on the value chain from the perspective of the online seller of products. This can include any type of online retailer as well as manufacturers maintaining direct-to-consumer (D2C) online operations. For selling products online, retailers and manufacturers need to run a website with an online shop.[1] Products sold online are typically stored in a warehouse. Once ordered, they are packaged, and delivered to the customer. The customer receives the product, keeps it, or may decide to return it.

---

[1] Some online sellers may choose to sell via downloadable apps or social media platforms such as Facebook or Instagram only. However, it can be assumed that most online sellers run a web-based online shop to reach the largest number of potential customers.

2.2 Sustainability in e-commerce

Research on the sustainability in e-commerce has addressed a wide array of aspects. This includes comparing the environmental impact of brick-and-mortar retail channels with online channels (Bertram & Chi, 2018; Pålsson, Pettersson, & Winslott Hiselius, 2017; P. Van Loon, McKinnon, Deketele, & Dewaele, 2014; Weber et al., 2009), sustainability in e-commerce packaging (Escursell, Llorach, & Roncero, 2020), sustainability of e-commerce logistics (Mangiaracina, Marchet, Perotti, & Tumino, 2015), lifecycle analysis of different fulfilment channel types (Patricia Van Loon, Deketele, Dewaele, McKinnon, & Rutherford, 2015), e-commerce customers' last-mile delivery preferences based on the sustainability impact (Ignat & Chankov, 2020), and product returns' effect on business, society, and the environment (Frei, Jack, & Brown, 2020). While these contributions are diverse in the topics they address and the methods applied, they have one important denominator in common: they address or can be related to specific segments or phases of the e-commerce value chain. This underlines the suitability of the value chain concept as a frame of reference for analyzing sustainability in e-commerce.

2.3 A framework for assessing the sustainability of online retailers

Adhering to the logic of the value chain and considering extant literature, we identified six core elements of the e-commerce value chain that need to be considered when assessing the sustainability of online shops: 1) *products*, 2) *online shop operation*s, 3) *intralogistics and storage*, 4) *packaging*, 5) *delivery*, and 6) *product returns*. Some of these elements may involve the cooperation with other companies (e.g. suppliers or fulfilment/logistics service providers). We thus included 7) *cooperation with partners* as a further core element. In our framework (see Figure 1), this element is depicted in parallel to the six linear elements since it can relate to any of them.

To emphasize the goal of this framework, namely to guide assessing online retailers' sustainability, we nest the core framework within the society and the environment. The link between online retailers and the society is established through interactions with their stakeholders which are part of the society. Such stakeholders can include partners of the company (e.g. suppliers), employees, and customers, but also consumer and trade organizations, and entities such as states with their regulatory requirements. Communication plays a major role for such interactions (Lim &

T. Wozniak, G. Lu, D. Georgi, A. Janoschka & A. Steigerwald
Exploring the Sustainability of Swiss Online Shops: Preliminary Evidence from a Clustering Approach

647

Greenwood, 2017). We thus include *communication with stakeholders* as further element in our framework.
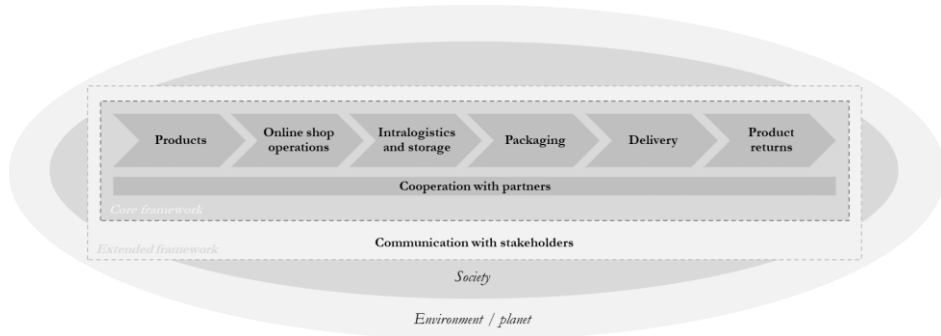

Figure 1: E-commerce sustainability assessment framework nested in society and environment

Nesting organizations and their processes and/or business models within external environments in the sense of context is not new. Prominent and well-known examples are the St. Gallen Management Model (Rüegg-Stürm & Grand, 2019) and the environment of the business model canvas (Osterwalder & Pigneur, 2010). Our approach is distinct from these examples in two ways: First, we purposely focus on society and environment, thereby simplifying the manifold and complex interactions between many types of stakeholders. Second, we nest our framework within society which is in turn nested within the environment. This is in line with that paradigm of sustainability which proposes the economy depends on the society and both depend on the environment, in that order (e.g. Griggs, 2013).

3   Methodology

3.1   Development of survey instrument, data collection, and sample

The framework developed in section 2.3 served as foundation for developing the survey instrument. For each of the framework's core elements, survey questions and items were designed to adequately address the specific nature of each core element. Further survey questions addressed communication and marketing, the assessment of customer needs, strategic priorities, and questions determining the size and nature

of the online shop (e.g. product categories and annual order volume). The final survey instrument was reviewed and pre-tested by experienced researchers, market research professionals, and e-commerce practitioners.

The online survey targeted online shops specifically operating in Switzerland and was administered from March to July 2020. Switzerland ranks first in the UNCTAD B2C E-commerce Index 2020, reflecting high scores in all four dimensions of the index, namely Internet penetration, banking coverage, secure server density, and postal reliability (UNCTAD, 2021b). An invitation to the online survey was sent to a large number of Swiss online shops. In addition, the link to the survey was distributed in social media and specific newsletters to reach out to the target group. The final sample consisted of n=227 completed online surveys.

In our sample, the most frequent product categories were fashion and food (27% each), houseware (21%), cosmetics and sporting goods (16% each), and toys (13%). In terms of business relationships, the majority stated to serve B2C customers (81%), which was the primary target group of the survey; 51% stated to serve B2B customers; 25% stated to follow a D2C model. Our sample included online shops with different annual order volumes: 32% reported less than 1,000 orders, 36% 1,000 to 9,999 orders, and 32% 10,000 and more orders in 2019.

3.2   Three-step clustering approach

3.2.1   Data preparation

From the overall survey instrument, we have selected a total of 10 questions to adequately cover all elements of our core framework, i.e. from products to returns, and cooperation with partners. The dummification of the relevant variables lead to a dataset containing 227 online shops and 150 features having only categorical values of "0" or "1".

An autoencoder-based neural network has been used to perform the dimension reduction on the dataset obtained from the previous step. This is beneficial to reducing the noises underlying the original data, which is helpful for the clustering. Compared to the conventional linear dimension reduction techniques using e.g. Principal Component Analysis (PCA), autoencoder-based neural networks are

T. Wozniak, G. Lu, D. Georgi, A. Janoschka & A. Steigerwald
Exploring the Sustainability of Swiss Online Shops: Preliminary Evidence from a Clustering Approach

649

claimed to be more powerful as they can capture nonlinear correlations between features and perform nonlinear transformations to the data (Alkhayrat, Aljnidi, & Aljoumaa, 2020).

Figure 2 displays the structure of the stacked autoencoder. To train this autoencoder, 20,000 epochs have been run, with a batch size 32, a validation split 0.1, and a shuffle variable being set as "true". Independent upon the size of the hidden layer, both the train and validation loss curves are getting flat after an epoch of approximately 10,000. At the end of the training process, the train loss and the validation loss are the smallest for a hidden layer size of 20, i.e. 0.1472 and 0.1450, respectively. Therefore, the original 150 features have been "condensed" into 20 new features in the latent space for further clustering analyses.
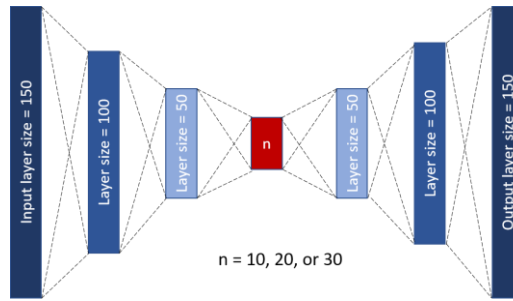

Figure 2: Structure of the stacked autoencoder. The input layers and the output layers are symmetric. *n* represents the number of neural nodes in the middle layer, which is varied using 10, 20, or 30

3.2.2 Clustering method

The conventional, unsupervised clustering method K-Means has been utilized to group the online shops that potentially have similar features. The elbow method has been used to determine the optimal number of clusters. Basically, the elbow method calculates the sum of squared distance between neighboring points (online shops) in the determined latent space. From Figure 3a, a clear "elbow" can be seen when the number of clusters equals 3. To verify the chosen number of clusters, a silhouette analysis has been performed for a cluster size of 2, 3, 4, 5, and 6. Correspondingly, the silhouette coefficients obtained are 0.289, 0.211, 0.159, 0.151, and 0.148. A cluster size of 4, 5, or 6 is relatively not good as it leads to a below-average silhouette
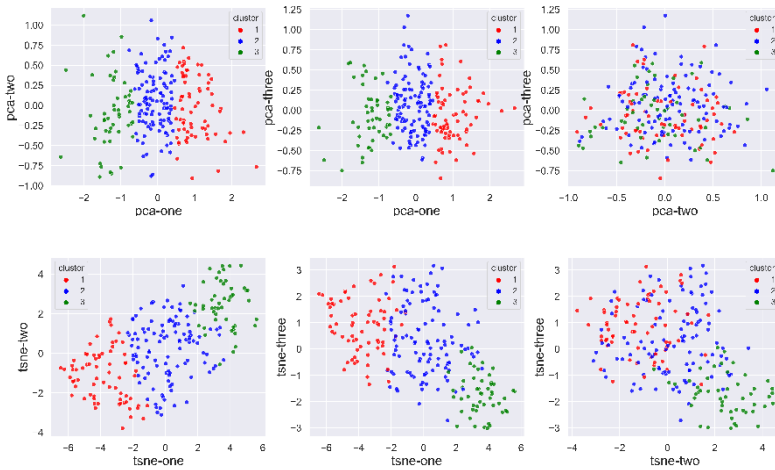
coefficient. In the end, the size of clusters is determined as 3. The online shops are grouped into three clusters (C1, C2, and C3) using the K-Means algorithm.
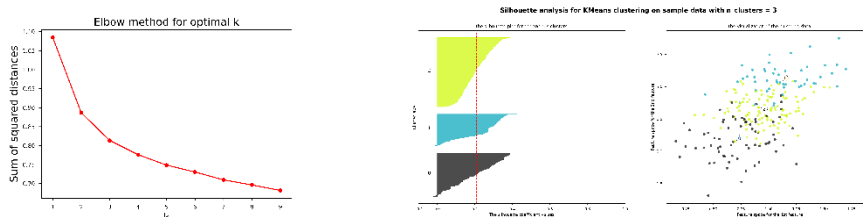
Figures 3a and b: Elbow method for determining the optimal number of clusters (a; left) and silhouette analysis for verifying the three-cluster solution (b; right)

3.2.3   Cluster visualization

To visualize the determined clusters of online shops on two-dimensional diagrams, both PCA and t-distributed stochastic neighbor embedding (t-SNE) methods are used. PCA is a deterministic, linear dimensionality reduction technique, which keeps the global structure of data by preserving their variance. Here, the first, second, and third principal components represent 78.4% of the variance in the original data, which is quite good. In comparison, t-SNE is a non-deterministic, non-linear dimensionality reduction technique, which is stronger at capturing local structures within the data and preserves distances between data points instead of their variances. In Figure 4, both the PCA and the t-SNE diagrams exhibit clear boundaries between clusters that can be distinguished mainly using the first principal component.

T. Wozniak, G. Lu, D. Georgi, A. Janoschka & A. Steigerwald
Exploring the Sustainability of Swiss Online Shops: Preliminary Evidence from a Clustering Approach

651

Figures 4a and b: PCA 2D plots (a; top) vs. t-SNE 2D plots (b; bottom)
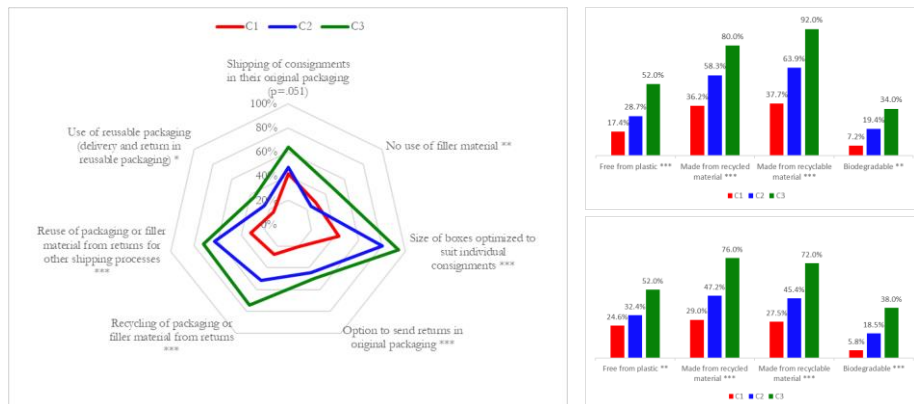
4   Results

In the survey instrument, we used five items on a 5-point scale to measure sustainability related to the element *products*. Similarly, five 5-point-scale items were used to assess the elements *online shop operations*, three items for *intralogistics and storage*, six items for *product returns*, and seven items for *cooperation with partners*. One-way ANOVA was performed to detect statistically significant differences between the clusters. Table 1 details the results for two exemplary items per framework element, highlighting distinct characteristics of each cluster.

Table 1: ANOVA summary table for two exemplary items for each of the following elements: products PROD, online shop operations OSO, intralogistics and storage ILS, product returns PR, and cooperation with partners CP

	Element of framework / Survey item	df	F *p*	C1 M *SD*	C2 M *SD*	C3 M *SD*
PROD	Environmental aspects (e.g. resource conservation … during manufacture) are just as important as product costs.	2, 223	6.1 *.003*	3.57 *1.23*	3.43 *1.28*	4.16 *1.17*
	Social aspects (e.g. fair working conditions during manufacture) are just as important as product costs.	2, 223	4.1 *.018*	4.03 *1.16*	4.06 *1.11*	4.54 *0.91*

OSO	We use IT solutions that leave the smallest possible carbon footprint (manufacture, operations, disposal).	2, 188	2.5 .081	2.95 1.33	2.81 1.32	3.37 1.54
	We use IT solutions that have minimal power requirements (incl. power for cooling).	2, 186	6.7 .002	2.85 1.35	2.78 1.40	3.67 1.34
ILS	Our warehouse is operated in a resource-friendly manner (e.g. low power consumption).	2, 210	8.9 .000	3.62 1.07	3.76 1.19	4.45 0.91
	We optimize transport capacities ecologically (e.g. to branches or distribution center).	2, 209	12.4 .000	3.37 1.18	3.39 1.28	4.36 0.99
PR	Returns are passed on to specialized retailers.	2, 194	9.5 .000	2.12 1.27	1.49 1.06	1.22 0.80
	Returns are destroyed.	2, 203	3.2 .042	2.11 1.30	2.13 1.36	1.59 1.02
CP	When choosing possible partners, we consider environmental aspects (e.g. resource conservation, avoidance of carbon emissions).	2, 220	7.5 .001	3.18 1.21	2.87 1.36	3.74 1.35
	When choosing possible partners, we consider social aspects (e.g. fair working conditions).	2, 216	6.5 .002	3.62 1.28	3.43 1.45	4.27 1.24

For the element *packaging*, we developed lists of specific characteristics of the packaging and the packaging system as multiple-answer checkox-type quesions (e.g. if the packaging material is free of plastic or if the size of boxes is optimized to suit individual consignments). For the element *delivery*, we developed a similar list for a variety of respective options (e.g. carbon-neutral shipping or bundling of partial deliveries). Figures 5 and 6 detail how the respective characteristics and options vary across the clusters. Cross tabulation and chi-square analyses revealed that most of the characteristics of the packaging systems are unequally distributed across the clusters at $p<.001$ (***), $p<.01$ (**), or $p<.05$ level (*; see Figure 5a). Characteristics of the packing material (see Figure 5b) and the filler material (see Figure 5c) are unequally distributed across the clusters at $p<.001$ (***) or $p<.01$ (**) level. Taken together, these results suggest that the identified clusters are quite distinct with regards to the *packaging* core element of our framework. Out of the seven *delivery* options displayed in Figure 6, three are unequally distributed across the clusters at $p<.001$ (***), $p<.01$ (**), or $p<.05$ level (*), namely consolidated shipping, use of locally adapted logistics solutions, and delivery to third-party collection points.

T. Wozniak, G. Lu, D. Georgi, A. Janoschka & A. Steigerwald
Exploring the Sustainability of Swiss Online Shops: Preliminary Evidence from a Clustering Approach

653

Figures 5a, b, and c: Packaging system (a; left), packaging material (b; upper right), and filler material characteristics (c; lower right) – percentages by cluster
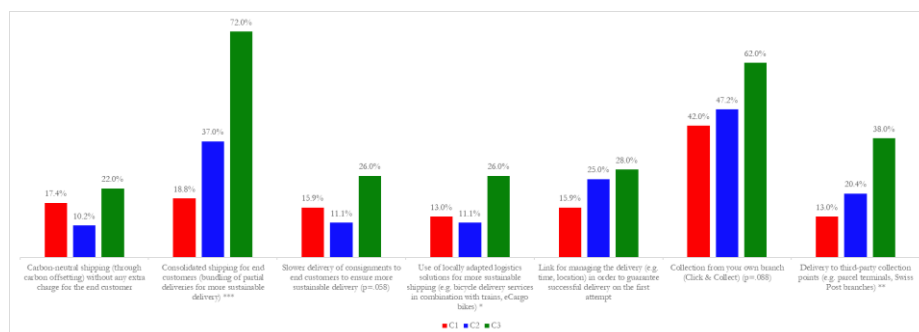

Figure 6: Delivery options offered to customers – percentages by cluster

5   Conclusion and outlook

Our clustering results indicate that online shops can be grouped into three clusters regarding their sustainability. With regard to many aspects of our proposed framework, the clusters are distinct from each other. Among them, cluster 3 has the most distinct characteristics and thus appears to comprise online shops that are most sustainable according to the criteria covered by our survey instrument. Further analysis of the present dataset needs to test to which extent the identified clusters also vary by structural properties of the online shops such as product categories, order volume, or for example D2C model. Future research should also investigate

the use of communication channels, possible differences between pure-play and omnichannel strategies, and potential latent drivers of sustainability engagement such as the values of the organizations behind the online shops.

Acknowledgements

References

Alkhayrat, M., Aljnidi, M., & Aljoumaa, K. (2020). A comparative dimensionality reduction study in telecom customer segmentation using deep learning and PCA. Journal of Big Data, 7(1). https://doi.org/10.1186/s40537-020-0286-0

Bertram, R. F., & Chi, T. (2018). A study of companies' business responses to fashion e-commerce's environmental impact. International Journal of Fashion Design, Technology and Education, 11(2), 254–264. https://doi.org/10.1080/17543266.2017.1406541

Cramer-Flood, E. (2020). Global Ecommerce 2020. Retrieved from https://www.emarketer.com/content/global-ecommerce-2020

Escursell, S., Llorach, P., & Roncero, M. B. (2020). Sustainability in e-commerce packaging: A review. Journal of Cleaner Production, 280. https://doi.org/10.1016/j.jclepro.2020.124314

European Commission. (2020). 2050 long-term strategy. Retrieved October 19, 2020, from https://ec.europa.eu/clima/policies/strategies/2050_en

Fichter, K. (2003). E-Commerce: Sorting Out the Environmental Cosequences. Journal of Industrial Ecology, 6(2), 25–41. https://doi.org/10.1162/108819802763471762

Frei, R., Jack, L., & Brown, S. (2020). Product returns: a growing problem for business, society and environment. International Journal of Operations & Production Management. https://doi.org/10.1108/IJOPM-02-2020-0083

Griggs, D. (2013). Sustainable developmnet goals for people and planet. Nature, 495, 305–307. https://doi.org/10.1038/495305a

Ignat, B., & Chankov, S. (2020). Do e-commerce customers change their preferred last-mile delivery based on its sustainability impact? International Journal of Logistics Management, 31(3), 521–548. https://doi.org/10.1108/IJLM-11-2019-0305

Lim, J. S., & Greenwood, C. A. (2017). Communicating corporate social responsibility (CSR): Stakeholder responsiveness and engagement strategy to achieve CSR goals. Public Relations Review, 43(4), 768–776. https://doi.org/10.1016/j.pubrev.2017.06.007

Lipsman, A. (2019). Global Ecommerce 2019. Retrieved October 13, 2020, from https://www.emarketer.com/content/global-ecommerce-2019

Mangiaracina, R., Marchet, G., Perotti, S., & Tumino, A. (2015). A review of the environmental implications of B2C e-commerce: a logistics perspective. International Journal of Physical Distribution and Logistics Management, 45(6), 565–591. https://doi.org/10.1108/IJPDLM-06-2014-0133

Osterwalder, A., & Pigneur, Y. (2010). Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. Hoboken, NJ: John Wiley & Sons.

Pålsson, H., Pettersson, F., & Winslott Hiselius, L. (2017). Energy consumption in e-commerce versus conventional trade channels - Insights into packaging, the last mile, unsold products and product returns. Journal of Cleaner Production, 164, 765–778.

T. Wozniak, G. Lu, D. Georgi, A. Janoschka & A. Steigerwald
Exploring the Sustainability of Swiss Online Shops: Preliminary Evidence from a Clustering Approach

655

https://doi.org/10.1016/j.jclepro.2017.06.242

Porter, M. E. (1985). The Competitive Advantage: Creating and Sustaining Superior Performance. New York: Free Press.

Reinartz, W., Wiegand, N., & Imschloss, M. (2019). The impact of digital transformation on the retailing value chain. International Journal of Research in Marketing, 36(3), 350–366. https://doi.org/10.1016/j.ijresmar.2018.12.002

Ricciotti, F. (2020). From value chain to value network: a systematic literature review. Management Review Quarterly, 70(2), 191–212. https://doi.org/10.1007/s11301-019-00164-7

Rüegg-Stürm, J., & Grand, S. (2019). Managing in a Complex World. Bern: Haupt.

Steffen, W., Rockström, J., Richardson, K., Lenton, T. M., Folke, C., Liverman, D., … Schellnhuber, H. J. (2018). Trajectories of the Earth System in the Anthropocene. Proceedings of the National Academy of Sciences of the United States of America, 115(33), 8252–8259. https://doi.org/10.1073/pnas.1810141115

UNCTAD. (2021a). COVID-19 and e-commerce: a global review.

UNCTAD. (2021b). The UNCTAD B2C E-commerce Index 2020.

Van Loon, P., McKinnon, A. C., Deketele, L., & Dewaele, J. (2014). The growth of online retailing: A review of its carbon impacts. Carbon Management, 5(3), 285–292. https://doi.org/10.1080/17583004.2014.982395

Van Loon, Patricia, Deketele, L., Dewaele, J., McKinnon, A., & Rutherford, C. (2015). A comparative analysis of carbon emissions from online retailing of fast moving consumer goods. Journal of Cleaner Production, 106(2015), 478–486. https://doi.org/10.1016/j.jclepro.2014.06.060

Weber, C. L., Hendrickson, C. T., Matthews, H. S., Nagengast, A., Nealer, R., & Jaramillo, P. (2009). Life cycle comparison of traditional retail and e-commerce logistics for electronic products: A case study of buy.com. 2009 IEEE International Symposium on Sustainable Systems and Technology, ISSST '09 in Cooperation with 2009 IEEE International Symposium on Technology and Society, ISTAS. https://doi.org/10.1109/ISSST.2009.5156681

SOCIAL ROBOTS FOR REDUCING MATHEMATICS HIATUSES IN PRIMARY EDUCATION, AN EXPLORATORY FIELD STUDY

MATTHIJS SMAKMAN,[1] KOEN SMIT,[1] ELINE LAN,[2]
THOMAS FERMIN,[2] JOB VAN LAGEN,[2] JULIA MAAS,[2]
DAVID VAN VLIET[2] & SAM LEEWIS[1]

[1] HU University of Applied Sciences Utrecht, Institute for ICT, Digital Ethics, the Netherlands; e-mail: name.surname@hu.nl
[2] HU University of Applied Sciences Utrecht, Institute for ICT, the Netherlands; e-mail: name.surname@student.hu.nl

Abstract Since the outbreak of COVID-19 schools have gone into lockdown and teachers have had to teach pupils online from home. When pupils go back to school, standard, contemporary learning methods do not seem to be enough to reduce incurred hiatuses. Social robots are slowly becoming an integral component of our society and have great potential as educational technology. This study explores how social robots in classrooms can contribute to reducing mathematics-related hiatuses in Dutch primary education (pupils from four till twelve years old). A social robot as a tutor is evaluated by means of a field study with children ($n = 43$) to compare a class working with the robot, to a class working without the robot. Multiple factors on learning effect are taken into account by using a survey. Our results demonstrate that a robot can take the role of a tutor and practice with pupils. The results are of interest to researchers in the field of human-robot interaction as well as to educational institutes who wish to understand the implications of adopting robots in education.

1   Introduction

Hiatuses have always been a reoccurring phenomenon within (primary) education and could have a negative impact on learning performance in the longer term (Luyten, Staman, & Vissch, 2013). Finding solutions to reduce hiatuses have always been an important subject in education, but has now become even more relevant due to the impact of the COVID-19 pandemic on schools (Rothan & Byrareddy, 2020). Many pupils are falling behind on their education because they are not able to physically attend school, which causes governments to take action and provide extra funds for reducing hiatuses (Ministerie van Algemene Zaken, 2020c). In addition, eight percent of the Dutch pupils in regular primary education in 2018/2019 were already suffering from hiatuses ("Leerlingen in (speciaal) basisonderwijs; migratieachtergrond, woonregio," 2020; Nederlandse Jeugdinstituut, 2019). One of the reasons for these hiatuses is the increasing shortage of teachers the Netherlands is currently facing (Cultuur en Wetenschap van Onderwijs, 2019). School leaders sending classes home, due to the absence of a teacher, possibly leading to performance loss, which also causes learning delays, is becoming a more regular phenomenon (Cultuur en Wetenschap van Onderwijs, 2019). A learning delay can have many adverse consequences, such as underperformance or inequality of opportunity (van Onderwijs, 2020).

COVID-19 has created large hiatuses in primary education (Keultjes, 2020; van der Heyden, 2020) and contemporary learning methods do not seem to be enough to reduce these hiatuses, therefore, new learning material must be developed. The Dutch government is currently investing 244 million euros for primary education into creating and developing extra learning materials to reduce COVID-19 related hiatuses (AVS, 2020; Ministerie van Algemene Zaken, 2020b). One of these solutions could be the use of social robots. They have been shown to be able to increase cognitive and affective outcomes and have achieved outcomes similar to those of human tutoring on restricted tasks (Belpaeme, Kennedy, Ramachandran, Scassellati, & Tanaka, 2018). One of the main benefits of social robots is their physical presence, which traditional learning technologies lack (Belpaeme et al., 2018).

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study

659

For this study, the focus will be on exploring the use of a social robot for teaching mathematics to pupils in Dutch primary education. This study aims to answer the following research question: *'can social robots be used for reducing mathematics hiatuses in primary education?'* By answering this question, we aim to provide insights into a potential new learning tool for primacy education for reducing hiatuses for mathematics, which is, especially now, an urgent social issue.

This paper is structured as follows. First, the background and related work will focus on hiatus factors and social robots as two main concepts. This is followed by a description of the utilized research methods. Next, the paper details how the data was collected. Then, the methods and techniques for data analysis are presented. Based on the data collection and analysis, the results of this study are presented. This is followed by the conclusions drawn from these results. Lastly, the limitations are discussed, and future research directions are described.

2   Background and Related Work

To explore the effect of social robots on reducing hiatuses, the definition of the term 'hiatus' must be established. A hiatus is a subjective concept. It is defined as the disadvantage in a particular learning area that a person has, compared to 'the average pupil' with the same age and the same level of education (Bannink, 2021). In the next section, we will detail the concept of hiatuses, as well as the concept of a social robot.

2.1   Hiatus factors

Many factors influence the forming of a hiatus in education. Research shows that the most recurring factors are *autonomy* (Meusen-Beekman, Joosten-ten Brinke, & Boshuizen, 2001; Ryan & Deci, 2000b; Sierens, Vansteenkiste, Goossens, Soenens, & Dochy, 2009), *motivation* (Lak, 2017; Ryan & Deci, 2000a), *learning environment* (Meusen-Beekman et al., 2001; Weiser & Riggio, 2010) and *causality* (Miller, Ferguson, & Byrne, 2000; U.S. Department of Education Office of Special Education Programs, 2003).

Autonomy refers to the experience of choice and psychological freedom with respect to one's study activities (Sierens et al., 2009). It involves being self-organizing and having a sense of choice over one's study behaviour (Sierens et al., 2009). Sub-factors here are self-regulation (Ryan & Deci, 2000b), self-effectiveness (Meusen-Beekman

et al., 2001) and basic needs (Meusen-Beekman et al., 2001). Competence, as an aspect of self-regulation and effectiveness, involves the experience of efficacy while completing a learning task (Meusen-Beekman et al., 2001). The need for relatedness, also an aspect of self-regulation and effectiveness, concerns feeling connected to others, like teachers or fellow pupils (Sierens et al., 2009). Home-tuition in The Netherlands was compulsory because of the COVID-19 pandemic (Ministerie van Algemene Zaken, 2020a). The feeling of being connected with others was reduced by studying at home (Odekerken-Schröder, Mele, Russo-Spena, Mahr, & Ruggiero, 2020). Through home tuition, the pupil will not be with others all day, but there will be online moments with others and, predominantly, offline moments without others ("Didactiek," 2020).

Motivation is an individual's drive or reason to achieve an action or performance, it can drive the person to a (desired) behavioural form (Karels, 2020). Motivation consists of a relationship between various factors, and can be divided into intrinsic and extrinsic motivation, including the biological (innate) and culture-dependent (learned) characteristics (Ryan & Deci, 2000a). In addition, the environment can play a role and several elements influence motivation. Descriptive studies have shown that some pupils enjoy mathematics, seek out mathematical problem situations, and excel in them (Middleton, 1995). Whereas others ('math anxious' students) have a fear of mathematics and avoid engaging in mathematical problem situations. In addition, although the utility and importance of mathematics are at least acknowledged by the majority of students even if not understood fully, this knowledge is not sufficient enough to motivate them to continue taking mathematics courses (Middleton, 1995).

Learning environment refers to the diverse physical locations, context, and cultures in which pupils learn (Education Reform, 2013). Since pupils may learn in a wide variety of settings, such as outside-of-school locations and outdoor environments, the term is often used as a more accurate or preferred alternative to a 'classroom' (Education Reform, 2013). The 'classroom' concept has more limited and traditional (Education Reform, 2013). Due to the COVID-19 pandemic, pupils have a disadvantage in social mobility (Cullinane & Montacute, 2020). The closing of schools has caused unpredicted challenges for everyone involved in the education domain (Cullinane & Montacute, 2020). These challenges mostly encompass (Cullinane & Montacute, 2020): 1) not being able to access additional support, 2) not

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study

661

having access to certain information online (due to the poor and rich gap), 3) widening the attainment gap due to extended lockdown(s), and 4) working in cramped housing conditions.

Causality is the demonstration of how one variable influences other variables in education. In this research, causality will refer to present disorders in the pupils. These disorders can be learning disorders such as Dyslexia or behavioural disorders such as Attention Deficit Hyperactivity Disorder (ADHD) or Oppositional Defiant Disorder (ODD). The presence of a disorder can have a possible negative impact on the learning ability of the pupils, which can lead to the pupil scoring worse than expected (de Meyer, 2019; Driessen, 1990; U.S. Department of Education Office of Special Education Programs, 2003).

2.2 Social robots

A social robot is an embodied object with a certain degree of autonomous behaviour that is specifically designed to socially interact with humans (Darling, 2012). The key elements for a social robot are the physical embodiment, social understanding and school behaviour, and interaction and communication with humans (Hameed, Tan, Thomsen, & Duan, 2016). There are different kinds of social robots, such as Pepper (humanoid and programmable robot) (SoftBank Robotics, 2021b), PARO (therapeutic robot) (Parorobots, 2014) and ROBEAR (nursing care robot) (Wilkinson, 2015). For this study, the focus will be on the NAO robot. The NAO robot is a humanoid and programmable robot with rounded features and is bipedal totalling 58 cm in height (SoftBank Robotics, 2021a). The NAO robot is the most often used robot in research related to social robots in education (Belpaeme et al., 2018).

Broadly speaking, there are three types of robots in education. These include 1) using a robot to learn pupils programming skills, 2) the robot being an object of learning as a means to understand what a robot is, and 3) a robot as a learning partner/social robot (for example, an assistant teacher or a fellow pupil).

The use of social robots has recently been explored in the educational domain and attention from researchers and practitioners is increasing (Kennedy, Baxter, & Belpaeme, 2014; Konijn & Hoorn, 2020). Robots can take the form of, for example, a tutor, learning buddy or teacher. Social robots have proved to have a positive effect

on pupils (Kennedy et al., 2014; Kory-Westlund & Breazeal, 2019). Furthermore, the potential advantages of social robots in education are a reduction in costs and time because teachers and pupils are assisted better in class (Pachidis et al., 2019). However, there are different factors regarding the positive effect on reducing hiatuses within the field of mathematics, such as emotional stability and openness (de Meyer, 2019; Driessen, 1990; Streur, 2016).

Social robots have the potential to make positive contributions to a range of human-centred activities in education (Belpaeme et al., 2013; Broadbent, Stafford, & MacDonald, 2009; Dautenhahn & Werry, 2004; Tapus, Mataric, & Scassellati, 2007). Research shows that one-on-one tutoring can lead to significant learning improvements (Moriguchi, Kanda, Ishiguro, Shimada, & Itakura, 2011). Furthermore, several studies found a positive learning effect on different areas of expertise such as learning words (Baxter, Ashurst, Read, Kennedy, & Belpaeme, 2017; Moriguchi et al., 2011; Mubin, Stevens, Shahid, Mahmud, & Dong, 2013), science or technology (Mubin et al., 2013), motor task training (Baxter et al., 2017), weight-loss programs (Baxter et al., 2017), and reducing puzzle-solving time (Baxter et al., 2017). These studies indicate that robots take advantage of, and amplify, the human-likeness to anthropomorphize inanimate objects (human-like, not alive objects).

3   Research methods

The goal of this study is to explore to what extent a social robot can contribute to reducing mathematics hiatuses in primary education. To do so, we will use a mixed-method approach, consisting of both quantitative and qualitative methods. This mix provides the opportunity to achieve method and data triangulation (Webster, 2007). In the next sub-sections, we will first describe the field experiment, followed by the quantitative and qualitative research methods.

3.1   Field Experiment

Six Dutch primary schools were invited to participate in this study. One primary school positively responded. The field experiment comprises three phases; 1) the Bareka pre-test, 2) pupils using the robot for a set time, and 3) the Bareka post-test. Two of those phases include a standardized test referred to as Bareka, which is a

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study

663

learning method for indicating a child's level of mathematics proficiency. The Bareka method is divided into exercises, ranging from addition and subtraction to multiplying, deviations and fractions, graphs and percentages, content, weight, decimal numbers, number concept, length, perimeter, and area.

The Bareka learning method was translated to the robotics platform *Robotsindeklas,* used by the NAO robot. This existing platform is used for teaching pupils and students with social robots such as the NAO robot and the Alpha Mini robot (Interactive Robotics, 2021). By using this existing platform, we created a Bareka learning application for the experiment. To ensure that the NAO's content was compatible with the teaching material of the school, two experienced primary school teachers reviewed the application. The teaching material consisted out of a short explanation for each subject accompanied by practice exercises.

The experiment consisted of a control- and an experimental group, wherein the experimental group was provided with a social robot as a supplement to the regular teaching methods.

3.2   Quantitative research methods

Two classes were included in the field experiment, ranging from pupils aged ten to twelve years old. The experimental group ($n = 20$) was taught by a social robot in addition to the teachings of their teacher. The control group ($n = 23$) was taught solely by their teacher, using standard teaching methods. The social robot was used from Monday until Friday for 10 minutes per group of pupils (maximum of four pupils per group). The experiment lasted about five weeks (from the 18th of November, 2020 till the 18th of December, 2020). The teacher in the experimental group, was 36 years of age, and had thirteen years of educational working experience. The teacher in the control group was 52 years of age and had 31 years of educational working experience.

To gain insight into the actual development of the pupils, a knowledge test was conducted before and after the experiment. The first Bareka test took place at the primary school under the supervision of the teachers. The second Bareka test took place digitally due to COVID-19 restrictions (home-tuition). The *Bareka* results range from one to four, where one is insufficient and four is very good.

To gain insight into the impact of the social robot on the factors that influence the forming of a hiatus in education, we used a validated questionnaire ("Autonomietool," 2019; "Motivatietool," 2019). The questionnaire was utilized in the context of the field experiment, by using an experimental design that occurs in a natural setting (in class/school) (Allen, 2017).

The questionnaire was sent in December 2020, during the experiment, to both the control and the experiment group. The questionnaire measured two factors related to hiatuses: 1) autonomy and 2) motivation. We excluded other hiatuses factors because including all other factors would result in a too long and complex questionnaire for children. This would potentially have resulted in missing data or a lack of valid data. However, we included the motivation factor because it is an overarching factor of the other factors. Factors such as causality and learning environment were excluded from this survey because details of these factors were indirectly required to answer the research question and are subject to future research.

Motivation is split up into the sub-factors 1) intrinsic motivation, 2) identified motivation, 3) extrinsic motivation, and 4) amotivation. The questionnaires were constructed using Leerling2020 ("Autonomietool," 2019; "Motivatietool," 2019) and were combined into one survey of 24 questions, eight questions about autonomy ("Autonomietool," 2019) and sixteen about motivation ("Motivatietool," 2019).

3.3   Qualitative research methods

Semi-Structured Interviews

Two semi-structured interviews were conducted with teachers during the field experiment. A semi-structured approach is well suited for the exploration of the perceptions and opinions of respondents, and enable probing for more information and clarification of answers (Barriball, 1994). The interviews were conducted to elicitate information about the pupils and teachers and their knowledge regarding Bareka and the way mathematics is taught. In the interviews, the measuring instrument (knowledge test) was discussed as well as the usage of the social robot. This was required to create the learning applications for the robot and to establish its role in the classroom.

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study

665

Focus group sessions

Because of COVID-19 restrictions during this study, it was not possible to physically observe the pupils while they were using the NAO robot. As an alternative, focus groups were set up, which enable discussion among pupils for a more holistic view regarding the use of the NAO and its learning materials. The focus group sessions were guided by the teachers.

A focus group was conducted with both groups at the end of the experiment. In each group, five focus groups were conducted. The focus groups consisted of four to six pupils who were around the same level of proficiency in mathematics. This number was chosen because it is the ideal group size (Kitzinger, 1995) and because of the total number of participants. During the focus group sessions, students were asked what the concept of motivation is and what motivates them, using the same focus group protocol.

4 Data analysis

The collected data was analysed using SPSS 27 and ATLAS.ti. SPSS was utilized for the analysis of quantitative data, while ATLAS.ti was utilized for the analysis of qualitative data.

4.1 Bareka pre-test, post-test and questionnaire analysis

To determine the difference between the performance of the control- and experimental group the data from the results of the Bareka test were analysed. A one-way analysis of variance (one-way ANOVA) was used to determine whether there were any statistically significant differences between the average of the two groups (control- and experimental group).

Using the Bareka data, the Hedges' g factor can be calculated in order to determine what the effect was of the progress of the pupils, taking into account the relatively small sample size (Hedges, 1981). The Hedges' g factor was calculated by using the sample size, average progress per group and the standard deviation of the progress measured. The value of the Hedges' g is related to the Cohens' d and can be used to determine the effect size (Becker, 2000). The effect size of around .20 is considered

a small effect, .50 a medium effect, and .80 a large effect (Becker, 2000; Cohen, 1992).

A Cronbach's analysis was conducted on each questionnaire factor to measure the reliability per item. A Cronbach's alpha with a value of .7 is an adequate level of inter-item reliability (Field, 2017).

4.2 ATLAS.ti

After the interviews had been conducted, they were first transcribed and put in ATLAS.ti. In ATLAS.ti, two coding rounds of thematic coding were conducted by two research team members to reduce bias. Based on the hiatus factors that were defined earlier, themes were chosen to use for thematic analysis. This method is focusing on identifying patterned meaning across a dataset (Braun & Clarke, 2021). For example, a code could be controlled teaching, which focuses on extrinsic motivation.

5 Results

The results of the field experiment (Bareka tests), questionnaire, and focus groups are presented in this section.

5.1 Bareka pre and post-tests

The results of Bareka were determined by the one-way ANOVA test. There was a statistically significant difference between the begin average (F $(1, 24.872) = 23.292$ $p < .001$) and end average (F $(1, 21.668) = 18.840$ $p < .001$) of the groups.
The contrast test revealed that group A had a significant mean difference compared to group B in the Begin Average, t $(24.872) = -4.826$, p $<.001$ and it revealed that the experimental group had a significant mean difference compared to the control group in the End Average, t$(21.668) = -4.340$, p $<.001$, see table 1.

Table 1: Contrast test results

	Contrast	Value of contrast	SE	t	df	Sig
Begin average	1	-.4472	.09265	-4.826	24.872	<.001
End average	1	-.4010	.09240	-4.340	21.668	<.001

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study

667

To determine the effect size, the Hedges' g was calculated. The calculated Hedges' g for the experimental group is 0.440 and for the control group is 0.6524. This means that there is a small positive effect in the experimental group and a medium positive effect in the control group, see table 2.

Table 2: Hedges' g results

Group	n	Mean	SD	Hedges' g
Exp.-Begin	18	2,673444	0,353707	0,4400
Exp.-End	18	2,835778	0,367611	
Control-Begin	23	3,120609	0,193880	0,6524
Control-End	23	3,236800	0,153894	

5.2 Questionnaire

A Cronbach's analysis was conducted on the Autonomy subscale of the survey. It was found that the subscale's alpha level did not have an adequate level of inter-item reliability (Cronbach's α = .54). However, further analysis revealed that by removing the item, *"My tutor always tells me what to do during class"*, the alpha could be raised (Cronbach's α = .64). Furthermore, the intrinsic motivation (Cronbach's α = .85), identified motivation (Cronbach's α = .78), extrinsic motivation (Cronbach's α = .70) and amotivation (Cronbach's α = .73) subscales of the survey had all an adequate level of inter-item reliability.

Table 3 shows the descriptive statics of each subscale, where the item *"My tutor always tells me what to do during class"* is removed. This ensured that the survey was more reliable.

Table 3: Questionnaire descriptive statics

Factor	Group	n	Mean	SD	SEM	Δ Mean
Autonomy	Experimental	20	3.875	.414	.092	.016
	Control	23	3.859	.399	.083	
Intrinsic motivation	Experimental	19	3.421	1.093	.251	.040
	Control	21	3.381	1.048	.229	
Identified motivation	Experimental	20	3.825	.770	.172	-.061
	Control	22	3.886	.763	.163	
Extrinsic motivation	Experimental	20	2.813	.996	.223	-.056
	Control	21	2.869	1.005	.219	
Amotivation	Experimental	20	1.625	.763	.171	-.011
	Control	22	1.636	.727	.155	

5.3   Semi-structured interviews

The interview method that was used is thematic analysis. There were themes generated that were focused on the hiatus factors. Example codes are shown in table 4. The results are based on the transcripts which are available upon request due to space limitations.

Table 4: Examples of coding results

Transcript	Code	Theme
"I work from my own agenda, because it has to be done."	Controlled teaching	Extrinsic motivation
"If children are very strong in mathematics and they have their own initiative and they want to do things, you can let them go more easily. And then I stimulate that."	Autonomy-supporting teaching	Intrinsic motivation, autonomy

The interviews mainly revealed that the teacher in the control group was in a more advanced phase of mathematics than the teacher in the experimental group. The pupils from the experimental group have a learning delay that takes more time to catch up to compared to the control group, according to the teachers.

The experimental group worked with the NAO robot. The value of the robot, in this context, would be to improve attention and motivation in the pupils, while the delivery and assessment are done by the human teacher.

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study

669

5.4 Focus group sessions

The focus group sessions revealed that the children's interest in subjects differed. Mathematics, English and physical exercises were the most frequently mentioned subjects of interest. Also, the children agreed that playing games were fun (with or without the use of the social robot). The team deliberately not asked about the robot, which was only mentioned once. No link was made by the pupils between motivation and a social robot that can help them deal with learning disadvantages/hiatuses.

6 Discussion and Conclusion

This paper aims to answer the following research question: *'can social robots be used for reducing mathematics hiatuses in primary education?'* The ANOVA identified a statistically significant difference between the begin- and end average of the groups (p <.001). The contrast test revealed that there was a significant difference between the begin- and end average of the experiment group, compared to the control group. However, the value contrast is higher in the begin average (= -.4472) compared to the end average (= -.4010). In addition, the calculated Hedges' g for the experimental group is 0.440 and for the control group is 0.6524. This means that there is a small positive effect for the experimental group and a medium positive effect for the control group. It may be concluded that the robot had a small effect on the average of the Bareka results. However, the control group showed a larger positive effect (medium >.5 versus small <.5) compared to the experimental group. Though, it cannot be concluded that this difference is caused by the social robot, as the effect of the social robot cannot be isolated to attribute to the learning performance of the pupils in this study. The interview and focus groups have shown that pupils in the experiment group had more difficulty with mathematics, which can also be seen in Bareka's average results.

This study has several limitations that should be discussed. firstly, this exploratory study consisted out of a limited number of participating children (*n* = 43). More participants are preferable and will result in increased reliability and generalizability in future research. Secondly, the teachers did not have previous experience with the use of social robots, and had limited time to practice with the robot before the experiment. This resulted in the teacher and pupils experimenting with the robot during the first lessons and having difficulty using the robot and application, which

affected the actual time to improve the learning effect using the robot. Lastly, while we argue that both groups are sufficient to determine a (potential) effect using the research methods presented, we discovered moderating variables that make it hard to isolate the effect of the social robot in this study setup, such as 1) the difference in intrinsic motivation regarding mathematics between both groups, 2) differences in mathematics progression and learning delays between both groups, 3) didactic styles and level of experience between both teachers, and 4) differences in diagnosis per pupil per group.

Overall, we conclude that a social robot can be used for reducing mathematic hiatuses, however, we could not conclude that a social robot is as effective as a human tutor. Furthermore, the high work pressure in primary education might have a negative effect on the use of social robots in primary education. Therefore, future research might focus on creating a plug and play social robot which allows teachers with limited robot experience to also experiment with this potentially promising technology.

Acknowledgements

References

Allen, M. (2017). Field Experiments. In The SAGE Encyclopedia of Communication Research Methods. 2455 Teller Road, Thousand Oaks California 91320: SAGE Publications, Inc. https://doi.org/10.4135/9781483381411.n200

Autonomietool. (2019). Leerling 2020.

AVS. (2020, May). 244 miljoen om leerachterstanden coronacrisis in te halen. Algemene Vereniging Schoolleiders.

Bannink, F. P. (2021). Wat is een leerachterstand? Retrieved March 31, 2021, from https://advanced-rt.nl/leerachterstand-2/

Barriball, L. (1994). Collecting data using a semi-structured interview: a discussion paper. Journal of Advanced Nursing (Vol. 19).

Baxter, P., Ashurst, E., Read, R., Kennedy, J., & Belpaeme, T. (2017, May). Robot education peers in a situated primary school study: Personalisation promotes child learning. PLOS ONE.

Becker, L. A. (2000). Effect Size (ES).

Belpaeme, T., Baxter, P. E., Read, R., Wood, R., Cuayáhuitl, H., Kiefer, B., … Humbert, R. (2013). Multimodal child-robot interaction: building social bonds. Journal of Human-Robot Interaction, 1(2), 33–53. https://doi.org/10.5898/JHRI.1.2.Belpaeme

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. Sci. Robot., 3(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study
671

Braun, V., & Clarke, V. (2021). Thematic analysis. Retrieved March 31, 2021, from https://www.psych.auckland.ac.nz/en/about/thematic-analysis.html

Broadbent, E., Stafford, R., & MacDonald, B. (2009). Acceptance of Healthcare Robots for the Older Population: Review and Future Directions. Int J of Soc Robotics, 1(4), 319. https://doi.org/10.1007/s12369-009-0030-6

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cullinane, C., & Montacute, R. (2020). COVID-19 and Social Mobility Impact Brief #1: School Shutdown, 11.

Cultuur en Wetenschap van Onderwijs. (2019, April). Oplopend lerarentekort bedreiging voor gelijke kansen in het onderwijs - De Staat van het Onderwijs - Inspectie van het onderwijs.

Darling, K. (2012, September). Extending Legal Protection to Social Robots. IEEE Spectrum: Technology, Engineering, and Science News.

Dautenhahn, K., & Werry, I. (2004). Towards interactive robots in autism therapy: Background, motivation and challenges. Pragmatics & Cognition, 12(1), 1–35. https://doi.org/10.1075/pc.12.1.03dau

de Meyer, B. (2019). Hoe zit ik in elkaar? Persoonlijkheid, temperament en rekenen bij kinderen met en zonder dyscalculatie., 66.

Didactiek. (2020, September). Les Op Afstand.

Driessen, G. W. J. M. (1990). De onderwijspositie van allochtone leerlingen: de rol van sociaal-economische en etnisch-culturele factoren, met speciale aandacht voor het onderwijs in eigen taal en cultuur. Nijmegen : Instituut voor Toegepaste Sociale Wetenschappen.

Education Reform. (2013). Learning Environment. Retrieved March 31, 2021, from https://www.edglossary.org/learning-environment/#:~:text=Learning environment refers to the,cultures in which students learn.

Field, A. (2017). Discovering statistics using IBM SPSS statistics (5th editio). Thousand Oaks, CA: SAGE Publications.

Hameed, I. A., Tan, Z.-H., Thomsen, N. B., & Duan, X. (2016). User Acceptance of Social Robots, pp. 274-27, 7.

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect size and Related Estimators. Journal of Educational Statistics, 6(2), 107–128. https://doi.org/10.3102/10769986006002107

Interactive Robotics. (2021). Educatieve robots voor het primair onderwijs. Retrieved March 31, 2021, from https://www.interactive-robotics.com/onderwijs/primair-onderwijs/

Karels, M. (2020, March). Motivatie - intrinsieke en extrinsieke motivatie.

Kennedy, J., Baxter, P., & Belpaeme, T. (2014). Using Social Robots in Child Education, 3.

Keultjes, H. (2020, October). Kabinet trekt 38 miljoen uit voor 'herfstscholen' tegen corona-leerachterstand. AD.Nl.

Kitzinger, J. (1995). Qualitative Research: Introducing focus groups. BMJ, 311(7000), 299–302. https://doi.org/10.1136/bmj.311.7000.299

Konijn, E. A., & Hoorn, J. F. (2020). Robot tutor and pupils' educational ability: Teaching the times tables. Computers \& Education, 157, 103970. https://doi.org/10.1016/j.compedu.2020.103970

Kory-Westlund, J. M., & Breazeal, C. (2019). Exploring the Effects of a Social Robot's Speech Entrainment and Backstory on Young Children's Emotion, Rapport, Relationship, and Learning. Front. Robot. AI, 6. https://doi.org/10.3389/frobt.2019.00054

Lak, E. (2017). Intrinsieke motivatie van kinderen in het speciaal basisonderwijs.

Leerlingen in (speciaal) basisonderwijs; migratieachtergrond, woonregio. (2020, May). Statline.

Luyten, H., Staman, L., & Vissch, A. (2013). Leerachterstanden van vertraagde leerlingen op normaalvorderende leeftijdgenoten. Pedagogische Studiën, 45–57.

Meusen-Beekman, K., Joosten-ten Brinke, D., & Boshuizen, E. (2001). De retentie van zelfregulatie, motivatie en self-efficacy in het voortgezet onderwijs na formatieve assessments in het basisonderwijs. Pedagogische Studiën, 136–153.

Middleton, J. A. (1995). A Study of Intrinsic Motivation in the Mathematics Classroom: A Personal
        Constructs Approach. Journal for Research in Mathematics Education, 26(3), 254.
        https://doi.org/10.2307/749130

Miller, A., Ferguson, E., & Byrne, I. (2000). Pupils' causal attributions for difficult classroom behaviour.
        British Journal of Educational Psychology, 70(1), 85–96.
        https://doi.org/10.1348/000709900157985

Ministerie van Algemene Zaken. (2020a, March). Aanvullende maatregelen onderwijs, horeca, sport.

Ministerie van Algemene Zaken. (2020b, May). Extra geld onderwijs door coronacrisis - Financiering
        onderwijs - Rijksoverheid.nl.

Ministerie van Algemene Zaken. (2020c, May). Vermijd drukte, houd 1,5 meter afstand - Nieuwsbericht
        - Rijksoverheid.nl.

Moriguchi, Y., Kanda, T., Ishiguro, H., Shimada, Y., & Itakura, S. (2011). Can young children learn
        words from a robot? IS, 12(1), 107–118. https://doi.org/10.1075/is.12.1.04mor

Motivatietool. (2019). Leerling 2020.

Mubin, O., Stevens, C. J., Shahid, S., Mahmud, A. Al, & Dong, J.-J. (2013). A review of the applicability
        of robots in education. Technology for Education and Learning 2013, 1(9).
        https://doi.org/10.2316/Journal.209.2013.1.209-0015

Nederlandse Jeugdinstituut. (2019, May). Cijfers over Onderwijsachterstanden.

Odekerken-Schröder, G., Mele, C., Russo-Spena, T., Mahr, D., & Ruggiero, A. (2020). Mitigating
        loneliness with companion robots in the COVID-19 pandemic and beyond: an integrative
        framework and research agenda. Journal of Service Management, 31(6), 1149–1162.
        https://doi.org/10.1108/JOSM-05-2020-0148

Pachidis, T., Vrochidou, E., Kaburlasos, V. G., Kostova, S., Bonković, M., & Papić, V. (2019). Social
        Robotics in Education: State-of-the-Art and Directions. In N. A. Aspragathos, P. N.
        Koustoumpardis, & V. C. Moulianitis (Eds.), Advances in Service and Industrial Robotics (pp.
        689–700). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-
        00232-9_72

Parorobots. (2014). PARO Therapeutic Robot. Retrieved March 31, 2021, from
        http://www.parorobots.com/

Rothan, H. A., & Byrareddy, S. N. (2020). The epidemiology and pathogenesis of coronavirus disease
        (COVID-19) outbreak. Journal of Autoimmunity, 109, 102433.
        https://doi.org/10.1016/j.jaut.2020.102433

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and Extrinsic Motivations: Classic Definitions and New
        Directions. Contemporary Educational Psychology, 25(1), 54–67.
        https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2000b). Self-Determination Theory and the Facilitation of Intrinsic
        Motivation, Social Development, and Well-Being. American Psychologist, 11.

Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B., & Dochy, F. (2009). The synergistic
        relationship of perceived autonomy support and structure in the prediction of self-regulated
        learning. British Journal of Educational Psychology, 79(1), 57–68.
        https://doi.org/10.1348/000709908X304398

SoftBank Robotics. (2021a). NAO the humanoid and programmable robot. Retrieved March 31, 2021,
        from https://www.softbankrobotics.com/emea/en/nao

SoftBank Robotics. (2021b). Pepper the humanoid and programmable robot. Retrieved March 31,
        2021, from https://www.softbankrobotics.com/emea/en/pepper

Streur, M. (2016, September). Een model hoe BIS en BAS chronisch pijngedrag beïnvloeden.
        Psychfysio Opleidingen.

Tapus, A., Mataric, M. J., & Scassellati, B. (2007). Socially assistive robotics [Grand Challenges of
        Robotics]. IEEE Robotics Automation Magazine, 14(1), 35–42.
        https://doi.org/10.1109/MRA.2007.339605

U.S. Department of Education Office of Special Education Programs. (2003). Identifying and Treating
        Attention Deficit Hyperactivity Disorder: A Resource for School and Home. Retrieved March
        31,   2021,   from

M. Smakman, K. Smit, E. Lan, T. Fermin, J. van Lagen, J. Maas, D. van Vliet & S. Leewis:
Social Robots for Reducing Mathematics Hiatuses in Primary Education, an Exploratory Field Study

673

> http://www.ldonline.org/article/Identifying_and_Treating_Attention_Deficit_Hyperactivity
> _Disorder%3A_A_Resource_for_School_and_Home

van der Heyden, W. (2020, November). Vooral kinderen van laagopgeleide ouders hebben leerachterstand door corona. RTL Nieuws.

van Onderwijs, C. en W. (2020, November). Onderwijs tijdens COVID-19: scholen en instellingen hebben zorgen over de continuïteit van het onderwijs en over leerachterstanden van leerlingen en studenten - Nieuwsbericht - Inspectie van het onderwijs.

Webster, M. (2007). Research Methods. Journal of Business, 5(3), 8.

Weiser, D. A., & Riggio, H. R. (2010). Family background and academic achievement: does self-efficacy mediate outcomes? Soc Psychol Educ, 13(3), 367–383. https://doi.org/10.1007/s11218-010-9115-1

Wilkinson, J. (2015). The strong robot with the gentle touch.

Artificial Intelligence Value Alignment Principles: The State of Art Review from Information Systems Research

Shengnan Han[1] & Shahrokh Nikou[2]

[1] Stockholm University, Department of Computer and Systems Sciences; e-mail: shengnan@dsv.su.se
[2] Åbo Akademi University & Stockholm University; e-mail: shahrokh.nikou@abo.fi

Abstract Information and communication technologies (ICTs) must be designed and used for humane ends. The rapid adoption of Artificial Intelligence (AI) has raised the critical question of whether we can ensure AI's alignment with human values to guide its design and use. We perform a selective literature review with the specific search terms of the papers published in the top information systems (basket of 8 journals and 5 AI journals in IS) from 2000-2020 to answer this question. The findings indicate that IS research has contributed insufficiently to a deeper understanding of human values and AI value alignment principles. Moreover, the mainstream IS research on AI is mostly dominated from its technical and managerial aspects. Thus, the future research agendas are proposed accordingly. The paper provides some food for thoughts in studying human values and AI alignment within the context of IS research.

1   Introduction

Berente et al. (2019) define AI as machines performing cognitive functions that we typically associate with humans, including perceiving, reasoning, learning, and interacting with others. They emphasize that "AI is not confined to one or a few applications, but rather is a pervasive economic, societal, and organizational phenomenon" (p. 1). To achieve what Walsham (2012) has argued that we must direct ICT at humane ends, AI should be aimed at making this a better world by using its highly optimized mechanistic functions and super intelligence to serve human needs, satisfy human desires and to maximize the realization of human values (e.g., Yudkowsky, 2011). This is also proposed as the AI value alignment principles (Russell, 2019) or as Sutrop (2020) put forward designing AI that conforms to human values is called 'value alignment'. One fundamental and critical question is raised and intensively debated: *how can we ensure AI alignment with human values through AI operations from design to use?*   Yamposkiy (2017) argued that, because of the unresolved disagreements in the disciplines of philosophy and axiology regarding the nature and content of human values, the question of how to align these values in AI development and use, is also moot. The IS community has not yet paid sufficient attention to this AI phenomenon and has contributed insufficiently to a deeper understanding of human values in general (Carman and Rosman, 2020, Lyytinen et al., 2020). In this paper, we first analyze the AI phenomenon as it is discussed in the top IS research outlets (basket of 8 journals and 5 AI journals in IS). It should be noted that the AI value alignment and its connection to ethical concerns is not included in the search because, while significant, this topic is not the paper's focus. Upon the results from the literature review, we propose the future research agendas.

2   AI Value Alignment Principles

Russell (2019) has proposed the three AI value alignment principles for creating a safe and beneficial AI. (1) A principle of altruism: the AI's only objective is to maximize the realization of human values. Here, human values are defined as what "we" would "prefer our life to be like". (2) A law of humility: AI as the digital agents is initially not certain of what human values are. But AI agents, in support of advanced machine learning capabilities, may learn those values and preferences by observing "our" behaviors. (3) To achieve the value alignments between AI and

humans, we, in this process, must learn to be better persons, or, perhaps, simpler. The aim should be ensuring that AI agents can learn the essential value-goods such as safety, healthcare, food and shelter, and meaningful work from "us". AI agents must be explicably programmed to make such values primary where and when needed. We acknowledge that advanced technical solutions are not sufficient for fulfilling the AI value alignment principles (e.g. Christian, 2020). Therefore, the multiple aspects of human values should be fully explored in AI design and use.

3   Research Method

We followed Lowry et al. (2004) recommendation to select the journals and articles. We searched Web of Science, INFORM, EBSCO, and Google Scholar. We separately also searched the basket of 8 IS journals and the top 5 AI journals in IS. As inclusion criteria, an article had to be original study and published in that top IS journal between 2000-2020 and written in English. Moreover, to be included in the review, articles had to match exact the search terms used during the publication search. As we used specific search terms (such as artificial intelligence AND human value*, or AI AND human value*), the initial database search retrieved 327 articles. In the next step, we excluded all duplicated and articles that did not adhere to our search criteria (n = 302 in total). The most frequent reason for excluding an article was that, although drawing to some extend on AI, the article did not primarily use AI in the context of human value or focused mainly on ethical issues. This final dataset is composed of 25 AI articles[1], which were downloaded in full text and reviewed by authors. After reviewing the 25 articles, it has become clear that none of the articles, although appeared to be relevant, discussed or approached "*AI and human value*" like our current approach. Nonetheless, the review results are briefly presented in the next section.

4   IS Research on "AI and Human Values"

The IS community has provided limited exploration of human values and of the possible AI alignment with those human values. Most of the reviewed articles limit their contributions towards AI technical problems (e.g., Li et al., 2009; Wong et al., 2020), and very few have implicitly discussed AI's impacts on humans, organizations,

---

[1] Due to page limitation, we cannot include a full list of all authors' information in the reference list.

and society in general. For example, Ransbotham et al. (2016, p. 1) argued that while IT provides many advantages to humans, their organizations, and to society in general, also have the potential to create new vulnerabilities such as online harassment, incivility, a merely algorithmic ethics, and bias towards minorities. In another study, Aleksander (2017) argued that, as robots and other machines operate in an algorithmic way and not in a truly cognitive and conscious human way, AI can present serious threats to humanity if the algorithms are not aligned with broader sets of values than those of pragmatic efficiency. Elkins et al. (2013) demonstrated that using artificial technology and integrating AI into advanced expert systems inadvertently imposes threats even to human experts and inhibits users from adopting the technology. Nonetheless, Aleksander (2004) argued that as AI technology develops more and more, it has greater potential for overcoming some of the unforeseen difficulties as humans pursue some very ambitious projects. Glezer (2003, p. 65) argued that using AI for automation of tasks is problematic, for the software agents often interfere with the human ability to specify the amount of control they would like to have over the agent's behavior. Nicolescu et al. (2018) investigated the emerging meanings of "value" associated with the Internet of Things (IoT) and argued that the multiple meanings of "value" are invariably articulated at the juncture of three domains: social, economic, and technical. Huysman (2020) asserted that we should create societal awareness about the rise of low quality of work due to AI rather than focusing merely on the effect of AI on job losses. Grønsund and Aanestad (2020, p. 14) argued that while research on algorithmic and intelligent technologies has generated insights about their potential to replace human work; however, the emergent configurations by which humans and algorithmic interplay emerge has not been investigated. In summary, the current AI research in IS field is restricted to technical developments and design issues. AI design and its alignment with human values are not yet being fully considered.

5 Discussion and Future Research Agendas

The study results clearly demonstrate that IS research has not yet sufficiently contributed to a deeper understanding of human values and AI value alignment and how to achieve the AI value alignment principles. Thus, we propose the following three research foci. First, we need to understand what are human values from different philosophical and ontological schools of thoughts within IS. Ågerfalk (2020) argues that IS research can contribute significantly to advance AI

development and use. We need to focus on the three key components, contextualization, communication, and practice to complete the inquires of AI phenomenon. AI phenomenon is much more complex with great uncertainties. We can explore this complexity from various school of thoughts with the aim of producing more comprehensive understanding of what are the human values and add IS perspectives to this multidisciplinary theme. Second, we need to understand what are the critical human values within the contexts of AI design and use. Human values are deeply rooted in cultural and social traditions. Gabriel (2020) points out that human preferences that are always embedded in a range of human values, may not be sufficient, though necessary, to give instructions to an AI agent for achieving desired outcomes. Our "immediate" preferences may differ largely than what we prefer in the longer time. AI systems are kind of IS artefacts (Chatterjee et al., 2017; Lee et al., 2015). The study of the critical human values can be conducted in the context of AI as an information artefact, AI as a technology artefact and AI as a social artefact. As well we need to study the interactions among the three artefacts components to reach the conclusion, i.e., what are the critical human values should be aligned with AI design. Third, we need to prioritize the critical human values in AI design and use for different user groups in various cultural, social, and personal contexts. To keep a positive reciprocal relationship of human/AI, we need to become a better person and use AI in a positive and ethical way. This is also what the AI alignment principles have proposed. Since the AI may learn from us (a law of humitity), we need to behave well and generate more positive values that benefit AI design. Thus, future studies can investigate the effects of the prioritized human values on the users' behaviors towards AI systems within a specific context.

6 Conclusions

We briefly address current IS research on human values and AI in this paper, as well as some perspectives on the importance of achieving AI value alignment principles. We also suggest three research directions that IS researchers can pursue, but they are tentative and might be naïve in their current state of development. We believe that the paper provides some food for thought about the significance of studying human values in AI design and application in IS research.

References

Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. European Journal of Information Systems, 29 (1), 1-8.

Aleksander, I. (2004). Advances in intelligent information technology: re-branding or progress towards conscious machines? Journal of Information Technology, 19 (1), 21-27.

Aleksander, I. (2017). Partners of humans: a realistic assessment of the role of robots in the foreseeable future. Journal of Information Technology, 32 (1), 1-9.

Berente, N., Gu, B., Recker, J., and Santhanam, R. (2019). Managing AI. Call for Papers, MIS Quarterly, 1-5.

Carman, M., and Rosman, B. (2020). Applying a principle of explicability to AI research in Africa: should we do it? Ethics and Information Technology. Available at: https://doi.org/10.1007/s10676-020-09534-2.

Chatterjee, S., Xiao, X., Elbanna, A., and Sarker, S. (2017). The Information Systems Artifact: A Conceptualization Based on General Systems Theory. Paper presented at the Hawaii International Conference on System Sciences.

Christian, B (2020). The alignment problem: machine learning and human values, W. W. Norton & Company, New York.

Elkins, A. C., Dunbar, N. E., Adame, B., and Nunamaker, J. F. (2013). Are users threatened by credibility assessment systems? Journal of Management Information Systems, 29 (4), 249-262.

Gabriel, I. (2020). Artificial Intelligence, Values, and Alignment. Minds and Machines, https://doi.org/10.1007/s11023-020-09539-2.

Glezer, C. (2003). A conceptual model of an interorganizational intelligent meeting-scheduler (IIMS). The Journal of Strategic Information Systems, 12 (1), 47-70.

Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. The Journal of Strategic Information Systems, 29(2), 101614.

Huysman, M. (2020). Information systems research on artificial intelligence and work: A commentary on "Robo-Apocalypse cancelled? Reframing the automation and future of work debate". Journal of Information Technology, 0268396220926511.

Lee, A., Thomas, M., and Baskerville, R. L. (2015). Going back to basics in design science: from the information technology artifact to the information systems artifact. Information Systems Journal, 25, 5-21.

Li, X., Chen, H., Zhang, Z., Li, J., and Nunamaker, J. F. (2009). Managing knowledge in light of its evolution process: An empirical study on citation network-based patent classification. Journal of Management Information Systems, 26 (1), 129-154.

Lowry, P., Romans, D., and Curtis, A. (2004). Global Journal Prestige and Supporting Disciplines: A Scientometric Study of Information Systems Journals. Journal of the Association for Information Systems 5(2), 29-77.

Lyytinen, K., Nickerson, J. V., and King, J. L. (2020). Metahuman systems= humans+ machines that learn. Journal of Information Technology. Available at: https://doi.org/10.1177/0268396220915917.

Nicolescu, R., Huth, M., Radanliev, P., and De Roure, D. (2018). Mapping the values of IoT. Journal of Information Technology, 33 (4), 345-360.

Ransbotham, S., Fichman, R. G., Gopal, R., and Gupta, A. (2016). Special section introduction—ubiquitous IT and digital vulnerabilities. Information Systems Research, 27 (4), 834-847.

Russell, S. (2019). Human compatible: Artificial intelligence and the problem of control. Penguin.

Sutrop, M. (2020). Challenges of Aligning Artificial Intelligence with Human Values. Acta Baltica Historiae et Philosophiae Scientiarum, 8(2).

Walsham, G. (2012). Are we making a better world with ICTs? Reflections on a future agenda for the IS field. Journal of Information Technology, 27(2), 87-93.

Wong, N., Ray, P., Stephens, G., and Lewis, L. (2012). Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results. Information Systems Journal, 22(1), 53-76.

Yampolskiy, R. V. (2017). AI Is the Future of Cybersecurity, for Better and for Worse. Harvard
    Business Review.
Yudkowsky, E. (2011). Complex Value Systems are Required to Realize Valuable Futures. The
    Singularity Institute, San Francisco, CA. http://intelligence.org/files/ComplexValues.pdf.

THE IMPACT OF COMPUTER-MEDIATED DELAYED FEEDBACK ON DEVELOPING ORAL PRESENTATION SKILLS: AN EXPERIMENTAL STUDY IN VIRTUAL REALITY

BO SICHTERMAN, MARIECKE SCHIPPER,
MAX VERSTAPPEN, PHILIPPINE WAISVISZ &
STAN VAN GINKEL

Utrecht University of Applied Sciences, Utrecht, The Netherlands; e-mail: bo.sichterman@hu.nl, mariecke.schipper@hu.nl, verstappen.m.a@gmail.com, philippine.waisvisz@hu.nl, stan.vanginkel@hu.nl

Abstract Previous studies emphasize that feedback is essential for acquiring presentation skills. However, it remains unknown whether computer-mediated delayed feedback, provided in Virtual Reality (VR) without the intervention of the teacher, impacts students' public speaking skills. Recent technological developments allowed to convert quantitative information from VR-systems into qualitative feedback messages that directly relate to the standards for high-quality feedback. This experimental field study, therefore, focuses on the impact of automated, qualitative feedback messages in a VR-system on students' presentation skills development (n = 60). The effects are compared with a validated condition in which the delayed VR-feedback is delivered by an expert. Mixed methods, including validated rubrics and self-evaluation tests, are used for data collection. This study aims to refine educational design principles concerning effective feedback in presentation curricula. Furthermore, the results should provide insights about supporting feedback processes while releasing the pressure on resources such as time and staffing.

1   Introduction

Presenting is considered as a crucial competence for higher-educated professionals (Van Ginkel, Gulikers, Biemans, & Mulder, 2015). It is a prerequisite for effective communication and successful performance within professional practice (De Grez, Valcke & Roozen, 2019). However, young professionals often fail to show effective presentation behaviors in these environments (Van Ginkel *et al.,* 2019).

Research on public speaking defines oral presentation competence as "a combination of knowledge, skills and attitudes needed to speak in public in order to inform, self-express, relate and to persuade" (De Grez, 2009, p.5). Developments in this competence are related to seven educational design principles (Van Ginkel *et al.,* 2015), including three principles focused on formative assessment strategies (i.e. expert feedback, peer feedback and self-assessment) (Van Ginkel, Gulikers, Biemans & Mulder, 2017a). While feedback is crucial for students' learning (Hatti & Timperley, 2007), research has shown that innovative technologies as VR are valuable for delivering feedback within oral presentation learning tasks. Specifically, VR offers the opportunity to practice oral presentations within a virtual learning environment that imitates a presentation environment (Merchant Goetz, Cifuentes, KeeneyKennicutt & Davis, 2014). Additionally, VR provides automated feedback on, for example, non-verbal communication aspects (e.g. eye contact and use of voice) both during (i.e. computer-mediated immediate feedback) as well as after (i.e. computer-mediated delayed feedback) presentation practice (Belboukhaddaoui & Van Ginkel, 2019; Van Ginkel *et al.,* 2019; Van Ginkel, Ruiz, Mononen, Karaman, De Keijzer & Sitthiworachart, 2020).

Studies have shown that a VR-based oral presentation task with computer-mediated delayed feedback is as effective for students' development of presentation competencies as traditional face-to-face approaches (Van Ginkel *et al.,* 2019). Nevertheless, the delayed feedback of the VR-system was reflected in quantitative data feedback reports and, consequently, interpreted by an expert. However, nowadays, technological developments allow the conversion of quantitative information into qualitative feedback messages that directly relate to the standards for high-quality feedback (Van Ginkel, Gulikers, Biemans & Mulder, 2017b). Therefore, this study aims to investigate the impact of computer-mediated delayed feedback with qualitative feedback messages – interpreted without the intervention

of an expert - on students' presentation skills. It is hypothesized that the effect of computer-mediated delayed feedback with qualitative feedback messages on students' presentation skills is equal to feedback delivered by an expert. Results of this study provide insights about optimizing feedback processes supported by AI-technologies while releasing the pressure on resources such as time and staffing, as mentioned by the UN (Adubra, Da Silva, Dhungana, Mohan, Saltsman, & Van Ginkel, 2019). Moreover, findings of this research could further refine existing educational design principles concerning effective feedback within oral presentation learning trajectories.

2   Theoretical Framework

Over the past decades, VR received much attention within the educational context. From a scientific perspective, several studies demonstrated the impact of this technology on students' learning (Merchant *et al.,* 2014). Moreover, from a logistical perspective, the implementation of VR has potentials in supporting teachers with instructions and providing feedback. Therefore, it has been suggested that VR might be beneficial in overcoming educational issues as teacher shortages and the increasing number of students in higher education worldwide (Adubra *et al.,* 2019; Parmigiani, Van Ginkel, Saltsman, & Dhunga, 2020).

Previous research has shown the effectiveness of VR-based presentation tasks including automated feedback on students' development of oral presentation competencies (Belboukhaddaoui & Van Ginkel, 2019; Boetje & Van Ginkel, 2020; Van Ginkel *et al.,* 2019; Van Ginkel *et al.,* 2020). Additionally, students highly appreciated the automated feedback and perceived the VR environment as motivating (Van Ginkel *et al.,* 2019). The present aimed experimental study contributes to the scientific literature in this field for several reasons. In summary, it can be stated that:

1. While previous experimental field studies focused on VR-based tasks for developing presentation skills and guaranteed ecological validity, several scientific discussions point to intervening factors in presentation modules (such as instructions or feedback from other sources) that negatively impact the internal validity (Van Ginkel *et al.,* 2019). Therefore, an experimental

study that focuses on the effect of computer-mediated delayed feedback with qualitative feedback messages, conducted in a controlled lab-setting, is needed to critically evaluate the value of this type of feedback.

2. While previous studies on VR and presentation skills focused on comparing the computer-mediated feedback condition with a situation in which students receive expert feedback (Van Ginkel *et al.,* 2019), VR-research focusing on the impact of computer-mediated delayed feedback on presentation skills - without the support of the expert - is lacking hitherto. Therefore, this experimental study tests the impact of computer-mediated delayed feedback on presentation skills solely provided by the VR-system.

3. Although previous studies focused on the impact of VR-technologies for developing presentation competencies, knowledge, skills and attitudes towards presenting were taken into account (Van Ginkel *et al.,* 2019). However, as addressed earlier in other VR-studies (Boetje & Van Ginkel, 2020), other factors might impact the learning outcomes as well. Therefore, in this study, students' perceptions towards the adopted VR-technology and feedback modalities are included, since these factors could serve as crucial intermediate variables according to the current literature (Merchant *et al.,* 2014).

4. Although previous studies on VR and presentation skills mainly discussed the impact of the learning environment on learning outcomes, student characteristics are scarcely integrated in these studies. However, as reported in VR-studies (e.g. Merchant *et al.,* 2014; Van Ginkel *et al.,* 2019), students' perceptions of VR differ depending on their preferred learning activities. In order to test the generalizability of the impact of VR on varying cohorts of students with regard to their presentation skills development, it is suggested to incorporate the following student characteristics in this experimental study: (1) students' traits (such as gender, age, and educational level) and (2) experienced versus non-experienced students regarding presenting in VR.

3   Method

In the last semester of the school year 2020-2021, students (n=60) of a Dutch University of Applied Sciences will be recruited for the experiment. Participants will be informed about the intention of the study and have to confirm their informed consent. Additionally, the Netherlands Code of Conduct for Scientific Practice is adopted to ensure research integrity (Van Ginkel *et al.* 2019).

The participants in this study will be randomly assigned to either the experimental condition or the control condition. The experimental condition (n=30) consists of a VR-based oral presentation learning task including computer-mediated delayed feedback with qualitative feedback messages. In this condition, students used a VR-headset to practice a five-minute presentation on a self-selected topic to a virtual audience consisting of virtual students in a virtual classroom. After the presentation practice, students will receive automated feedback with qualitative feedback messages - delivered by the VR-system - on non-verbal communication aspects including eye contact and use of voice (i.e. pace, volume and frequency). This feedback will be interpreted by the students without the intervention of an expert.

The control condition (n=30) involves a similar VR-based learning task that consists of oral presentation practice within a similar VR classroom environment and computer-mediated delayed feedback. However, feedback on the VR-registered non-verbal communication aspects will be delivered by an expert.

The present study involves an experimental post-test design. Firstly, participants in both conditions will receive a brief introduction about the experiment, the presentation task and how to use the VR set-up. Secondly, participants will practice their presentation within one of the two conditions in the VR environment. Finally, as a final session, students in both conditions present their presentation in front of a small audience – comparable to the audience within the VR classroom environment - within a face-to-face setting. During this final presentation, students' presentation performance will be assessed by adopting a validated rubric (Van Ginkel, Laurentzen, Mulder, Mononen, Kyttä, & Kortelainen, 2017). Additionally, students have to complete additional questionnaires to assess students' sensitivity towards the feedback source and perceptions towards the (1) development of presentation skills, (2) VR environment and use of VR and (3) the value of computer-mediated feedback.

For data analysis, MANCOVA will be carried out to determine potential differences in impact, between the experimental condition and the control condition, on students' presentation skills and perceptions towards the VR-technology, value of feedback and sensitivity towards the feedback source. Further, students' traits are adopted as co-variates in this experiment.

Acknowledgements (optional)

References

Adubra, E., Da Silva, I., Dhungana, B., Mohan, N., Saltsman, G., & Van Ginkel, S. (2019). Exploring the use of Artificial Intelligence to support teachers and teacher development. UNESCO International Task Force on Teachers, 1(19), 1-16.

Belboukhaddaoui, I. & Van Ginkel, S. (2019). Fostering oral presentation skills by the timing of feedback: an exploratory study in virtual reality. Research on Education and Media, 11(1), 25-31. doi: 10.2478/rem-2019-0005.

Boetje, J. & Van Ginkel, S. (2020). The added benefit of an extra practice session in virtual reality on the development of presentation skills: A randomized control trial. Journal of Computer Assisted Learning, 1-12. doi: 10.1111/jcal.12484.

De Grez, L. (2009). Optimizing the instructional environment to learn presentation skills (Dissertation). Universiteit Gent.

De Grez, L., Valcke, M., & Roozen, I. (2009). The impact of an innovative instructional intervention on the acquisition of oral presentation skills in higher education. Computers & Education, 53(1), 112–120.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. Computers & Education, 70, 29–40.

Parmigiani, D., Van Ginkel, S., Saltsman, G., & Dhungana, B. (2020). Innovation in Teacher Education. Italian Journal of Educational Technology. doi: 10.17471/2499-4324/1187

Van Ginkel, S., Ruiz, D., Mononen, A., Karaman, C, De Keijzer, A., & Sitthiworachart, J. (2020). The impact of computer-mediated immediate feedback on developing oral presentation skills: an exploratory study in virtual reality. Journal of Computer Assisted Learning, 412-422. doi: 10.1111/jcal.12424.

Van Ginkel, S., Gulikers, J., Biemans, H., Noroozi, O., Roozen, M., Bos, T., Van Tilborg, R., Van Halteren, M., & Mulder, M. (2019). Fostering oral presentation competence through a virtual reality-based task for delivering feedback. Computers & Education, 134, 78-97. https://doi.org/10.1016/j.compedu.2019.02.006.

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2017a). The impact of the feedback source on developing oral presentation competence. Studies in Higher Education, 42(9), 1671-1685. doi:10.1080/03075079.2015.1117064.

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2017b). Fostering oral presentation performance: does the quality of feedback differ when provided by the teacher, peers or peers guided by tutor? Assessment and Evaluation in Higher Education, 42(6), 953-966. doi: 10.1080/02602938.2016.1212984.

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2015). Towards a set of design principles for developing oral presentation competence: A synthesis of research in higher education. Educational Research Review, 14, 62-80. doi:10.1016/j.edurev.2015.02.002.

Van Ginkel, S., Laurentzen, R., Mulder, M., Mononen, A., Kyttä, J., & Kortelainen, M. J. (2017). Assessing oral presentation performance: designing a rubric and testing its validity with an expert group. Journal of Applied Research in Higher Education, 9(3), 474-486

DESIGNING DATA GOVERNANCE MECHANISMS FOR DATA MARKETPLACE META-PLATFORMS

ANTRAGAMA EWA ABBAS

Delft University of Technology, Faculty of Technology, Policy and Management, Delft, The Netherlands; e-mail: a.e.abbas@tudelft.nl

Abstract Data Marketplace Meta-platforms (DMMPs) federate the fragmented set of data marketplaces and are expected to become a pivotal instrument to realize a single European Data Market in 2030. However, one critical hindrance to foster the adoption of business data sharing via DMMPs is data providers' risk of losing control over data. Generally, the literature on inter-organizational data sharing has highlighted that data governance mechanisms can help data providers to retain control over their data. Nevertheless, data governance mechanisms in the DMMP context are yet to be explored. Therefore, this research aims to design data governance mechanisms for business data sharing in DMMPs by employing the Design Science Research (DSR) approach. This study contributes to the literature by identifying root causes and consequences of losing control over data and defining prescriptive knowledge regarding design requirements, design principles, and a framework for designing data governance mechanisms within the novel setting of meta-platforms.

1   Introduction

Data marketplaces are increasingly recognized as a pivotal instrument for accomplishing the EU vision to create a single European Data Market in 2030 (European Commission, 2020). A data marketplace is a multi-sided platform that matches data providers and data buyers; that facilitates data sharing and transactions via features provided by data marketplace owners and third-party providers (TPPs) (Fruhwirth, Rachinger, & Prlja, 2020; Koutroumpis, Leiponen, & Thomas, 2020; Spiekermann, 2019). Data marketplaces' core aim is to facilitate business data sharing among companies (Agahari, 2020). Thereby, business data become a trading commodity. Nevertheless, considerable heterogeneity of data marketplace initiatives exists and causes fragmentation. The fragmentation causes multiple aspects of data marketplaces (e.g., business models, governance arrangements, and technical standards) to diverge uncontrollably, leading to a decrease of trust in data marketplaces as a whole (TRUSTS, 2019). For potential data buyers, for instance, the fragmentation triggers difficulties in data discovery processes. Data buyers also suffer from vendor lock-in (i.e., unable to switch to other data marketplace providers due to high switching costs). In general, the fragmentation has slowed down the platforms' commercialization due to a lack of users (i.e., data providers and buyers) (Basaure, Vesselkov, & Töyli, 2020).

Nascimbeni (2020) refers to meta-platforms as a promising solution to tackle fragmentation. A meta-platform is a platform of platforms that coordinates and integrates different platforms' resources and solutions. (Billhardt et al., 2020; Burkhardt, Frey, Hiller, Neff, & Lasi, 2019; Savković, Schweigkofler, Savković, Riedl, & Matt, 2020). Meta-platforms are centralized efforts to organize collective actions by enforcing common policies, standards, and infrastructures (Chia, Keogh, Leorke, & Nicoll, 2020; Floetgen et al., 2021). Meta-platforms functionalities include *one-stop-shop* via standardized portals, information dissemination & aggregation, and the establishment of shared services (Floetgen et al., 2021; Hoffmann, Rupp, & Sander, 2020). Meta-platforms enable the increase of demand-side users (e.g., data providers and data sellers) to discover data, avoid switching costs and demonstrate legal compliance (Basaure et al., 2020).

However, a critical hindrance to foster the adoption of business data sharing via data marketplaces is data providers' risk of losing control over data (Richter & Slowinski, 2019; Spiekermann, 2019). Losing control over data triggers many consequences for data providers. For instance, competitors may benefit from their data in unanticipated ways (Gelhaar & Otto, 2020; Richter & Slowinski, 2019). Moreover, it also brings privacy risks (Schomakers, Lidynia, & Ziefle, 2020) and triggers data providers' reputational damage. In our exact context, i.e., Data Marketplace Meta-Platforms (DMMPs), we argue that data providers also possess the same risk. DMMPs inherit unresolved data control problems in data marketplaces. Even more, the nature of DMMPs where data flows from a data marketplaces to others (and vice versa) may increase the risk.

The literature on inter-organizational data sharing has highlighted that data governance can potentially help data providers to retain control over their data (van den Broek & van Veenstra, 2015). We define data governance as the activities of exercising control (i.e., defining what, who, and how) over data ownership, access, and data usage decisions to minimize the risks associated with data sharing (De Prieëlle, De Reuver, & Rezaei, 2020; Lis & Otto, 2020; Nokkala, Salmela, & Toivonen, 2019). Examples of data governance mechanisms are defining data ownership and access, formal contract selling, user consent, or data stewards. These mechanisms are beneficial to overcome the barrier of losing control over data (Günther, Rezazade Mehrizi, Huysman, & Feldberg, 2017; Lee, Zhu, & Jeffery, 2017; Suver, Thorogood, Doerr, Wilbanks, & Knoppers, 2020). Nevertheless, data governance mechanisms in the DMMP context are yet to be explored. Based on the previous elaboration, the objective of this study is **to design data governance mechanisms for business data sharing via DMMPs to reduce data providers' risk of losing control over data.**

2   Research Design

This study will employ the Design Science Research (DSR) approach. The DSR aims to develop innovative Information System (IS) artifacts to solve real-world problems (Hevner, 2007). The DSR approach is appropriate because it allows the creation of innovative artifacts (i.e., data governance mechanisms) that could reduce data providers' risk of losing control over data. The creation of artifacts and also it implementations can lead to knowledge creation. Based on the study of Peffers, Tuunanen, and Niehaves (2018) that defines DSR genres, we position our DSR study under the classification of the *Design Science Research*

Methodology (DSRM). **The DSRM is a well-adopted methodology in DSR. The DSRM focuses on artifact development, and its evaluation is outcome-oriented and practical. It should aim for generalizability and reasoned arguments on why the designed artifacts could work (Peffers et al., 2018). Building from the DSRM (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007), we will explain the relationship between research phases, questions, and instruments. The research will be conducted within five phases (see** Figure).
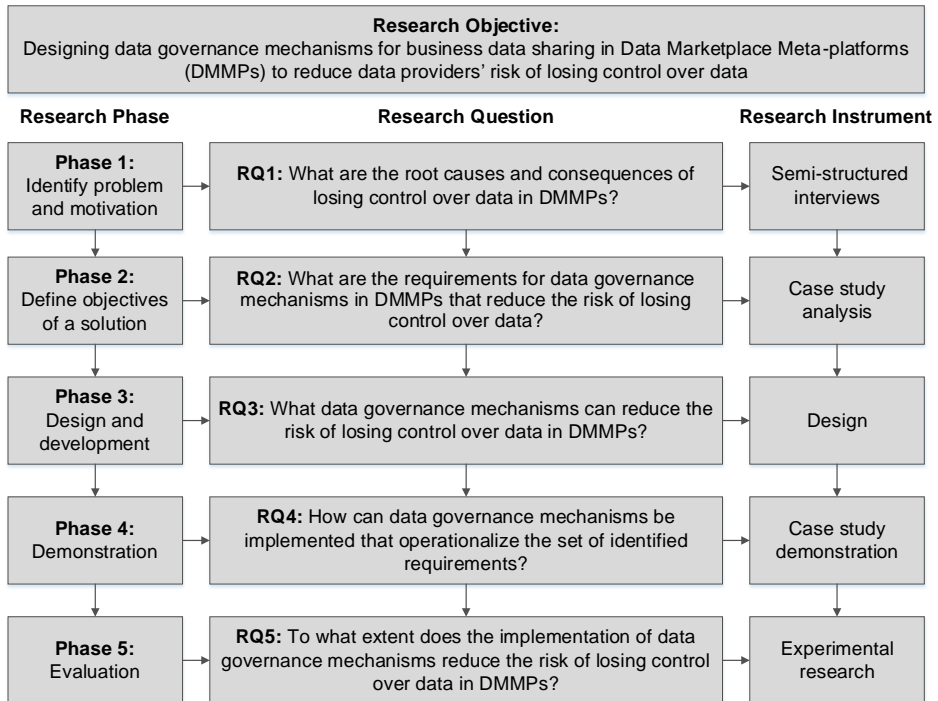
Figure 1: Research design

We will begin the research by identifying the root causes and consequences of losing control over data in the context of DMMPs. These factors will be the foundation when defining requirements in the later stage. We will conduct an exploratory study to answer the RQ1. An exploratory study is suitable because not much is known about this phenomenon (Sekaran & Bougie, 2016). Our exploratory study will employ the qualitative approach using semi-structured interviews as a primary data collecting method. Semi-structured interviews will be guided by the preliminary literature review result we conducted in the introduction section. Next, the second

phase of our DSR study defines requirements for designing data governance mechanisms in DMMPs. Requirements here refer to the detailed descriptions of what users want from the designed solution (Dym, Little, & Orwin, 2014). We will use a single case study to derive data governance requirements. We will select the TRUSTS project[1] since that project aims to develop a DMMP. We will use several procedures, such as interviews, document analysis, and observation, to collect the data. We are aware that it will be challenging to come up with one final design. Therefore, we will create business data sharing scenarios via DMMPs to scope our research.

Once we have defined the requirements, we will answer the RQ3 by designing the artifacts. We will produce two major artifact types in this phase, namely 1) design principles and 2) a framework. First, we will construct design principles, which refer to "core principles and concepts to guide design" (Vaishnavi et al., 2004, p. 16). Design principles inform the designers to develop instances from artifacts that belong to a similar class (Gregor, Kruse, & Seidel, 2020). Second, we will develop a framework to design data governance mechanisms. The framework here can be described as "real or conceptual guides to serve as support…" (Vaishnavi et al., 2004, p. 16). DMMPs owners can use the framework to develop data governance mechanisms. The design principles will be used as input to create this framework. The framework includes the guideline to decide and select which mechanisms will be appropriate; to operationalize the mechanisms in a specific context.

In phase 4, we will apply the designed framework to build data governance mechanisms. Phase 4 will be beneficial to demonstrate data governance mechanisms in practices. We will use a case study (i.e., the TRUSTS project) to operationalize our solution. We will follow the frameworks from Phase 3 to develop an instantiation. In the final phase, we will evaluate whether demonstrable data governance mechanisms can achieve the goal or not. We will evaluate our data governance mechanisms using a summative evaluation. The summative evaluation aims to test whether our designed artifact creates the desired impact or not. Summative evaluation is often conducted after the artifact has been developed (i.e., *ex-post* evaluation) (Sonnenberg & Vom Brocke, 2012; Venable, Pries-Heje, & Baskerville,

---

[1] https://www.trusts-data.eu/ accessed on March 02, 2021

2016). We will conduct an experiment to test our designed artifacts (e.g., quasi-experiments).

3   Theoretical framework

The section below discusses relevant theories that will be used to develop data governance mechanisms.

Data governance theory – data governance can help organizations to retain control over their data (van den Broek & van Veenstra, 2015). Data governance mechanisms such as defining data ownership and access, RACI chart, formal contract selling, user consent, governance mode, and data stewards can be used as instruments to retain control over data (Günther et al., 2017; Lee et al., 2017; Suver et al., 2020). We will build upon these works and will explore which mechanisms are appropriate for our novel settings.

Platform control theory – platform control can be described as platform owner's attempts to influence complementors (e.g., application developers) to behave according to the platform's objectives (Goldbach, Benlian, & Buxmann, 2018; Tiwana, 2013). There are two categories of platform control mechanisms: a) formal and b) informal control. The formal control can be further subdivided into input, behavior, and output control. In contrast, informal control can be subcategorized into self- and clan control. Previous studies show how platform controls affect digital platforms. For example, Goldbach et al. (2018) discover that enforcing self-control to third-party developers positively influences the quality of the application and their continuance intention to participate in an ecosystem. In addition, Zheng, Xu, and Lin (2019) reveal that formal control (e.g., seller reputation) and social control (e.g., number of fans and members), in the context of the crowdsourcing platform, are decreasing the likelihood of opportunistic behavior by the seller.

Coordination theory – coordination is the act of managing dependencies amongst activities to accomplish a goal. The coordination consists of four components: the *goals, activities, actors,* and *dependencies.* The common coordination processes associated with those components include: a) defining goals, b) plotting goals to activities, c) defining actors and assigning them to relevant activities, and d) managing dependencies (Malone & Crowston, 1990, 1994). Though studies that draw the

connection between coordination and control are limited (Ko, Lee, Keil, & Xia, 2019), some researchers have started investigating this matter. For example, in the project management context, Remus, Wiener, Mähring, Saunders, and Cram (2015) inform that control-focus on coordination can be achieved by "empowerment, guidance, facilitation, and trust" (p. 5). Ko et al. (2019) explore the different impacts of formal control (i.e., outcome and process control) on coordination and explain how coordination mediates these formal controls. The coordination theory will be beneficial to complement the previous platform control theory. For instance, if we want to employ the clan control that requires interactions between complementors, we need to define what appropriate activities should be proposed, which dependencies may occur, and how to assemble those dependencies.

Accountability theory – Vance, Lowry, and Eggett (2015) have synthesized the definition of *accountability*: "a process in which a person has a potential obligation to explain his/her actions to another party who has the right to pass judgment on those actions and to administer potential positive or negative consequences in response to them" (p. 347). Accountability theory suggests numerous mechanisms that increase accountability perceptions. These are: *a) identifiability*, *b) expectation of evaluation, c) awareness of monitoring,* and *d) social presence.* Various studies have revealed the relationship between accountability and control. For example, accountability theory can be used to increase accountability perception, and in consequence, decrease the attention to violate the data access policy (Vance, Lowry, & Eggett, 2013; Vance et al., 2015). Yaokumah, Walker, and Kumah (2019) explain how Security Education, Training, and Awareness programs (SETA) can improve employee security behavior, mediated by employee accountability. A study conducted by Y. Zheng, Huang, Lee, and Bao (2017) shows that extra-role behaviors (derived from social control and accountability theory) positively influence such bright internet policy adaptions. In general, individuals with a high accountability perception are more likely to develop cognitive awareness. Accordingly, it will lead to the conformity of expected actions, pro-social behaviors, and decrease risk behaviors (Zhang, Wei, & Zeng, 2020).

The summary of the context, relevant theories, and DSR can be seen in Figure 2. It describes the idea of having kernel and provides justificatory knowledge for prescriptive knowledge (Kuechler & Vaishnavi, 2008). We will use multiple theories to gather justificatory knowledge. Finally, we will also use those theories to derive criteria on the utility of the artifact (Niehaves & Ortbach, 2016).
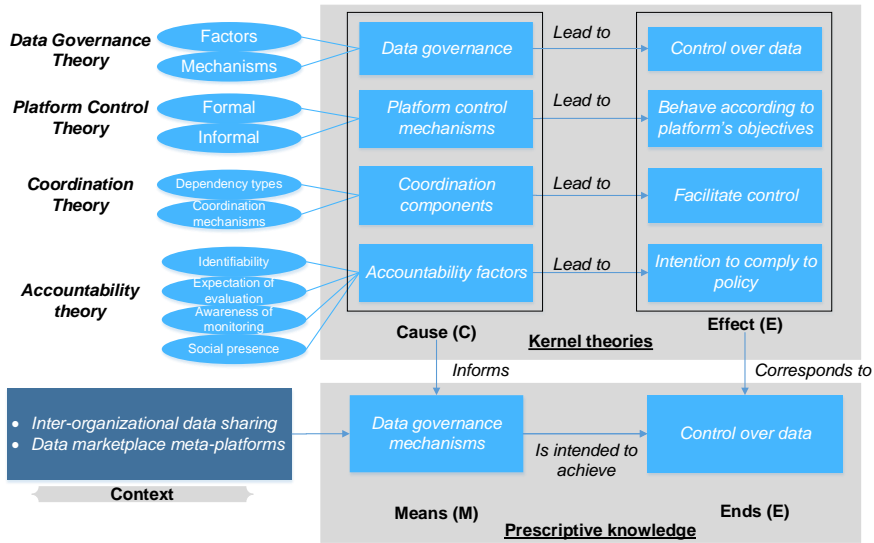

Figure 2: The role of theories in the DSR study

4 Preliminary results

A Systematic Literature Review (SLR) study for business data sharing via data marketplace has been conducted. This study is currently under review at an international conference. This study helps to set the stage and position our proposed research in data marketplace literature. Therefore, we will use it as an input for the literature section. After reviewing numerous articles, we found no comprehensive overview of data marketplace research available in the literature. Consequently, we have no clear understanding of what is known about data marketplaces, and we are unable to spot neglected research topics that may contribute to advancing data marketplaces towards commercialization. This study provides an overview of the state of the art of data marketplace research. We employ the SLR approach and structure our analysis using the Service-Technology-Organization-Finance (STOF) model. We studied 135 articles from the Scopus database and found that the extant

data marketplace literature is primarily dominated by technical research, such as discussions about *computational pricing* and *architecture*. To move past the first stage of the platform's lifecycle, i.e., platform design, to the second stage, i.e., platform adoption, we call for empirical research in non-technological areas, such as *value networks* and *organizational arrangement*. The findings, therefore, in line with our goal (i.e., contribute to a non-technical topic and incorporate empirical data) because the proposed research will design data governance mechanisms that likely discuss the interaction between actors and required arrangements in DMMPs.

Moreover, we conducted an initial workshop with TRUSTS internal participants on 22 October 2020 to explore potential value creations and perceived adoption barriers of DMMPs. In total, 15 participants from different organizations attended the workshop. The former objective is relevant to this study because it strengthens the argument we elaborate in the introduction section, i.e., related to the benefit of DMMPs towards resolving the fragmentation. The latter objective is also beneficial to give additional insights for exploring the root causes and impacts of losing control over data. On the one hand, potential value creations of DMMPs include 1) traffic forwarding and commissioned brokerage services, 2) the increase of dataset/data source numbers within the federation, 3) gradual harmonization of technology stack through coordination and common standards, and 5) provision of non-differentiating capabilities (e.g., billing) as shared services. On the other hand, perceived adoption barriers of DMMPs include 1) unclear and unproven value propositions, 2) unclear and unproven network-effects and added value from small-size and domain-specific data marketplaces, 3) unexplored economics of various data marketplace setups with a federation (e.g., revenue sharing mechanisms), 4) increased complexity and cost for technology integration, and 5) fear of losing data marketplaces' Unique Selling Proposition (USP).

In addition, we also developed a business model taxonomy for data marketplaces. The developed taxonomy is relevant because we need to know what kind of data marketplaces a DMMP platform will federate or interoperable with. The study is also beneficial in the literature review section of the dissertation. In this study, we argued that data marketplaces are vastly different so that a taxonomy can be used to classify data marketplaces. Existing data government arrangements that the DMMP federates or interoperable with may be incompatible. For instance, a data marketplace may focus on automotive data while the other focuses on healthcare

data. The degree of sensitivity of these data types are different, and in consequence, it may impact the definition of data access right. We employed a DSR approach and a standard taxonomy development method by Nickerson, Varshney, and Muntermann (2013) to develop the taxonomy. Four meta-dimensions, 17 business model dimensions, and 59 business model characteristics have been identified in the final taxonomy.

5 Future development and expected contributions

The future development of this research is to execute the plan provided in section 2. These are: 1) identifying root causes and impacts of losing control over data, 2) identifying requirements for data governance mechanisms, 3) creating a framework for designing data governance mechanisms, 4) demonstrating data governance mechanisms, and 5) evaluating the usefulness of the developed mechanisms.

This study contributes to science by identifying root causes and impacts of losing control over data in DMMPs, thus serving as a basis for designing solutions. We also contribute to defining prescriptive knowledge regarding: a) design requirements, b) design principles, and c) a framework for designing data governance mechanisms in DMMPs. Societally, DMMPs' owners may benefit from this research by applying data governance mechanisms to reduce data providers' risk of losing control over data. Data providers will feel safe and trust the ecosystem because of their positive perception. Consequently, data providers' adoption of DMMPs may increase, potentially leading to more value generation through business data sharing and use.

References

Agahari, W. (2020). Platformization of data sharing: Multi-party computation (MPC) as control mechanism and its effect on firms' participation in data sharing via data marketplaces. Paper presented at the 33rd Bled eCONFERENCE.

Basaure, A., Vesselkov, A., & Töyli, J. (2020). Internet of things (IoT) platform competition: Consumer switching versus provider multihoming. Technovation, 90, 102101.

Billhardt, H., Santos, J.-A., Fernández, A., Moreno-Rebato, M., Ossowski, S., & Rodríguez-García, J. A. (2020). Legal Implications of Novel Taxi Assignment Strategies. In (pp. 361-372): Springer International Publishing.

Burkhardt, D., Frey, P., Hiller, S., Neff, A., & Lasi, H. (2019). Distributed Ledger Enabled Internet of Things Platforms: Symbiosis Evaluation. In (pp. 77-118): Springer International Publishing.

Chia, A., Keogh, B., Leorke, D., & Nicoll, B. (2020). Platformisation in game development. Internet Policy Review, 9(4), 1-28.

De Prieëlle, F., De Reuver, M., & Rezaei, J. (2020). The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. IEEE Transactions on Engineering Management.

Dym, C. L., Little, P., & Orwin, E. J. (2014). Engineering design: a project-based introduction.

European Commission. (2020). Work stream on Data [Press release]. Retrieved from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68354

Floetgen, R. J., Strauss, J., Weking, J., Hein, A., Urmetzer, F., Böhm, M., & Krcmar, H. (2021). Introducing platform ecosystem resilience: leveraging mobility platforms and their ecosystems for the new normal during COVID-19. European Journal of Information Systems, 1-18. doi:10.1080/0960085x.2021.1884009

Fruhwirth, M., Rachinger, M., & Prlja, E. (2020). Discovering Business Models of Data Marketplaces. Paper presented at the Proceedings of the 53rd Hawaii International Conference on System Sciences.

Gelhaar, J., & Otto, B. (2020). Challenges in the Emergence of Data Ecosystems. Paper presented at the PACIS.

Goldbach, T., Benlian, A., & Buxmann, P. (2018). Differential effects of formal and self-control in mobile platform ecosystems: Multi-method findings on third-party developers' continuance intentions and application quality. Information & Management, 55(3), 271-284.

Gregor, S., Kruse, L. C., & Seidel, S. (2020). The Anatomy of a Design Principle. Journal of the Association for Information Systems.

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems, 26(3), 191-209. doi:10.1016/j.jsis.2017.07.003

Hevner, A. R. (2007). A three cycle view of design science research. Scandinavian journal of information systems, 19(2), 4.

Hoffmann, E., Rupp, J., & Sander, K. (2020). What Do Users Expect from Climate Adaptation Services? Developing an Information Platform Based on User Surveys. In (pp. 105-134): Springer International Publishing.

Ko, D.-G., Lee, G., Keil, M., & Xia, W. (2019). Project Control, Coordination, and Performance in Complex Information Systems Outsourcing. Journal of Computer Information Systems, 1-11. doi:10.1080/08874417.2019.1606687

Koutroumpis, P., Leiponen, A., & Thomas, L. D. W. (2020). Markets for data. Industrial and Corporate Change, 29(3), 645-660. doi:10.1093/icc/dtaa002

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: anatomy of a research project. European Journal of Information Systems, 17(5), 489-504.

Lee, S. U., Zhu, L., & Jeffery, R. (2017). Data governance for platform ecosystems: Critical factors and the state of practice. arXiv preprint arXiv:1705.03509.

Lis, D., & Otto, B. (2020). Data Governance in Data Ecosystems–Insights from Organizations.

Malone, T. W., & Crowston, K. (1990, 1990). What is coordination theory and how can it help design cooperative work systems? Paper presented at the Proceedings of the 1990 ACM Conference on Computer-Supported Cooperative Work, CSCW 1990, Los Angeles.

Malone, T. W., & Crowston, K. (1994). The interdisciplinary study of coordination. ACM Computing Surveys, 26(1), 87-119. doi:10.1145/174666.174668

Nascimbeni, F. (2020). Open Educational Resources in Italy. In (pp. 49-62): Springer Singapore.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. European Journal of Information Systems, 22(3), 336-359. doi:10.1057/ejis.2012.26

Niehaves, B., & Ortbach, K. (2016). The inner and the outer model in explanatory design theory: the case of designing electronic feedback systems. European Journal of Information Systems, 25(4), 303-316. doi:10.1057/ejis.2016.3

Nokkala, T., Salmela, H., & Toivonen, J. (2019). Data Governance in Digital Platforms. Paper presented at the AMCIS.

Peffers, K., Tuunanen, T., & Niehaves, B. (2018). Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. In: Taylor & Francis.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of management information systems, 24(3), 45-77.

Remus, U., Wiener, M., Mähring, M., Saunders, C., & Cram, W. A. (2015). Why do you control? The concept of control purpose and its implications for IS project control research.

Richter, H., & Slowinski, P. R. (2019). The Data Sharing Economy: On the Emergence of New Intermediaries. IIC - International Review of Intellectual Property and Competition Law, 50(1), 4-29. doi:10.1007/s40319-018-00777-7

Savković, B. V., Schweigkofler, A., Savković, O., Riedl, M., & Matt, D. T. (2020). Validation Methodology for a Citizen-centric Smart-City Platform. Paper presented at the ISPIM Conference Proceedings.

Schomakers, E.-M., Lidynia, C., & Ziefle, M. (2020). All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. Electronic Markets, 1-17.

Sekaran, U., & Bougie, R. (2016). Research methods for business: A skill building approach: John Wiley & Sons.

Sonnenberg, C., & Vom Brocke, J. (2012). Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research. Paper presented at the International Conference on Design Science Research in Information Systems.

Spiekermann, M. (2019). Data Marketplaces: Trends and Monetisation of Data Goods. Intereconomics, 54(4), 208-216. doi:10.1007/s10272-019-0826-z

Suver, C., Thorogood, A., Doerr, M., Wilbanks, J., & Knoppers, B. (2020). Bringing Code to Data: Do Not Forget Governance. Journal of Medical Internet Research, 22(7), e18087. doi:10.2196/18087

Tiwana, A. (2013). Platform ecosystems: Aligning architecture, governance, and strategy: Newnes.

TRUSTS. (2019). Trusted Secure Data Sharing Space. In. European Union: European Union's Horizon 2020 (No. 871481).

Vaishnavi, V., Kuechler, W., & Petter, S. (2004). Design science research in information systems. Retrieved from http://www.desrist.org/design-research-in-information-systems/

van den Broek, T., & van Veenstra, A. F. (2015). Modes of governance in inter-organizational data collaborations.

Vance, A., Lowry, P. B., & Eggett, D. (2013). Using accountability to reduce access policy violations in information systems. Journal of management information systems, 29(4), 263-290.

Vance, A., Lowry, P. B., & Eggett, D. L. (2015). Increasing accountability through the user interface design artifacts: A new approach to addressing the problem of access-policy violations. MIS quarterly, 39(2), 345-366.

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. European Journal of Information Systems, 25(1), 77-89.

Yaokumah, W., Walker, D., & Kumah, P. (2019). SETA and Security Behavior - Mediating Role of Employee Relations, Monitoring, and Accountability. Journal of Global Information Management, 27, 102-121. doi:10.4018/JGIM.2019040106

Zhang, M., Wei, X., & Zeng, D. D. (2020). A matter of reevaluation: Incentivizing users to contribute reviews in online platforms. Decision Support Systems, 128, 113158. doi:10.1016/j.dss.2019.113158

Zheng, H., Xu, B., & Lin, Z. (2019). Seller's creditworthiness in the online service market: A study from the control perspective. Decision Support Systems, 127, 113118. doi:10.1016/j.dss.2019.113118

Monitoring Remote Service Platforms Using Artificial Intelligence-Based Distributed Intrusion Detection

Thorsten Weber[1] & Rüdiger Buchkremer[2]

[1] UCAM Universidad Católica San Antonio de Murcia, Spain; e-mail: thorsten.weber@fom-net.de
[2] Institute for IT Management and Digitization, FOM University of Applied Sciences, Düsseldorf, Germany; e-mail: ruediger.buchkremer@fom-net.de

Abstract Due to their flexibility, remote support platforms are ideal for contributing to companies' digital strategy. Simultaneously, this flexibility of use cases makes it difficult to reliably detect attacks on the network infrastructure. This paper presents a proposal for the detection of fraud patterns on remote service platforms through artificial intelligence. A blockchain-based approach will be used to adapt these attack signatures to the specific use cases of remote service platform users. By employing a blockchain-based attack signature selection mechanism, remote service platform users will be able to adjust the attack signatures flexibly and in a tamper-proof manner.

1   Introduction

Effective plant reliability is of utmost importance for manufacturing and other industrial pursuits. Due to industrial plants' high-profile nature, unplanned downtime events can easily result in extraordinary costs (Christer & Waller, 1984). The causes of such breakdowns are numerous, and troubleshooting is typically performed by engineers or experienced technicians (Hiltunen et al., 2008). To ensure the lowest possible downtime, a company must have suitable service technicians as soon as possible on-site and available. Due to a plant's complexity, deploying an emergency service for troubleshooting can quickly turn into a planning problems (Vossing, 2017); digitization may improve the planning process's accuracy.

Remote service platforms (RSPs) are digital solutions that help companies better plan service deployment in plants. Companies can implement RSPs to train and educate workers remotely on new machines, plants, or systems. Analog monitoring processes, such as maintenance, quality assurance, and auditing, can be performed remotely, as well (Werner & Bechini, 2019). Moreover, RSPs allow remote guidance of workers and transmission of instructions. Service technicians and engineers can use RSPs to transmit real-time advice to on-site workers and repair problems from a distance without traveling. This results in less downtime and thus to a faster re-start of production after an incident.

This digitization of analog processes causes additional economic side-effects on companies. On the one hand is the direct saving of travel costs (e.g., costs for cars, flights, trains, cabs, and hotel accommodation). On the other hand, companies can redeploy their service technicians much more quickly. Service technicians must no longer "waste" time traveling and can be deployed more frequently in the same time frame. Last, saving on travel impacts a company's carbon dioxide (CO_2) footprint and can be a competitive advantage.

In summary, RSPs offer many benefits to companies. They can ensure that a service technician can quickly get to where they are needed, even if that technician might not be able or allowed to travel.

T. Weber & R. Buchkremer:
Monitoring Remote Service Platforms Using Artificial Intelligence-Based Distributed Intrusion Detection

707

1.1   RSP Architecture

Figure illustrates a generic approach for an RSP. The architecture typically consists of three main components (Yin et al., 2006). On the one hand, it is an individual exchange and management platform to which both the service technician and the customer have access via the Hypertext Transfer Protocol Secure (HTTPS). The platform management server enables two or more participants to communicate and exchange data with each other. The management functionalities refer to access control and user management.

On the other hand, a client-side application allows users to connect to the central platform management server. Typically, these are desktop, browser, or smartphone/tablet/smart glasses applications. In a basic configuration of the communications infrastructure, two or more participants communicate via peer-to-peer (P2P)networks (Ripeanu, 2001), using the Web Real-Time Communication Protocol (Johnston et al., 2013). If a P2P connection is not feasible for technical reasons, participants switch to alternate settings (Mahy et al., 2010).
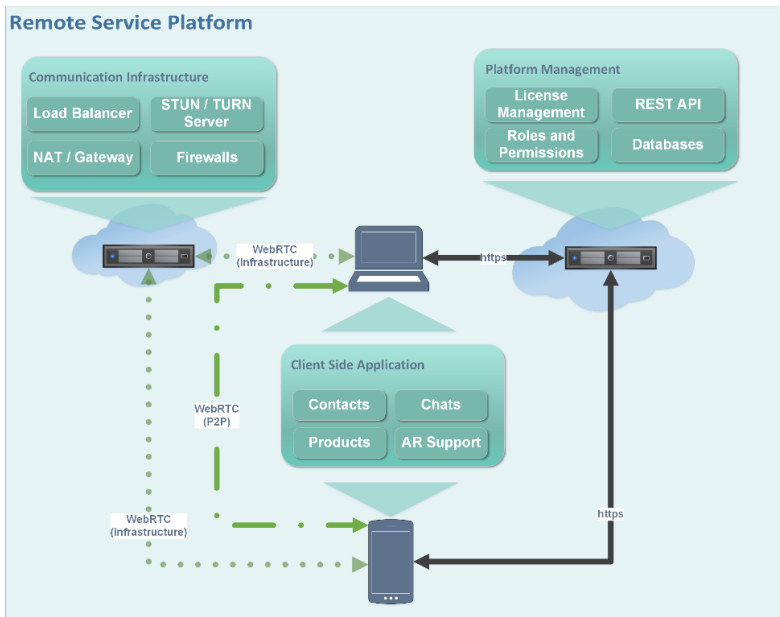

Figure 1: Typical RSP components in a nutshell.

1.2   RSP Security

An essential prerequisite for the successful implementation of RSPs is, in addition to pure functionality, confidence in the security settings of the platform's network (i.e., confidence in its security goals: confidentiality, integrity, and availability [CIA]) (Can & Sahingoz, 2015). Security assets are often critical for selecting software packages (Academy et al., 2007) and is often assumed to be naturally given (Sahay & Gupta, 2003) by a software provider.

Basis security measures of RSPs can be achieved by applying state-of-the-art security protocols, such as HTTPS or other authentication mechanisms (Kiraz, 2016). However, this basic security is not always appropriate, and advanced security mechanisms are needed. For example, the primary security mechanisms do not include protection against network-based attacks and do not allow monitoring whether a system has been exploited or tampered (Brown & Heikki, 2005; Jatti & Kishor Sontif, 2019; Liao et al., 2013). Some authors recommend implementing network intrusion detection systems (NIDSs) as the first choice for detecting network-based attacks (Debar et al., 2000; El-Bakry & Mastorakis, 2008).

The idea of intrusion detection systems (IDSs) was described in 1987 by Denning (Denning, 1987). Henceforth, the topics of IDSs were well researched by the scientific community (Khraisat et al., 2019). Today, there is a specialization trend in those systems, such as for wireless sensor networks (Can & Sahingoz, 2015), the Internet of Things (Zarpelão et al., 2017), smart grids (Jow et al., 2017), and cloud computing (Chiba et al., 2016). Specialization has the advantage that systems' unique characteristics can be considered. It is conceivable that an IDS designed for Internet of Things applications could have significantly higher requirements in terms of power consumption than, for example, an IDS developed for cloud systems.

On the other hand, many RSPs require security measures. Unauthorized platform access, attacks on communication infrastructure, and unauthorized use of premium services are only a few potential threat scenarios that could reduce confidence in RSPs. For these reasons, it is logical and consequent to develop an IDS tailored to RSPs.

T. Weber & R. Buchkremer:
Monitoring Remote Service Platforms Using Artificial Intelligence-Based Distributed Intrusion Detection

709

2 Problem Definition

In general, the implementation of an IDS for RSPs requires attention to three main aspects. These are the mathematical requirements, the challenges for tailoring an IDS for RSPs, and the possibility of customizing and notarizing the selected configuration on the customer's part.

2.1 Mathematical Boundaries

A significant problem encountered by IDSs is the so-called base rate fallacy (Axelsson, 2000), a statistical error that may occur when determining conditional probabilities. This problem can be easily explained by applying Bayes' theorem.

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A)\Pr(A) + \Pr(B|\neg A)\Pr(\neg A)}$$

Assuming that 1% ($\Pr(\neg A)$) of traffic constitutes "bad traffic," such as a synchronize message flood (SYN flood), while 99% ($\Pr(A)$) constitutes a valid connection. The IDE detection rate is 90% ($\Pr(B|\neg A)$), and the false alarm rate is 10% ($\Pr(B|A)$).

Research question: What is the conditional probability that a connection marked by the IDS as an SYN flood is valid? What is the conditional probability that traffic is valid under the condition that the IDS triggers an alarm? Using the values mentioned above in Bayes' theorem yields the following:

$$\frac{0.10 \cdot 0.99}{10 \cdot 0.99 + 0.90 \cdot 0.01} \approx 92\%$$

Thus, if an alarm triggers the IDS, the probability is around 92% that it is a false alarm, which is an extremely high value. Ultimately, a high value can result in employees ignoring the alarm, leading to current attacks being ignored.

2.2   Tailoring Intrusion Detection to RSPs

According to (Liao et al., 2013), IDSs are divided into signature-based, anomaly-based, and specification-based systems. Signature-based and specification-based systems belong to the knowledge-based systems, while anomaly-based systems belong to the behavior-based systems. Anomaly-based IDSs detect typical user behavior and network connections; if the behavior deviates from this pattern, anomaly-based IDSs react accordingly.

Currently, there exists a trend towards specialization when developing an IDS. Research gap: a scientific approach that handles specific requirements of an IDS in the environment of RSPs is missing. The challenge for defining an IDS for RSP is the broadness of the RSP use cases, such as remote training (Masoni et al., 2017), remote audits (Teeter et al., 2010), and remote assembly (Elvezio et al., 2017). One challenge for an IDS is the ability to be adapted as flexibly as possible to existing and future RSP use cases and at the same time meet all users' data protection requirements.

To be more precise, two artificial intelligence (AI) methods are needed. In the first step, the network traffic must be classified correctly. Using AI, received network traffic must be classified based on its properties. E.g., being HTTPS, ping, or another kind of traffic. clustering algorithms can do so (Liu et al., 2008; Münz et al., 2007). There are two approaches in principle: Supervised and Unsupervised Learning Algorithms (Sathya & Abraham, 2013). Even though their differences have been analyzed in the past, in the use case of RSPs, a priori, it is not clear which method can be used most reliably to classify the network traffic in the use cases of RSPs.

In the next step, the classified network traffic must then be analyzed and predicted whether the examined network traffic is a possible attack. The prediction of an attack can be made in various ways, for example, by analyzing the transmitted packet information using text analysis algorithms (Min et al., 2018; Stone, 2007) or using regression (Altwaijry & Algarny, 2012; Wang, 2005). Again, a priori, it is not clear which method is best suited for predicting possible attacks on RSPs. It might also be the case that a hybrid solution might be most promising.

T. Weber & R. Buchkremer:
Monitoring Remote Service Platforms Using Artificial Intelligence-Based Distributed Intrusion Detection

711

2.3   Customization of IDSs

Typically, IDSs are configured utilizing policies (Bace & Mell, 2001). Based on the example presented in

Figure , the IDS would raise an alert containing the alert message "IP Package detected" if an IP packet from any source IP and Port would be sent to any destination IP and Port.
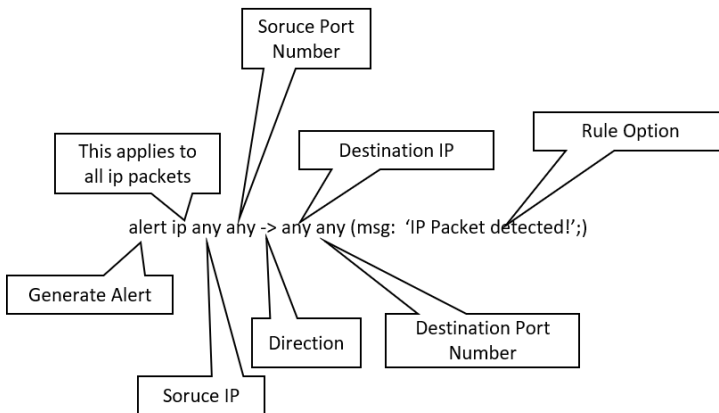
Figure 2: IDS Policy Example for Snort ("Snort 2.1 Intrusion Detect.," 2004)

Therefore, IDS policies can determine attack patterns and read off allowed network activities. This knowledge can serve as beneficial information for an attacker to plan an attack. Securing and configuring IDS policies are therefore crucial in terms of securing infrastructures. Consequently, RSP customers are interested in confining these policies independently and need a monitoring option for selected policies and attack signatures. As such, RSP customers also need a guarantee (notarized confirmation) that the RSP provider has indeed implemented the established IDS policy and attack signatures.

2.4   Resulting Research Questions

The central question to be answered by this dissertation project is as follows*: "Is it possible to develop a privacy-compliant and customizable artificial intelligence (AI)-based attack*

detection system for remote service platforms with the highest possible detection rate and lowest possible false-positive rate, optimization of data exchange, and an intuitive visualization and reaction to detected attacks?"

Further research questions (RQ) that this project includes are the following:

1. What are legal requirements for RSP's IDS?
2. What relevant intrusion detection system approaches already exist?
3. How can client-side applications be used to detect intrusions on RSPs?
4. How long does the learning phase of an AI-based IDS guarantee the greatest possible likelihood of attack detection?
5. How should a neural network be adjusted to distinguish between different application areas within the RSP?
6. How should the IDS react upon attack?

3 Methodology

The purpose of this chapter is to clearly outline *what* (implementation) is being done to solve each research question and *how* (means) it is being done. Moreover, this chapter addresses how the data is collected and what data can be accessed to answer the research questions.

RQs 1 and 2 serve as the basis for this dissertation, as they establish the research scope. Both research questions will be addressed via qualitative research or, to be more precise, by systematically reviewing the literature. RQ1 clarifies the legal framework in which this dissertation must operate to develop a legally secure and data- protection-compliant IDS for RSPs. The approach for solving RQs 1 and 2 is literature research, as described by vom Brocke et al. (Vom Brocke et al., 2009).

RQ 3 concerns the architecture of the software to be developed within the scope of this dissertation. The central task of the IDS is to detect attacks by examining deviations from normal behavior (Umer et al., 2017). Therefore, the IDS must receive status information of all entities being monitored. To guarantee error-free monitoring, this dissertation additionally must develop an architecture that monitors all entities reliably. Hence, RQ3 will be investigated through both qualitative and quantitative methods. A qualitative literature review must identify which IDS

architectural approaches already exist and which approaches should be considered when analyzing this research question.

On the other hand, quantitative experiments must collect and evaluate network load data and create attack signatures. An RSP typically has several connected devices, such as laptops, servers, smartphones, and smart glasses (Kao et al., 2014). By evaluating the network traffic, it is possible to check which approaches to architecture and communication with the IDS prove to be the most reliable in practice.

Since there are no reliable values for RQs 4 and 5, they must be investigated in an explorative study (Shields & Rangarjan, 2013). Therefore, a neural network will be created and trained over several periods in an attack-free test network. The attack vectors to be defined for this purpose will subsequently investigate whether the trained network recognizes attacks and how many it recognizes. Qualitative methods must be used to determine which training times are realistically achievable for actual companies (i.e., interviews with various stakeholders).

The final research question is highly individual, and it might not be possible to answer it in general terms. Instead, this dissertation aims to develop a set of recommendations based on a comprehensible presentation of various automated attack reactions. This is intended to present to users the possibilities of reacting to an attack and the consequences of these reactions.

4   Expected Results

On the one hand, this dissertation's expected results are a data protection compliant intrusion detection system that includes a set of attack signatures that are continuously improved by utilizing artificial intelligence. On the other hand, this dissertation expects to deliver a procedure allowing for the customer of the RSP to monitor the selected attack signatures and adjust them independently, if necessary.

4.1   Continuously Attack Signature Generation and Evaluation

To successfully detect an attack, the IDS must distinguish between "regular" and "attack" behaviors. Creating a continuous attack signature through log files utilization on both the client and server is, therefore, one expected outcome. Although through this dissertation, a substantial amount of actual RSP data will be available, this data will be (according to current knowledge) "attack-free." Another expected result of this dissertation is the creation of "attack" data and RSP-tailored attack signatures through penetration tests.

4.2   IDS Management via Blockchain

Besides, the goal is to develop a procedure that allows the customer of the RSP to monitor the selected attack signatures and adjust them independently, if necessary. A possible solution for this is the definition of the attack signature via blockchain. Blockchain can be used to establish a notarized definition of the selected attack signatures on the one hand and, on the other, be able to adjust the attack signature without the intervention of the platform operator.

5   Future Development

At the core of this dissertation, new attack patterns are created to detect attacks on RSPs. In particular, the data from the mobile devices that are part of the RSP will be used for this purpose. Generally, it can be assumed that client devices are mostly connected end-to-end encrypted—both with the management server and with each other for communication.

However, encrypted data packets can be examined for attack patterns to only a restricted level (Sherry et al., 2015). The data in the data packets can be analyzed for malicious content to only a limited extent. In the first step, a cloud infrastructure must be set up with which it is possible to simulate attacks on an RSP. The infrastructure must decrypt the devices' end-to-end encryption and forward the decrypted data packets to an IDS. Thus, the infrastructure must allow for the decryption of the network traffic, which must be analyzed. The decryption of encrypted network traffic can typically be achieved using a reverse proxy (Radivilova et al., 2018). In the second step, the newly built infrastructure must detect and

T. Weber & R. Buchkremer:
Monitoring Remote Service Platforms Using Artificial Intelligence-Based Distributed Intrusion Detection

715

classify new attack signatures. By performing targeted penetration tests, predefined attack patterns can be generated. Based on the performed penetration tests, the data packets analyzed by the IDS can then be stored as a new attack signature. For example, suppose a penetration test is used to conduct a brute force attack for guessing management server login data. In this case, these data packets can be uniquely recognized by the IDS and stored as a new attack signature. Later, an AI will be trained to improve the generated attack signatures continuously. Once it is possible to generate targeted attack signatures and improve them via AI, the cloud infrastructure will be connected to a blockchain. With the help of the blockchain, it should then be possible to select and monitor the various generated attack signatures in a tamper-proof manner. The result will be an IDS that specializes in RSPs, can detect attacks, and can be configured and monitored independently of the RSP operator via blockchain.

References

Academy, R. M., Studies, E., Software, E. R. P., Software, R. P., & Man-, C. M. (2007). Criteria for the selection of ERP software. Informatica Economica, XI(2), 63–66.

Altwaijry, H., & Algarny, S. (2012). Bayesian based intrusion detection system. Journal of King Saud University - Computer and Information Sciences. https://doi.org/10.1016/j.jksuci.2011.10.001

Axelsson, S. (2000). The Base-Rate Fallacy and the Difficulty of Intrusion Detection. ACM Transactions on Information and System Security. https://doi.org/10.1145/357830.357849

Bace, R., & Mell, P. (2001). NIST special publication on intrusion detection systems. In Nist Special Publication.

Brown, C. V., & Heikki, T. (2005). Information systems management handbook (8th ed.).

Can, O., & Sahingoz, O. K. (2015). A survey of intrusion detection systems in wireless sensor networks. 6th International Conference on Modeling, Simulation, and Applied Optimization, ICMSAO 2015 - Dedicated to the Memory of Late Ibrahim El-Sadek. https://doi.org/10.1109/ICMSAO.2015.7152200

Chiba, Z., Abghour, N., Moussaid, K., El Omri, A., & Rida, M. (2016). A Cooperative and Hybrid Network Intrusion Detection Framework in Cloud Computing Based on Snort and Optimized Back Propagation Neural Network. Procedia Computer Science, 83, 1200–1206. https://doi.org/10.1016/j.procs.2016.04.249

Christer, A. H., & Waller, W. M. (1984). Reducing production downtime using delay-time analysis. Journal of the Operational Research Society. https://doi.org/10.1057/jors.1984.103

Debar, H., Dacier, M., & Wespi, A. (2000). Revised taxonomy for intrusion-detection systems. Annales Des Telecommunications/Annals of Telecommunications. https://doi.org/10.1007/BF02994844

Denning, D. E. (1987). An Intrusion-Detection Model. IEEE Transactions on Software Engineering. https://doi.org/10.1109/TSE.1987.232894

El-Bakry, H. M., & Mastorakis, N. (2008). A real-time intrusion detection algorithm for network security. WSEAS Transactions on Communications, 7(12), 1222–1228.

Elvezio, C., Sukan, M., Oda, O., Feiner, S., & Tversky, B. (2017). Remote collaboration in AR and VR using virtual replicas. ACM SIGGRAPH 2017 VR Village, SIGGRAPH 2017. https://doi.org/10.1145/3089269.3089281

Hiltunen, P., Bligh, R., Klett, C., Missalla, M., & Schmidt, H. W. (2008). How to achieve high availability with large calciners and avoid unforseen downtime. TMS Light Metals.

Jatti, S. A. V., & Kishor Sontif, V. J. K. (2019). Intrusion detection systems. International Journal of Recent Technology and Engineering. https://doi.org/10.35940/ijrte.B1540.0982S1119

Johnston, A., Yoakum, J., & Singh, K. (2013). Taking on webRTC in an enterprise. IEEE Communications Magazine. https://doi.org/10.1109/MCOM.2013.6495760

Jow, J., Yang, X., & Han, W. (2017). A survey of intrusion detection systems in smart grid. International Journal of Sensor Networks. https://doi.org/10.1504/IJSNET.2017.083410

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity, 2(1). https://doi.org/10.1186/s42400-019-0038-7

Kiraz, M. S. (2016). A comprehensive meta-analysis of cryptographic security mechanisms for cloud computing. Journal of Ambient Intelligence and Humanized Computing. https://doi.org/10.1007/s12652-016-0385-0

Liao, H. J., Richard Lin, C. H., Lin, Y. C., & Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. In Journal of Network and Computer Applications. https://doi.org/10.1016/j.jnca.2012.09.004

Liu, Y., Li, W., & Li, Y. (2008). Network Traffic Classification Using K-means Clustering. https://doi.org/10.1109/imsccs.2007.52

Mahy, R., Matthews, P., Alcatel-Lucent, & Rosenberg, J. (2010). Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN). In Internet Engineering Task Force (IETF).

Masoni, R., Ferrise, F., Bordegoni, M., Gattullo, M., Uva, A. E., Fiorentino, M., Carrabba, E., & Di Donato, M. (2017). Supporting Remote Maintenance in Industry 4.0 through Augmented Reality. Procedia Manufacturing. https://doi.org/10.1016/j.promfg.2017.07.257

Min, E., Long, J., Liu, Q., Cui, J., & Chen, W. (2018). TR-IDS: Anomaly-Based Intrusion Detection through Text-Convolutional Neural Network and Random Forest. Security and Communication Networks. https://doi.org/10.1155/2018/4943509

Münz, G., Li, S., & Carle, G. (2007). Traffic Anomaly Detection Using K-Means Clustering. GI/ITG Workshop MMBnet.

Radivilova, T., Kirichenko, L., Ageyev, D., Tawalbeh, M., & Bulakh, V. (2018). Decrypting SSL/TLS traffic for hidden threats detection. Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies, DESSERT 2018. https://doi.org/10.1109/DESSERT.2018.8409116

Ripeanu, M. (2001). Peer-to-peer architecture case study: Gnutella network. Proceedings - 1st International Conference on Peer-to-Peer Computing, P2P 2001. https://doi.org/10.1109/P2P.2001.990433

Sahay, B. S., & Gupta, A. K. (2003). Development of software selection criteria for supply chain solutions. Industrial Management and Data Systems. https://doi.org/10.1108/02635570310463429

Sathya, R., & Abraham, A. (2013). Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. International Journal of Advanced Research in Artificial Intelligence. https://doi.org/10.14569/ijarai.2013.020206

Sherry, J., Lan, C., Popa, R. A., & Ratnasamy, S. (2015). BlindBox: Deep Packet Inspection over Encrypted Traffic. Computer Communication Review. https://doi.org/10.1145/2785956.2787502

Shields, P., & Rangarjan, N. (2013). A Playbook for Research Methods: Integrating Conceptual Frameworks and Project Management. New Forums Press.

Snort 2.1 Intrusion Detection. (2004). In Snort 2.1 Intrusion Detection. https://doi.org/10.1016/b978-1-931836-04-3.x5000-0

T. Weber & R. Buchkremer:
Monitoring Remote Service Platforms Using Artificial Intelligence-Based Distributed Intrusion Detection

717

Stone, A. (2007). Natural-language processing for intrusion detection. Computer. https://doi.org/10.1109/MC.2007.437

Teeter, R. A., Alles, M. G., & Vasarhelyi, M. A. (2010). The remote audit. Journal of Emerging Technologies in Accounting. https://doi.org/10.2308/jeta.2010.7.1.73

Vossing, M. (2017). Towards managing complexity and uncertainty in field service technician planning. Proceedings - 2017 IEEE 19th Conference on Business Informatics, CBI 2017. https://doi.org/10.1109/CBI.2017.50

Wang, Y. (2005). A multinomial logistic regression modeling approach for anomaly intrusion detection. Computers and Security. https://doi.org/10.1016/j.cose.2005.05.003

Werner, M., & Bechini, G. (2019). Customer technical support: OEM collaboration in a digitalized world. Society of Petroleum Engineers - Abu Dhabi International Petroleum Exhibition and Conference 2018, ADIPEC 2018. https://doi.org/10.2118/192623-ms

Yin, Y., Zhou, Z., Chen, Y., Liu, Q., & Long, Y. (2006). Realization of a web-based remote service platform. Proceedings - 2006 10th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2006. https://doi.org/10.1109/CSCWD.2006.253079

Zarpelão, B. B., Miani, R. S., Kawakani, C. T., & de Alvarenga, S. C. (2017). A survey of intrusion detection in Internet of Things. In Journal of Network and Computer Applications. https://doi.org/10.1016/j.jnca.2017.02.009

Sensory-Marketing-Evaluation of E-Commerce Websites with Artificial Intelligence

Kevin Hamacher[1] & Rüdiger Buchkremer[2]

[1] UCAM Universidad Católica San Antonio de Murcia, Spain; e-mail: khamacher@alu.ucam.edu
[2] Institute of IT Management and Digitization, FOM University of Applied Sciences, Düsseldorf, Germany; e-mail: ruediger.buchkremer@fom.de

Abstract Multisensory consumer engagement on e-commerce websites is technically limited to visual, acoustic, and written elements. Consumers communicate, buy, and share products and services via digital environments in which sensory information is limited. To improve consumers' online sensory experience, media types and the content need to be quantitatively assessed and adapted. This project aims to develop a quantitative model, an Online Sensory Marketing Index (OSMI), which assesses e-commerce websites in multisensory communication quality. The OSMI will be supported by an automatic procedure that is based on artificial intelligence. Content of texts, images, and videos is evaluated by natural language processing (NLP), natural language generation (NLG) as well as automatic machine learning (AutoML) procedures. Multiple e-commerce websites from various industries are examined.

1   Introduction

In the past three decades, researchers have reported that all five human senses have an immense significance towards consumers' purchasing decisions, and many organizations modify their marketing activities accordingly. 'Sensory marketing' is becoming increasingly important in research and practice and is underpinned by the progressive debate in the scientific literature during the last decade (Bleier, Harmeling, & Palmatier, 2019; Krishna, 2012; Peck & Childers, 2008). It is partly due to the increasing number of exchangeable products and an information overload for consumers. Besides, the increasing intensity of competition makes it more difficult for suppliers to attract consumers' attention. Consumer and buyer behavior have also changed dramatically in recent years because today, consumers strive for individualization and personalization. Sensory information about products and services can influence people's attitudes, purchasing intentions, and consumption (Petit, Cheok, Spence, Velasco, & Karunanayaka, 2015). All perceptible stimuli of a product, e. g., a new dress or car, offer valuable information about the product's perceived quality and can have a significant influence on the purchasing decision process (Elder & Krishna, 2010; Krishna, Cian, & Sokolova, 2016). After all, (advertising) messages always have a more substantial effect if they affect the consumer through more than one sense (Krishna, 2012).

As a result of consumer and buyer behavior changes, sensory marketing has increasingly become part of the scientific discourse. Krishna defines sensory marketing as "marketing that engages the consumers' senses and affects their perception, judgment and behavior" (Krishna, 2012). Impressions about seeing, hearing, feeling, tasting, and smelling are decoded in the human brain (Barsalou, 2008). People can react and possibly buy a product or a service. A sensory approach's advantages are obvious: an increase in attention for the product and/or the brand, a resulting higher memorability value, and ultimately faster recognition, combined with the possibility of differentiation from the relevant competition (Krishna, 2012).

The question arises as to which e-commerce strategy should be implemented for sales via the Internet. - The fundamental problem with marketing goods digitally is that some human stimuli cannot be directly addressed or only to a minimal extent because direct contact with the consumer is not apparent. This limitation applies

especially to haptics when consumers cannot feel the quality of clothing, for example, as they are used to in terms of weight or material properties (Klatzky, Lederman, & Matula, 1993). Nevertheless, current statistics confirm that the e-commerce channel is on a growth course, particularly in the US, Asia, and Northern Europe (Pappas, Kourouthanassis, Giannakos, & Lekakos, 2017). Furthermore, forecasts for the coming financial year point to even more significant growth of up to $4.93 trillion in 2021, which will be more than double compared with the 2017 year's expenses of $2.3 trillion, driven by about ¼ of the world's population already shopping online (Adam, Alhassan, & Afriyie, 2020; OECD, 2020).

2 Problem definition

2.1 Problem Definition

Sensory marketing is currently at a tipping point because the challenge lies in transferring the digital world's address. This leads to a complex environment in which e-commerce business constantly demands new communication and style to be resistant to future challenges (Bleier idr., 2019). To enable such a tremendous change, it is required to design a new understanding of sensory marketing in terms of digitization. The literature offers various ways to acknowledge changes and presents some hints towards them. Nevertheless, acknowledging research in sensory marketing of the last two decades leads to the assumption that an overarching viewpoint needs to be taken in a digitally changing world. However, current research does not provide a holistic and measurable view of how an e-commerce website needs to look like from a sensory perspective. The main problem is that current sensory marketing models do not quantitatively incorporate digitization trends and sensory aspects of communication. There is no generic guidance on how to design an e-commerce website in an appropriate way regarding sensory marketing that allows the evaluation of the sensory communication quality.

The question arises of how sensory elements can be measured on a website. Sensory elements can be accommodated in pictures, texts, and videos and a collection with manual methods is possible and practicable. In recent years, however, artificial intelligence (Ekbia, 2010) has developed to such an extent that automatic acquisition and identification of sensory elements should be possible. The methods of choice for our project are "natural language processing" (Buchkremer idr., 2019) and

"automatic machine learning" (Braka, Buchkremer, & Ebener, 2020) for the acquisition and analysis of text, images, sounds, and videos (Kacprzyk & Zadrozny, 2010; Truong idr., 2019).

Thus, online sensory marketing research is still in its infancy, and a deeper analysis must be carried to fill this gap. Thus, the research project will pick up on the need described above and create an automated assessment framework named Online Sensory Marketing Index (OSMI). The following paragraph highlights the scientific objectives and related research questions.

2.2   Scientific Objectives

The related problem scope frames the research area and presents questions that the project will answer. It furthermore details aspects where the OSMI delivers specific answers to the solution. The OSMI is based on two interdisciplinary models. First of all, the Web Quality Index (WQI) can be considered as an assessment measure for websites in general (Fernández-Cavia, Rovira, Díaz-Luque, & Cavaller, 2014). Secondly, the model developed by Hultén (Hultén, 2011) serves as a framework to describe sensory elements in marketing. The digitization trends have no long history, and thus, no generic (online) sensory evaluation model has yet been developed. For this reason, research is based on the problem of the extent to which compensation can be achieved by combining two or even more sensory stimuli and how this affects consumers and their purchasing behavior keeping in mind that direct sensory consumer appeal on e-commerce websites is limited to visuality and acoustics (Bleier idr., 2019; Petit, Velasco, & Spence, 2019; Yazdanparast & Spears, 2013).

An evaluation framework would provide the following benefits to the body of knowledge in online sensory marketing design:

- Pattern search for best practice success factors: initially, an evaluation framework allows to assess existing data. Various dimensions of the OSMI can be evaluated regarding patterns that different e-commerce websites provide. Therefore, presenting guidelines for further implementation of other e-commerce websites is a prerequisite.

- Sales increase: an OSMI-based optimization of e-commerce websites leads to a better consumer approach, a viable generic argument of using the framework for business purposes.
- Cost savings: A practical and validated online sensory consumer experience could, among other things, reduce the cost to retain the customer.

The most significant disadvantage of the digital customer journey is that customers are not directly able to inspect the product and have to wait a particular time for it, under uncertainty of the accuracy of the shown representation (Hong & Pavlou, 2014). A sensory consumer approach could remedy this. For instance, a virtual touch becomes a crucial aspect in online shopping and will be examined in more detail later (Brasel & Gips, 2015).

This research project aims to create an automatic assessment framework for sensory communication quality that answers the problem mentioned above. To reach this goal, the following research questions (RQ) are addressed as the underlying structure that guides the overall research process:

RQ1: To what extent can industry-specific taxonomies be determined based on OSMI measurements?

Based on the current status, OSMI evaluations of e-commerce websites need to be acquired manually. The question arises whether it is feasible to encode the whole website automatically into a keyword catalog/taxonomy related to sensory stimuli.

RQ2: To what extent is the automated extraction and identification of OSMI elements possible through artificial intelligence?

The automated examination of e-commerce websites includes additional components beyond the pure text, such as the image material's quality and the acoustic elements. For this reason, the OSMI needs to be expanded to all media types to be evaluated by artificial intelligence.

RQ3: To what extent is it feasible to quantitatively determine and compare industry-specific OSMI scores?

Building upon RQ1 and RQ2, the next step is to generate an automatically running OSMI framework, which is, in the first instance, able to crawl websites for sensory keywords. Using, e.g., the keyword catalog, a comparison is made to determine whether the automatic assessment result is identical or even better compared to the index from the manual analysis.

RQ4:   How can guiding principles for the determination and application of OSMI indices be generated?

Finally, the last research question relates to OSMI recommendations that can be derived predominantly from quantitative research to make the OSMI most useful for practitioners.

3   Methodology

We plan to use qualitative and quantitative research methods to specify the final OSMI evaluation framework. Qualitative research is used to develop a set of hypotheses that are tested with a quantitative study. The initial qualitative research uses qualitative content analysis (Hsieh & Shannon, 2005) as the foundation of an exploratory study to develop indicators for each framework's parameter. A qualitative approach is chosen for the exploratory study (Petit idr., 2019). Sensory marketing trends have been researched in terms of digitization to existing frameworks. Within this research scope, it is considered a success factor to include existing sensory marketing insights and assessment tools from scientific disciplines in general into consideration. It is planned to derive the OSMI from analyzing existing literature related to online consumer experience and sensory marketing trends. The chosen research method furthermore concentrates on the reduction of data by paraphrasing and categorizing. Analogous to the study by Fernandez-Cavia et al. (2014) (Fernández-Cavia idr., 2014), the evaluation model makes use of indicators that apply different scales: The scale 0-1 in case a specific characteristic is present or completely absent. The 0-2 or 0-3 scale, on the other hand, in the case of a more specific assessment basis, which allows more factual statements to be made about the quality of the characteristic, e.g., Weak (0); Standard (1); Good (2); Excellent (3). The combined index across all senses can be seen as a general evaluation of the sensory quality of the analyzed e-commerce website. The closer the OSMI value is to 1, the more successful the website communicates concerning

sensory marketing theoretically by excluding possible sensory overloading effects. A score close to 0, on the other hand, suggests that the website does not contain sensory elements and does not meet the indicators for successful sensory consumer demands that are necessary and relevant for an e-commerce website. The general advantage of the method chosen here is that these indicators allow comparisons to be made between the websites analyzed, and those can be identified that have received a positive rating and can be compared with one another based on individual indicators. In this way, examples of good and less good sensory designs can be identified. An analysis of the improvement potential is easily possible by the specific details to be used fundamentally for pointing out sensory optimizations. According to this, the research methods of the research can be classified as follows:

At the early stage of research, it is planned to conduct expert interviews with marketing leaders to obtain opinions and suggestions for improvement on the status quo of the OSMI. Besides, a taxonomy for the five human senses (classification scheme) must be built manually about qualitative research. In this, a keyword catalog is created, which bundles related sensory terms per sense. The associated goal is to visualize patterns for various branches and relationships of sensory information represented in texts used on these e-commerce platforms to point out best-in-class sensory communication quality.

In a second step, the manually based keyword catalog will be developed within the framework of an automated tool, enabling to crawl off the e-commerce websites according to the compiled terms, and a score is automatically generated. At this research project level, the Online Sensory Marketing Index will be combined with Artificial Intelligence (AI). The aim of this idea is that the OSMI becomes an automatically running system to check and evaluate an e-commerce website in terms of sensory communication quality. AI has to evaluate e-commerce websites fully automatically, including presenting an overview of rankings for each examined parameter. Based on rankings, further recommendations are possible and should also be automatically displayed by the OSMI and its use of AI. In detail, it is planned to make use of natural language processing (NLP) as a text mining method (Manning, Bauer, Finkel, & Bethard, 2014). It is not intended to use a cluster analysis as an approach since the dimensions investigated have already been determined with the five human senses. With the use of NLP, it will be possible to be focused on keyword extraction on the one hand but also spoken language in audio or video

elements, the so-called speech processing, on the other hand. Further, it is under consideration to extend the new OSMI framework towards sentiment analysis to extract meanings of sensory formulated texts automatically (Collobert idr., 2011). This information extraction technique should also include metaphors, which usually require learned interpretation patterns.

To complement the new OSMI approach, the automated assessment tool will also be extended to analyze other relatively objective website components and sensory indicators like, for instance, 3D-elements, background colors, or the resolution of product or mood pictures in general via automated machine learning using Cloud AutoML. This allows generating a grammatical score to be recorded with a "complication index" based on the text length. In this context, the weighting of individual keywords and their positioning within texts or web pages is to be examined, which could then be taken into account in automatic indexing.

Finally, the automated AI OSMI framework is to be compared quantitatively with the analogously determined values employing a field test of approximately 80 to 100 e-commerce websites of mainly global brands. Implementation will be done in collaboration with companies from various industries to determine the OSMI in a feasible way.
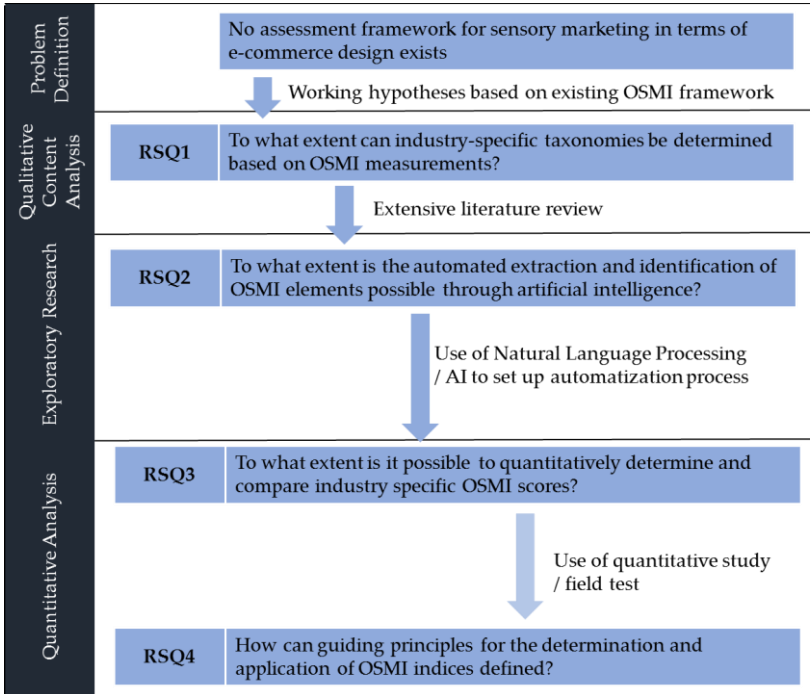
Figure 1: Research Process

As depicted above, the research process is vertically divided into four phases, four research questions, and a plan for writing papers. It starts from the problem definition and qualitative content analysis over the exploratory research towards the quantitative analysis and summing up with the research results. The size of each box represents the scope of investigation of the respective phase of the research. Starting with the first phase, the research process initially needs to be framed within guiding hypotheses as a foundation. The problem definition phase is mainly intended to show, based on current scientific literature, that no model exists that is capable of generating qualitative assessments of the sensory consumer appeal on e-commerce websites. As mentioned above, also expert interviews are to take place at this point to critically review the current OSMI approach and to get further improvement.

The second phase, the qualitative content analysis, follows this and examines the current scientific literature concerning the topicality of the sensory indicators. It will also be examined whether additional sensory indicators could supplement the OSMI approach.

The exploratory research, the third phase, will then deal with the new OSMI approach based on artificial intelligence. Therefore, as mentioned above, NLP methods, including keyword extraction, and AutoML will be examined. This phase's results will be published within a paper to document the current status of the research.

With the previously developed automated evaluation model, a quantitative analysis can then be aimed at the fourth and last phases. In this section, the automated evaluation of numerous practical examples with the manual evaluation of sensory communication quality will be compared. Finally, we will critically examine the effectiveness of the new OSMI approach to give practitioners advice.

4   Preliminary/Expected results

Within the experimental quantitative phase, we expect that some sensory communication aspects can be evaluated via machine-based methods in principle. We recently conducted an extensive test in a big data project with 40 students to support our assumption. Here, students had to independently analyze different senses on given, international e-commerce platforms with big data methods. After crawling text and images using various methods, they could intensively evaluate data and assess the quality of sensory communication. For this purpose, texts were evaluated using word embedding technologies on the one hand (Horn, Erhardt, Di Stefano, Bosten, & Buchkremer, 2020), (Hussain idr., 2020), and elements of images were recognized using object recognition via a Google cloud framework on the other. Interesting results can already be seen, for example, in the fact that concerning examined e-commerce websites for wines, tendencies are recognizable to the effect that the taste and associations of red wines are often related to images with chocolate - white wines, on the other hand, relates to pictures of lemons. We already presented this first investigation's approach in a technical article published at the INTED 2021 conference (Hamacher & Buchkremer, 2021).

5   Future development

Some researchers present ongoing developments towards digitizing of senses - see also Spence, Obrist, Velasco, & Ranasinghe for an overview (Spence, Obrist, Velasco, & Ranasinghe, 2017). Knowledge from the field of sensory marketing can often be transferred to the digital space. Thus, e-commerce consumer experience, as

a particular field of online consumer experience, has enormous potential to be enriched with sensory communication aspects. For instance, a good mood can compensate frustration caused by a lack of haptics (Yazdanparast & Spears, 2013). An appealing design of the online shop, positive product descriptions or ratings, chats with friendly consultants, plenty of humor, and/or stimulating images can attract people with a high need for touch to online shopping (Roggeveen, Grewal, Townsend, & Krishnan, 2015; San-martín, González-benito, Martos-partal, & San-martín, 2017; Yazdanparast & Spears, 2013). For this and other reasons, the intersection of these two dimensions is extensive.

Assessment frameworks for the two dimensions only exist to a minimal extent. Some key performance indicators in online consumer experience are known, for example, a bounce rate or a click rate, but there is a lack of qualitative metrics, especially in terms of online sensory communication quality. An overall assessment framework for offline and online communication elements is missing in sensory marketing but very important due to the findings described in this paper. Going along with this argument, artificial intelligence is the key to fill existing technology knowledge gaps. Accordingly, AI will have a crucial role in designing online sensory consumer experiences in the future. For the time being, it is evident that sensory marketing, in particular, does not have extensive research contacts with AI. Therefore, the overall goal is to connect these dimensions more closely and thus close the mentioned knowledge gap.

References

A. Krishna, "An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior," J. Consum. Psychol., vol. 22, no. 3, pp. 332–351, 2012.

J. Peck and T. L. Childers, "Sensory factors and consumer behavior," Handb. Consum. Psychol., pp. 193–219, 2008.

A. Bleier, C. M. Harmeling, and R. W. Palmatier, "Creating effective online customer experiences," J. Mark., vol. 83, no. 2, pp. 98–119, 2019.

O. Petit, A. D. Cheok, C. Spence, C. Velasco, and K. T. Karunanayaka, "Sensory marketing in light of new technologies," Proc. 12th Int. Conf. Adv. Comput. Entertain. Technol. - ACE '15, pp. 1–4, 2015.

R. S. Elder and A. Krishna, "The Effects of Advertising Copy on Sensory Thoughts and Perceived Taste," J. Consum. Res., vol. 36, no. 5, pp. 748–756, 2010.

A. Krishna, L. Cian, and T. Sokolova, "The power of sensory marketing in advertising," Curr. Opin. Psychol., vol. 10, pp. 142–147, 2016.

L. W. Barsalou, "Grounded Cognition," Annu. Rev. Psychol., vol. 59, no. 1, pp. 617–645, 2008.

R. L. Klatzky, S. J. Lederman, and D. E. Matula, "Haptic Exploration in the Presence of Vision," J. Exp. Psychol. Hum. Percept. Perform., vol. 19, no. 4, pp. 726–743, 1993.

I. O. Pappas, P. E. Kourouthanassis, M. N. Giannakos, and G. Lekakos, "The interplay of online shopping motivations and experiential factors on personalized e-commerce: A complexity theory approach," Telemat. Informatics, vol. 34, no. 5, pp. 730–742, 2017.

I. O. Adam, M. D. Alhassan, and Y. Afriyie, "What drives global B2C E-commerce? An analysis of the effect of ICT access, human resource development and regulatory environment," Technol. Anal. Strateg. Manag., vol. 0, no. 0, pp. 1–16, 2020.

OECD, "E-commerce in the time of COVID-19," no. October, pp. 1–10, 2020, [Online]. Available: http://www.oecd.org/coronavirus/policy-responses/e-commerce-in-the-time-of-covid-19-3a2b78e8/#biblio-d1e705.

H. R. Ekbia, "Fifty years of research in artificial intelligence," Annu. Rev. Inf. Sci. Technol., vol. 44, no. 1, pp. 201–242, 2010.

R. Buchkremer et al., "The Application of Artificial Intelligence Technologies as a Substitute for Reading and to Support and Enhance the Authoring of Scientific Review Articles," IEEE Access, vol. 7, pp. 65263–65276, 2019, doi: 10.1109/ACCESS.2019.2917719.

D. Braka, R. Buchkremer, and S. Ebener, "Informationsextraktion und kartografische Visualisierung von Haushaltsplänen mit AutoML-Methoden," R. Buchkremer, T. Heupel, and O. Koch, Eds. Wiesbaden: Springer Fachmedien Wiesbaden, 2020, pp. 107–128.

J. Kacprzyk and S. Zadrozny, "Computing with words is an implementable paradigm: Fuzzy queries, linguistic data summaries, and natural-language generation," IEEE Trans. Fuzzy Syst., vol. 18, no. 3, pp. 461–472, Jun. 2010.

A. Truong, A. Walters, J. Goodsitt, K. Hines, C. B. Bruss, and R. Farivar, "Towards Automated Machine Learning: Evaluation and Comparison of AutoML Approaches and Tools," Aug. 2019.

J. Fernández-Cavia, C. Rovira, P. Díaz-Luque, and V. Cavaller, "Web quality index (WQI) for official tourist destination websites. proposal for an assessment system," Tour. Manag. Perspect., vol. 9, pp. 5–13, 2014.

B. Hultén, "Sensory marketing: the multi-sensory brand-experience concept," Eur. Bus. Rev., vol. 23, no. 3, pp. 256–273, May 2011.

A. Yazdanparast and N. Spears, "Can Consumers Forgo the Need to Touch Products? An Investigation of Nonhaptic Situational Factors in an Online Context," Psychol. Mark., vol. 30, no. 1, pp. 46–61, 2013.

O. Petit, C. Velasco, and C. Spence, "Digital Sensory Marketing: Integrating New Technologies Into Multisensory Online Experience," J. Interact. Mark., vol. 45, pp. 42–61, 2019, doi: 10.1016/j.intmar.2018.07.004.

Y. Hong and P. A. Pavlou, "Product fit uncertainty in online markets: Nature, effects, and antecedents," Inf. Syst. Res., vol. 25, no. 2, pp. 328–344, 2014.

S. A. Brasel and J. Gips, "Interface Psychology: Touchscreens Change Attribute Importance, Decision Criteria, and Behavior in Online Choice," Cyberpsychology, Behav. Soc. Netw., vol. 18, no. 9, pp. 534–538, 2015.

H. F. Hsieh and S. E. Shannon, "Three approaches to qualitative content analysis," Qual. Health Res., vol. 15, no. 9, pp. 1277–1288, 2005.

C. D. Manning, J. Bauer, J. Finkel, and S. J. Bethard, "The Stanford CoreNLP natural language processing toolkit," Aclweb.Org, pp. 55–60, 2014.

R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural Language Processing (Almost) from Scratch," J. Mach. Learn. Res., no. 12, pp. 2493–2537, 2011.

N. Horn, M. S. Erhardt, M. Di Stefano, F. Bosten, and R. Buchkremer, "Vergleichende Analyse der Word-Embedding-Verfahren Word2Vec und GloVe am Beispiel von Kundenbewertungen eines Online-Versandhändlers," in Künstliche Intelligenz in Wirtschaft & Gesellschaft, 2020, pp. 559–581.

A. Hussain et al., "Artificial intelligence-enabled analysis of UK and US public attitudes on Facebook and Twitter towards COVID-19 vaccinations (Preprint)," J. Med. Internet Res., Dec. 2020, doi: 10.2196/26627.

K. Hamacher and R. Buchkremer, "MEDIATION OF ONLINE SENSORY MARKETING THROUGH ONLINE COLLABORATION SOFTWARE," in 15th annual International Technology, Education and Development Conference, Mar. 2021, pp. 1387–1396, doi: 10.21125/inted.2021.0320.

C. Spence, M. Obrist, C. Velasco, and N. Ranasinghe, "Digitizing the chemical senses: Possibilities & pitfalls," Int. J. Hum. Comput. Stud., vol. 107, pp. 62–74, 2017.

A. L. Roggeveen, D. Grewal, C. Townsend, and R. Krishnan, "The Impact of Dynamic Presentation Format on Consumer Preferences for Hedonic Products and Services," J. Mark., vol. 79, no. 6, pp. 34–49, 2015.

S. San-martín, Ó. González-benito, M. Martos-partal, and S. San-martín, "To what extent does need for touch affect online perceived quality?," Int. J. Retail Distrib. Manag., vol. 45, no. 9, pp. 950–968, 2017.

KNOWLEDGE-BASED PLANNING AND CONTROLLING WITH METHODS OF ARTIFICIAL INTELLIGENCE TO INCREASE EFFICIENCY IN IT PROJECTS

SASCHA BRÜGGEN[1] & ALEXANDER HOLLAND[2]

[1] UCAM Universidad Católica San Antonio de Murcia, Av. de los Jerónimos, Guadalupe de Maciascoque, Murcia, Spain; e-mail: sascha.brueggen@outlook.com
[2] FOM University of Applied Sciences, Herkulesstraße 32, 45127 Essen, Germany; e-mail: alexander.holland@fom.de

Abstract Because the success of IT projects in companies is increasingly becoming a competitive factor, this paper aims to analyze how knowledge-based tasks in an IT project can be supported with the help of artificial intelligence methods to carry out the IT project more efficiently. To answer the research question, a qualitative method in form of expert interviews will be used on the one hand and a discrete event simulation on the other hand to achieve quantitative results. In the simulation, it is intended to create a model without the use of AI elements. Under the same conditions, a project will be developed to support the knowledge-based tasks in the second step. The different constellations of the model can be adapted depending on the focus and the question. These models will be operationalized with parameters and will be compared under different constellations to measure the efficiency quantitatively. The findings of this study allow a statement whether it makes sense to support specific tasks in an IT project with the help of AI methods.

1   Introduction

The ability to recognize and conduct transformations in the organization is one of the most important for companies nowadays. Knowing about the requests from their markets with their participants and the need to offer the right products to the right place, at the right time, and quality is essential for the survival of a company.

1.1   IT Project management

The pressure on IT departments to implement projects that contribute to the company's success remains high to reach their goals in the dimensions of time, quality, and effort. For the question what project management is about Reich describes it "as an area in which action is paramount and in which tasks, budgets, people and schedules must be managed and controlled to achieve expected results" (Reich, B.H., Gemino, A, Sauer, C., 2008a). Nevertheless, the achievement of the project objectives is not an end in itself. A research study in 2019 from Capgemini Germany figure out that only 15% of the interviewed companies will decrease their IT budget comparing the previous year. 50% of them will an increased budget, and nearly 35% will have the same budget level as last year. The main target of the IT budget with 72% will be the digitization of the processes and increasing the organization's efficiency (Dumslaff & Heimann, 2018). The main issue for these companies is that there are no sufficient human resources working in the IT and project departments to manage the upcoming demands. Also, the need to handle the knowledge of the employees is an increasingly severe topic where companies are fully aware of this circumstance. With these challenges, it is becoming increasingly challenging to achieve the stated goals of IT projects. This result is supported by The Standish Group International from 2015, which shows that only 6% of all grand projects were completed as planned, and only 11% of all large projects were completed as planned (The Standish Group International, Inc, 2015).

1.2   Knowledge management

An essential part of planning and conducting project management is the part of knowledge management (KM). That is also valuable to examine as already Gemino, Lee and Reich mentioned (Reich, B.H., Gemino, A, Sauer, C., 2008a) (Gemino et al., 2007) (Lee & Lee, 2000). As Wang et al. point out, projects are knowledge-

intensive, but IT projects in a particular way (Wang, E., Lin, C., Jiang, J.J., Klein, G., 2007). This could be justified because IT projects often aim to develop new software, which is an imaginary product. On the other hand, the number of IT trends with knowledge-intensive skills and competence requirements is growing rapidly. Independent from the type of IT project or selected procedure model, often there are the following parameters given to hit the project goals. Lech uses the knowledge taxonomy of project knowledge to describe the parameters of resources, time, and cost to reach the target of a specific project, schedules, milestones, and other artifacts (Lech, 2014). In projects, all forms of knowledge are used in many places, and someone in projects must be aware of it. Hanisch also refers to the need that managing knowledge in projects is one of the most important tasks of a project manager, and the necessity to focus research in the field of project-based companies is growing (Hanisch et al., 2009).

Nevertheless, the project manager is responsible for many tasks, but two components must be distinguished in terms of knowledge management. Lech mentions that there are two types of knowledge that a project manager must have at his disposal. The one is the generic project management knowledge available in the PM body of knowledge guides also PM process models, and the other one is the product-related project management knowledge. This includes best practices for performing the projects involving implementing a specific system or topic (Lech, 2014). Judging by these facts, it should also be in the project manager's interest to manage knowledge as efficiently as possible in the projects.

2   Problem definition

The meaning of knowledge management is mentioned as a success factor in a large number of documentations. In general, there are some descriptions of how knowledge management is part of projects. One of them is the definition of Reich, Gemino and Sauer: "Knowledge management in the context of a project is the application of principles and processes designed to make relevant knowledge available to the project team. Effective knowledge management facilitates the creation and integration of knowledge, minimizes knowledge losses, and fills knowledge gaps throughout the duration of the project" (Reich, B.H., Gemino, A, Sauer, C., 2008b). Lech refers that identifying knowledge, mapping, and sourcing that does not take place in a structured manner (Lech, 2014). Gasik developed a

model of project knowledge, which includes different views on knowledge within a project. He linked the micro-knowledge, which means the needed knowledge to perform one task, with the macro-knowledge view, which includes a complete knowledge possessed by a given person to increase efficiency (Gasik, 2011).

The issue regarding knowledge or rather the management of knowledge in projects is well-known, and there are many frameworks and modeling, e.g., from Gasik or the examinations from Lech. However, a quantitative simulation of a project management approach is still missing to prove that knowledge-based components with the support of artificial intelligence can run projects more efficiently.

This paper investigates to what extent the use of selected artificial intelligence or available services in knowledge-based tasks within an IT project leads to increased efficiency. The classic project management dimensions such as time, costs, and quality serve as variables, although other dimensions can also be used, e.g., quantity according to the functionalities in a software project.

3 Methodology

To ensure that both quantitative and qualitative questions are answered, two types of methods will be planned. For the questions of qualitative nature, the methodology of expert interviews is planned. The research project first provides for selecting relevant experts, e.g., CIOs., CEOs, heads of Project management, and high-level decision-makers. The interviews will be analyzed the answers by qualitative data analysis (QDA), according to Mayring. The interviews will be conducted as semi-structured interviews in order to be able to respond spontaneously to the development of the interview. To maintain a high degree of objectivity, reliability, and validity, each role is interviewed twice with different interviewees to obtain possible differences within a role. The results will use in the following discussion to underline similarities and deviations and try to explain these.

To answer quantitative questions, it will create and conduct a discrete event simulation model to evaluate the effects of using AI methods in IT projects' knowledge-based components. Adler, Mandelbaum, Nguyen, and Schwerer also used simulation to study processes. To identify weak points, bottlenecks, and performance losses, this methodology is preferable because the rapid modifiability

of the simulation model allows new results and insights (Adler et al., 1995). Bassil also used a discrete simulation to evaluate software development life cycles, a software methodology for designing, building, and maintaining information and industrial systems (Bassil, 2012). This work shows that this type of simulation can investigate complex processes and make new findings visible.

MATLAB® Release 2020a software, a comprehensive numerical mathematics software package, is used to create and run the discrete simulation. The strengths lie in vector and matrix calculations, whereby the software is divided into an introductory module and numerous extension packages. Furthermore, in comparison to other simulation programs, MATLAB® offers the advantages of automated evaluation of larger data sets at high speed, the possibility of efficient multiple evaluations of the same data sets with programmed analysis routines, and integrated, flexibly applicable graphics and statistics functions. A particular position of the extensions takes Simulink, which provides a graphical interface for modeling and simulation of systems using signal flow graphs. Besides, the extension StateFlow is to be mentioned. It is a tool for modeling event-driven reactive systems with a finite number of states (Angermann et al., 2021).

The simulation structure represents the work breakdown structure of an IT project, which includes the classic phases of initialization, definition, planning, controlling, and finalization. Each of these phases, in turn, includes defined tasks that are primarily planned and carried out sequentially. Possible dependencies between the tasks are taken into account in the same way as possible returns to previous tasks if these are necessary. The individual tasks are provided with the classic parameters of project management quality, costs, and time (duration) and assigned corresponding values. After simulating the reference model, three values are calculated for each parameter, reflecting the project's total value. This model with the values mentioned above serves as a reference model with which all of the following models can be measured and compared, which is illustrated in figure 1.

Figure 1: Reference model vs. AI supported KM model
Source: own illustration

In the second step, the same model structure is used to identify the tasks related to knowledge management outlined in red in figure 1. Every task related to knowledge management is examined to determine which artificial intelligence methods are suitable for performing this task more efficiently. Once a method has been found, the values of the parameters in quality, costs, and time (duration) also change. Objectively it can only be said that the use of AI in knowledge-based tasks leads to either "no improvement", "medium improvement," or "great improvement" in the individual parameters. However, these three categories cannot be compared quantitatively. To make this possible, predefined ranges are defined in the three categories, and the values are determined using random variables. This procedure is carried out with all knowledge-based tasks. For tasks with no relation to knowledge management, the parameters quality, costs, and time (duration) remain identical to the reference model. With the comparison between the reference model and the new knowledge management-optimized model, a sufficient number of simulations runs with statistical methods can now be used to investigate the extent to which efficiency can be increased using AI methods for knowledge-based tasks within IT projects, which is shown as an example in figure 2.
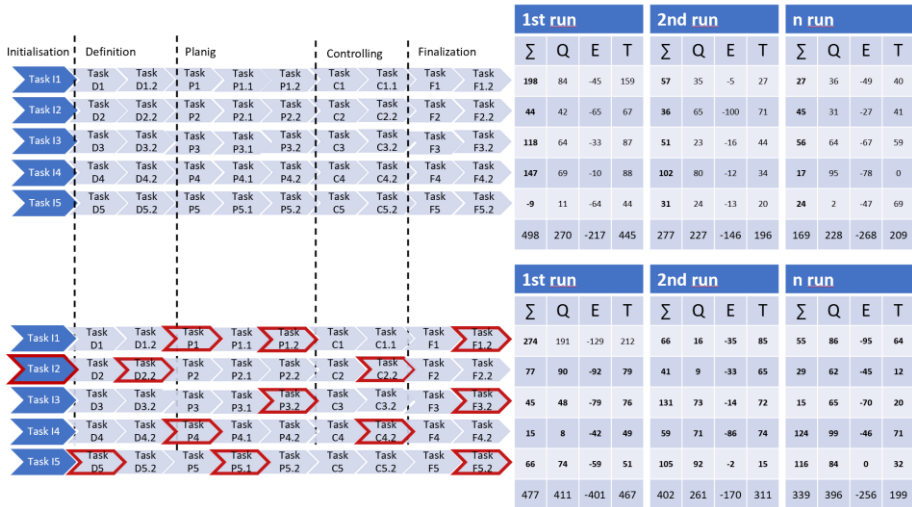
Columns (Initialisation | Definition | Planig | Controlling | Finalization):

	Initialisation	Definition		Planig			Controlling		Finalization	
Task I1	Task D1	Task D1.2	Task P1	Task P1.1	Task P1.2	Task C1	Task C1.1	Task F1	Task F1.2
Task I2	Task D2	Task D2.2	Task P2	Task P2.1	Task P2.2	Task C2	Task C2.2	Task F2	Task F2.2
Task I3	Task D3	Task D3.2	Task P3	Task P3.1	Task P3.2	Task C3	Task C3.2	Task F3	Task F3.2
Task I4	Task D4	Task D4.2	Task P4	Task P4.1	Task P4.2	Task C4	Task C4.2	Task F4	Task F4.2
Task I5	Task D5	Task D5.2	Task P5	Task P5.1	Task P5.2	Task C5	Task C5.2	Task F5	Task F5.2

1st run				2nd run				n run			
Σ	Q	E	T	Σ	Q	E	T	Σ	Q	E	T
198	84	-45	159	57	35	-5	27	27	36	-49	40
44	42	-65	67	36	65	-100	71	45	31	-27	41
118	64	-33	87	51	23	-16	44	56	64	-67	59
147	69	-10	88	102	80	-12	34	17	95	-78	0
-9	11	-64	44	31	24	-13	20	24	2	-47	69
498	270	-217	445	277	227	-146	196	169	228	-268	209

	Initialisation	Definition		Planig			Controlling		Finalization	
Task I1	Task D1	Task D1.2	Task P1	Task P1.1	Task P1.2	Task C1	Task C1.1	Task F1	Task F1.2
Task I2	Task D2	Task D2.2	Task P2	Task P2.1	Task P2.2	Task C2	Task C2.2	Task F2	Task F2.2
Task I3	Task D3	Task D3.2	Task P3	Task P3.1	Task P3.2	Task C3	Task C3.2	Task F3	Task F3.2
Task I4	Task D4	Task D4.2	Task P4	Task P4.1	Task P4.2	Task C4	Task C4.2	Task F4	Task F4.2
Task I5	Task D5	Task D5.2	Task P5	Task P5.1	Task P5.2	Task C5	Task C5.2	Task F5	Task F5.2

1st run				2nd run				n run			
Σ	Q	E	T	Σ	Q	E	T	Σ	Q	E	T
274	191	-129	212	66	16	-35	85	55	86	-95	64
77	90	-92	79	41	9	-33	65	29	62	-45	12
45	48	-79	76	131	73	-14	72	15	65	-70	20
15	8	-42	49	59	71	-86	74	124	99	-46	71
66	74	-59	51	105	92	-2	15	116	84	0	32
477	411	-401	467	402	261	-170	311	339	396	-256	199

Figure 2: Simulation result
Source: own illustration

With a combination of the results by expert interviews and simulation will the research question be answered.

4   Preliminary/Expected results

The expectations in terms of answering the research question are divided into two forms of methods. On the one hand, results are obtained from expert interviews and quantitative content analysis. According to Kuckartz, the quality of the information depends on selecting the interview participants, who take part in the expert interviews, and whose answers are subsequently evaluated (Kuckartz, 2018). A basic breakdown of experts can be divided into three areas. On the one hand, there are those experts who have the technical know-how. Secondly, the experts have process knowledge, which goes hand in hand with informal or hidden knowledge. The third group of experts consists of interpretive knowledge, which has ideas, ideologies, and explanatory patterns (Kruse et al., 2015). To conduct an open-ended study on the one hand and to validate the results, on the other hand, at least two experts of the same role are interviewed from each role. This constellation should ensure the quality of the statements. In the case of significant deviations within the same role,

an attempt is made to find explanatory approaches, which represents a gain in knowledge.

The expected results of the simulation essentially depend on the identification of the knowledge-based tasks in the IT project, which can be supported with AI methods. With the classic parameters of quality, cost, and time (duration), the cost parameter is probably the only negative evaluation parameter since AI methods will cause an initial effort in the company before they are used. Therefore, the simulation remains to be seen whether AI methods can form a significant added value. On the other hand, the effort could be reduced to a certain extent through cost degression.

5   Future development

Future research fields concerning the current topic can be, on the one hand, the change management process in the planning and implementation of IT projects. In particular, the skills of project managers about artificial intelligence methods would be worth mentioning. Likewise, the training and development of project staff should also be investigated in more detail. Here, there is the possibility of searching for and finding suitable candidates within the own company or relying on externally purchased specialists. However, the introduction and ongoing operation of AI methods for IT projects are not enough. After the processes around project planning and project execution with AI methods have been implemented in knowledge management-relevant areas, it is essential to set oneself apart from the competition by safeguarding innovations in this area and secure the technological lead for as long as possible. The standardization of these processes should then be made as efficient as possible before they are commoditized, and the optimization of IT projects begins anew with new innovative approaches in this area (Moore, 2002).

References

Adler, Paul S. u. a. (From Project to Process Management: An Empirically-Based Framework for Analyzing Product Development Time, 1995): From Project to Process Management: An Empirically-Based Framework for Analyzing Product Development Time, in: Management Science 41 (1995), 3, 458–484, https://doi.org/10.1287/mnsc.41.3.458
Angermann, Anne u. a. (MATLAB - Simulink - Stateflow, 2021): MATLAB - Simulink - Stateflow: Grundlagen, Toolboxen, Beispiele, 10. Auflage, Berlin/Boston: De Gruyter Oldenbourg, 2021
Bassil, Youssef (A Simulation Model for the Waterfall Software Development Life Cycle, 2012): A Simulation Model for the Waterfall Software Development Life Cycle, 2012

Dumslaff, Uwe/Heimann, Thomas (IT-Trends 2018, 2018): IT-Trends 2018 (2018),
     https://www.capgemini.com/de-de/wp-content/uploads/sites/5/2018/02/it-trends-studie-
     2018.pdf

Gasik, Stanislaw (A Model of Project Knowledge Management, 2011): A Model of Project Knowledge
     Management, in: Project Management Journal 42 (2011), 3, 23–44

Gemino, Andrew/Reich, Blaize Horner/Sauer, Chris (A Temporal Model of Information Technology
     Project Performance, 2007): A Temporal Model of Information Technology Project
     Performance, in: Journal of Management Information Systems 24 (2007), 3, 9–44,
     https://doi.org/10.2753/MIS0742-1222240301

Hanisch, Bastian u. a. (Knowledge management in project environments, 2009): Knowledge
     management in project environments, in: J of Knowledge Management 13 (2009), 4, 148–160,
     https://doi.org/10.1108/13673270910971897

Kruse, Jan u. a. (Qualitative Interviewforschung, 2015): Qualitative Interviewforschung: Ein
     integrativer Ansatz, 2., überarbeitete und ergänzte Auflage, Weinheim/Basel: Beltz Juventa,
     2015

Kuckartz, Udo (Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung, 2018):
     Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung, 4. Auflage,
     Weinheim/Basel: Beltz Juventa, 2018

Lech, Przemyslaw (Managing knowledge in IT projects: a framework for enterprise system
     implementation, 2014): Managing knowledge in IT projects: a framework for enterprise system
     implementation, in: Journal of Knowledge Management 18 (2014), 3, 551–573

Lee, Zoonky/Lee, Jinyoul (An ERP implementation case study from a knowledge transfer perspective,
     2000): An ERP implementation case study from a knowledge transfer perspective, in: Journal
     of Information Technology 15 (2000), 4, 281–288,
     https://doi.org/10.1080/02683960010009060

Moore, Geoffrey A. (Living on the fault line, 2002): Living on the fault line: Managing for shareholder
     value in any economy, Rev. ed., New York, NY: HarperBusiness, 2002

Reich, B.H., Gemino, A, Sauer, C. (Modelling the knowledge perspective of IT projects, 2008):
     Modelling the knowledge perspective of IT projects, in: Project Management Journal 2008, 39,
     4–14

The Standish Group International, Inc (CHAOS REPORT 2015, 2015): CHAOS REPORT 2015
     (2015),   https://www.standishgroup.com/sample_research_files/CHAOSReport2015-
     Final.pdf

Wang, E., Lin, C., Jiang, J.J., Klein, G. (Improving enterprise resource planning (ERP) fit to
     organizational process through knowledge transfer, 2007): Improving enterprise resource
     planning (ERP) fit to organizational process through knowledge transfer, in: International
     Journal of Information Management 2007 (2007), 27

TRANSFORMATION OF THE BPMN BUSINESS PROCESS MODEL INTO SMART CONTRACTS FOR THE HYPERLEDGER FABRIC ENVIRONMENT

JANKO HRIBERŠEK

University of Maribor, Faculty of organizational sciences, Kranj, Slovenia; e-mail: jhribersek@gmail.com

Abstract Transformation of BPMN business process model will be a very important topic in the future of the Hyperledger Fabric blockchain environment. Machine transformation can increase the quality of transformation and thus reduce errors. The research paper first describes BPMN and the Hyperledger Fabric environment, what smart contracts are and why they are so important. In the second part, the raw transformation model is described, where the inputs for the transformations are BPMN and the metadata file, and the result is a smart contract written in Java that can be imported into the Hyperledger Fabric environment.

1   Introduction

Blockchain technology is revolutionizing the creation of scalable information systems and diverse applications by incorporating increasingly popular artificial intelligence, cloud computing, and large databases (Lu, 2019). The possibilities of using blockchain technology are being explored by various industries, so it is expected to spread to all sectors of industry. The proposal addresses the problem of machine conversion of a business process model into an executable language with Hyperledger Fabric smart contracts for blocks with permissions, where the model is written in the standard notation for BPMN 2.0 modelling. Solving this problem would significantly reduce the time required to convert a business process into a smart contract. This would reduce conversion costs, increase the security of business processes, and reduce the number of errors when converting to executable code. The adoption of blockchain technology would increase the traceability of the process. There would be no need to create an additional audit trail because records in the blockchain are never deleted. They are only added, and, in this way, it is always possible to revisit all stored values in the blockchain. There are few entries in the literature on this transformation (López-Pintado et al., 2018), and a similar experiment from 2019 (Lopez-Pintado et al., 2019) is related to BPMN and the Ethereum environment. Therefore, in the following problem, we analyzed the properties of BPMN and the blockchain environment with a focus on Hyperledger Fabric.

2   Problem definition

2.1   Business Process Modelling and Notation (BPMN)

The Object Management Group (OMG, 2013) has developed a standard Business Process Model and Notation (BPMN). The main goal of BPMN is to provide a record that is easily understood by all business users, from the business analysts who create the initial process designs to the technical developers responsible for implementing the technology that will implement those processes to the business people who will manage and monitor those processes. BPMN creates a standardized bridge for the gap between business process design and implementation. Another goal of BPMN is to ensure that eXtensible Markup Language (XML) languages designed for business process execution, such as WSBPEL (Web Services Business

Process Execution Language), can be visualized with a business-oriented data set. BPMN provides an easy way to share process information with other companies, users, process contractors, customers, and suppliers.

BPMN is widely used to graphically represent business process artifacts such as start, end, flow, activity, event, and transition. The model, written in BPMN syntax, formally represents a graph with nodes and links. The BPMN syntax also allows the definition of conditions for the transition from one node to another. Any markup language can be used to write it. The Object Management Group used the scalable XML markup language (OMG, 2013) in its development. The markup language itself has no execution capability, so BPMN is used as the top, contextual layer for automated mapping into an executable language. The middle layer of mapping is called Business Process Execution Language (BPEL) and at the time of its creation supported Service Oriented Architecture (SOA) or even more explicitly Web-based Services (WEB service) via the Web Service Description Language (WSDL). The lowest layer is an executable language that can communicate with a WEB service usually this is bytecode translated into Java. The process of orchestrating a business process—that is, transforming BPMN to executable (Java) code is only partially automated, time-consuming, and the risk of errors is high.

2.2 Blockchain technology

In the near future, blockchain technology will become a powerful tool for use in Industry 4.0 as it integrates and brings together architecture, technology, devices, and other related things to deliver high-quality products and services (Lu, 2019). Blockchain will also be important in the upcoming Industry 5.0. It is a new production model based on the interaction between people and machines. It emphasizes the standard of living, creativity and high quality of customized products. (Rada, 2018)

Blockchain technology became interesting with the release of the first cryptocurrency, Bitcoin. The electronic payment system is based on cryptological evidence rather than trust and allows two willing parties to transact directly without the involvement of a trusted third party (Nakamoto, 2008). From the cryptocurrency environment, blockchain technology has leapt to a higher level with the advent of the Ethereum platform, which allows for decentralized application replication. The

novelty of the Ethereum platform was the smart contracts implemented in the Solidity and Vyper languages. The further development of blockchain technology is moving towards private blockchain environments, where users are authenticated when they integrate into the blockchain environment (private environment or permissioned environment).

The immutability and transparency of blockchain technology reduces human error and the need for manual intervention in database conflict situations. Blockchain can help streamline business processes by eliminating the duplication of data management activities (Tuan et al., 2018). Despite its popularity, blockchain technology still faces fundamental problems in managing transactions that are somewhat like those of traditional databases. Sharma et al. (2019) cite the Hyperledger Fabric blockchain platform as an example where parallel transaction processing can be compared to concurrency control mechanisms in traditional databases (Sharma et al., 2019). In a permissioned environment, a smart contract plays an important role. The Hyperledger Fabric platform is one of the representatives of the latest generation of blockchain environments, which aims to upgrade cross-industry technologies. In this cross-industry environment, smart contracts supported by Hyperledger Fabric play an important role.

In addition, the Hyperledger Fabric platform contains many advanced functionalities. It is possible to run distributed applications written in general standard programming languages without depending on the internal cryptocurrency. In previous blockchain platforms, smart contracts were written in specific languages. The flexibility is enabled in Hyperledger Fabric with the blockchain design itself and does not consume resources or reduce hardware performance (Androulaki et al., 2018).

Information solutions based on blockchain technology have been introduced in quite different sectors over the last five years. Solutions have been released in electricity marketing (Knirsch et al., 2019; Silva et al., 2019), finance (Duong-Trung et al., 2019; G. S. Group, 2016; Kabra et al., 2020; Nguyen et al., 2020; Wang et al., 2019), computing (Elghaish et al., 2020), law (Truong et al., 2019), food production and processing (Kumar et al., 2020), manufacturing (Sund et al., 2020), transportation (Naerland et al., 2017), and healthcare (Niu et al., 2020; Tanwar et al., 2020).

2.3 Smart contracts in the Hyperledger Fabric environment

A smart contract is a business logic that operates on a chain of blocks. They can be as simple as updating data or as complex as implementing built-in contract terms. There are two different types of smart contracts (*Hyperledger Architecture, Volume II, Smart Contacts*, n.d.):

- With built-in smart contracts the business logic of network certifiers is established before the network goes live.
- On-chain smart contracts introduce the business logic as a transaction that is added to the blockchain and executed each time the transaction is invoked. With on-chain smart contracts, the execution code of the business logic becomes part of the ledger.

The implementation of smart contracts in the Hyperledger Fabric environment is divided into three segments:

- The input contains the contract code, the transaction request, the possible dependency of the transaction, and the current state of the general ledger.
- The contract interpreter contains the current state of the general ledger and the program code of the smart contract.
- Outputs are generated only if the request was correct and confirmed. The output contains the new status and the side effects of the executed smart contract.

The smart contract layer is responsible for processing transaction requests and determining whether transactions are valid according to the specified business logic (*Hyperledger Architecture, Volume II, Smart Contacts*, n.d.). It validates each request by ensuring compliance with the policy and contract for a given transaction. Invalid requests are rejected and may be excluded from the block depending on the framework.

Transactions are considered either not yet begun or not yet completed. The transaction cannot be completed in parts. This ensures the integrity of transactions. Smart contracts can determine the dependencies between multiple transactions that need to be executed individually. These dependencies can be implicit or explicit.

With implicit reference, it is usually impossible to determine the order of transactions. Bitcoin solves this by repeatedly attempting to use the transaction, which results in transactions being executed one at a time until their preconditions are met. Such behaviour requires a transaction transfer component (pool, mempool) to take care of uselessly expiring transactions. With an explicit reference, the user specifies the transaction identifier.

2.1.1 Communication of a smart contract with other architectural layers

The smart contract layer works closely with the consensus layer. Specifically, the smart contract layer receives a proposal from the consent layer. This proposal specifies which contract to execute, the details of the transaction, including the identity and credentials of the entity requesting the execution of the contract, and any transaction dependencies.

The smart contract layer uses the current general ledger balance and the entry from the consent layer to confirm the transaction.

During transaction processing, the smart contract layer uses the identity service layer to authenticate and authorize the entity requesting the smart contract execution. This ensures two things: The entity is known on the blockchain network and the entity has sufficient access to execute the smart contract. Identity can be provided through a variety of methods: simple identities, ledger-managed identities and credentials, anonymous credentials, or managed identity services from a third-party certificate authority.

After processing the transaction, the smart contract layer returns the transaction acceptance status (transaction accepted or rejected). If the transaction is accepted, the smart-contract layer also returns the regularity certificate, delta state, and any optional ordering instructions required for transaction dependency compliance. Delta states include sets of changes and any side effects that should occur once the partners have successfully completed the transaction.

2.1.2   Integrity and availability of smart contracts

To ensure the integrity and availability of the blockchain network and the smart contract layer, enterprise blockchain networks must control access to certain resources. Since smart contracts are programs, they are exposed to malicious attacks, coding errors, and poor design. Disrupting the execution of program code in any of these areas can compromise the integrity or availability of the blockchain system.

Hyperledger recommends the following four security precautions for use with a smart contract layer to ensure integrity and availability:

- Denial of Service Protection
- Sandboxing
- Resource Management / Flow Control and Application Lifecycle Management (ALM)

In the design phase of this predisposition, we found two recent studies on the successful transformation of BPMN to Ethereum (Lopez-Pintado et al., 2019; López-Pintado et al., 2018), and none on the transformation of Hyperledger Fabric, which is a more technology-aware technology.

3   Methodology

To conduct the research, we will use:

- Study the relevant literature on BPMN and Hyperledger Fabric.
- Software engineering, which will include problem definition, programming and testing using the white box method.
- Verification and validation of the developed solution on selected cases of business processes. Verification (checking conformance to specification) will be performed after each activity that results in a test artifact and validation will be performed with key stakeholders of a given business process.

In the context of software engineering, a core and language for software engineering methods called Essence is used, which was developed by the Object Management Group (Object Management Group, 2018). It is the result of the SEMAT initiative (Software Engineering Methods and Techniques), which states on its website (*Http://Semat.Org*, n.d.) that Essence can be used by researchers to define a problem that needs to be understood and explored as part of an effort to develop a general theory of software engineering (cit. " Researchers can use Essence as a definition of the problem they want to understand and explore in their efforts to develop a General Theory of Software Engineering."). The current version of the Essence core and language is 1.2, with beginnings dating back to 2012 (Jacobson et al., 2012). Ivar Jacobson et. al. (Jacobson et al., 2019) state that software complexity is not the only reason for the "software crisis". They believe that developing solutions is not only about programming, but also about planning activities, managing a group of stakeholders, and effective communication and cooperation (p. 18). Ian Sommerville in the preface of the publication talks about Essence as a metamethod because the universality of the concepts involved allows it to be used in a broader range of domains than current methods, and Grady Booch says that he and his friends Ivar Jacobson and Jim Rumbaugh have tried to help developers with UML (Universal Modelling Language), in which some things are set right, others wrong, and that Essence is considered a set of basic software engineering abstractions. Jacobson et al. (2019) freely admit that software engineering is still in the process of developing a suitable theory that encompasses both descriptive and predictive aspects. Essence is seen as an important descriptive theory, while Tarpit (Johnson & Ekstedt, 2016) is ascribed the role of predictive theory. The core of Essence otherwise includes:

- *Alphas* - descriptions of things we manage, develop, and use in the process of development, maintenance, and support. Alphas can have subalphas,
- *Activity Spaces* - representations of important things that need to be done in the process of developing, maintaining, and supporting software solutions; and
- *Competencies* - a representation of the key skills required for all activities.

In the current state of research (dissertations) are:

- *Alpha* - business process models in BPMN, transformation rules, smart contracts in the chosen programming language (probably Java and/or Javascript), a set of metadata about the business processes under study,
- *Activity spaces* - detailed acquaintance with Hyperledger Fabric, contexts of business process models, variants of developed program code and white-box testing, and
- *Competences* - explicit and implicit knowledge of programming languages BPMN, Hyperledger Fabric, Java and Javascript.

The Essence language specification includes:

- *Headings* - formal names of language elements,
- *Descriptions* - an informal description of a language element,
- *Generalizations* - a structure of classes and parent classes,
- *Attributes* - a list of attributes with their data types,
- *Associations* - a list of links that the language element has,
- *Invariants* - informal verbal descriptions of correctly formed rules and Object Constraint Language (OCL) terms,
- *Additional Operations* - descriptions of additional operations for correctly formed rules, and
- *Semantics* - a detailed description of the elements in natural language.

Language specifications with Essence cannot be defined well enough at the current stage of research, but they will be developed at a later stage.

4   Expected results

The primary research question of the dissertation proposal is:

How can the business process model described in BPMN with certain metadata be automatically and formally converted into the execution code of smart contracts in the Hyperledger Fabric blockchain environment?

In line with the research question, the objectives are:

- investigate BPMN and Hyperledger Fabric.
- in the case of a business process, define the steps and procedures for transforming BPMN into smart contracts.
- test the sequence of steps and procedures for the transformation of several other standard business processes.
- identify the minimum set of metadata relevant for the transformation.
- develop software that formally correctly performs the transformation from BPMN to Hyperledger Fabric smart contracts with permitted access.

The achieved goals of the research will enable the understanding of the process of language conversion for data specification into a Turing universal machine. The developed approach and application will be useful in:

- the basic research of business informatics
- practice for business management, specifically for enhancing business performance.

The transformation of the business process model from BPMN to a smart contract will be presented and tested in detail. A minimal set of metadata is identified to define the details in the Hyperledger Fabric environment.

5 Future development

At this point, development is still in the first stage. Further development will include the following development steps:

- development of a software module for the identification of individual BPMN building blocks (event, activity, gateway, sequence flow)
- development of a transformation library for individual basic BPMN building blocks.

- development of a solution that prepares a transcription of the BPMN model from the XML format into an intermediate format arranged in the order of implementation of the individual activities captured in the BPMN.

- a transformation program that will create software in Java from a list of sequential activities and using a metadata file.

References

Androulaki, E., Barger, A., Bortnikov, V., Cachin, C., Christidis, K., De Caro, A., Enyeart, D., Ferris, C., Laventman, G., Manevich, Y., Muralidharan, S., Murthy, C., Nguyen, B., Sethi, M., Singh, G., Smith, K., Sorniotti, A., Stathakopoulou, C., Vukolić, M., … Yellick, J. (2018). Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains. 1–15. https://doi.org/10.1145/3190508.3190538

Duong-Trung, N., Ha, X. S., Phan, T. T., Trieu, P. N., Nguyen, Q. N., Pham, D., Huynh, T. T., & Le, H. T. (2019). Multi-sessions mechanism for decentralized cash on delivery system. International Journal of Advanced Computer Science and Applications, 10(9), 553–560. https://doi.org/10.14569/ijacsa.2019.0100973

Elghaish, F., Abrishami, S., & Hosseini, M. R. (2020). Integrated project delivery with blockchain: An automated financial system. Automation in Construction, 114(November 2019), 103182. https://doi.org/10.1016/j.autcon.2020.103182

G. S. Group. (2016). Blockchain: The New Technology of Trust. https://www.goldmansachs.com/insights/pages/blockchain/

http://semat.org. (n.d.). Retrieved September 10, 2020, from http://semat.org

Hyperledger Architecture, Volume II, Smart Contacts. (n.d.). Retrieved January 6, 2019, from https://www.hyperledger.org/wp-content/uploads/2017/08/Hyperledger_Arch_WG_Paper_1_Consensus.pdf

Jacobson, I., Lawson, H. "Bud," Ng, P.-W., McMahon, P. E., & Goedicke, M. (2019). The Essentials of Modern Software Engineering: Free the Practices from the Method Prisons! In The Essentials of Modern Software Engineering: Free the Practices from the Method Prisons! Association for Computing Machinery. https://doi.org/10.1145/3277669

Jacobson, I., Ng, P. W., McMahon, P. E., Spence, I., & Lidman, S. (2012). The essence of software engineering: The SEMAT kernel. Queue, 10(10), 40–51. https://doi.org/10.1145/2381996.2389616

Johnson, P., & Ekstedt, M. (2016). The Tarpit - A general theory of software engineering. Information and Software Technology, 70, 181–203. https://doi.org/10.1016/j.infsof.2015.06.001

Kabra, N., Bhattacharya, P., Tanwar, S., & Tyagi, S. (2020). MudraChain: Blockchain-based framework for automated cheque clearance in financial institutions. Future Generation Computer Systems, 102, 574–587. https://doi.org/10.1016/j.future.2019.08.035

Knirsch, F., Unterweger, A., & Engel, D. (2019). Implementing a blockchain from scratch: why, how, and what we learned. Eurasip Journal on Information Security, 2019(1). https://doi.org/10.1186/s13635-019-0085-3

Kumar, A., Liu, R., & Shan, Z. (2020). Is Blockchain a Silver Bullet for Supply Chain Management? Technical Challenges and Research Opportunities. Decision Sciences, 51(1), 8–37. https://doi.org/10.1111/deci.12396

Lopez-Pintado, O., Dumas, M., Garcia-Banuelos, L., & Weber, I. (2019). Interpreted execution of business process models on blockchain. Proceedings - 2019 IEEE 23rd International

Enterprise Distributed Object Computing Conference, EDOC 2019, 206–215. https://doi.org/10.1109/EDOC.2019.00033

López-Pintado, O., García-Bañuelos, L., Dumas, M., Weber, I., & Ponomarev, A. (2018). Caterpillar: A business process execution engine on the Ethereum blockchain. Software - Practice and Experience, 00, 1–45. https://doi.org/10.1002/spe.2702

Lu, Y. (2019). The blockchain: State-of-the-art and research challenges. Journal of Industrial Information Integration, 15(January), 80–90. https://doi.org/10.1016/j.jii.2019.04.002

Naerland, K., Müller-Bloch, C., Beck, R., & Palmund, S. (2017). Blockchain to Rule the Waves - Nascent Design Principles for Reducing Risk and Uncertainty in Decentralized Environments. Icis, 1–16. https://www.researchgate.net/profile/Christoph_Mueller Bloch/publication/319990622_Blockchain_to_Rule_the_Waves_- _Nascent_Design_Principles_for_Reducing_Risk_and_Uncertainty_in_Decentralized_Envir onments/links/59c8df97aca272c71bcdc61f/Blockchain-to-Rule-t

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. Journal for General Philosophy of Science, 39(1), 53–67. https://doi.org/10.1007/s10838-008-9062-0

Nguyen, B. M., Dao, T. C., & Do, B. L. (2020). Towards a blockchain-based certificate authentication system in Vietnam. PeerJ Computer Science, 2020(3). https://doi.org/10.7717/peerj-cs.266

Niu, S., Chen, L., Wang, J., & Yu, F. (2020). Electronic Health Record Sharing Scheme with Searchable Attribute-Based Encryption on Blockchain. IEEE Access, 8, 7195–7204. https://doi.org/10.1109/ACCESS.2019.2959044

Object Management Group. (2018). Kernel and Language for Software Engineering Methods (Essence). 2007 4th IEEE International Workshop on Visualizing Software for Understanding and Analysis, Versión 1.2, 300. https://www.omg.org/spec/Essence/1.2

OMG. (2013). Object Management Group, Business process modeling notation (BPMN) version 2.0.2. Object Management Group, December, 134. https://doi.org/10.1007/978-3-642-33155-8

Rada, M. (2018). INDUSTRY 5.0 definition – Michael Rada – Medium. https://medium.com/@michael.rada/industry-5-0-definition-6a2f9922dc48

Sharma, A., Agrawal, D., Schuhknecht, F. M., & Dittrich, J. (2019). Blurring the lines between blockchains and database systems: The case of hyperledger fabric. Proceedings of the ACM SIGMOD International Conference on Management of Data, 105–122. https://doi.org/10.1145/3299869.3319883

Silva, F. C., Ahmed, M. A., Martínez, J. M., & Kim, Y. C. (2019). Design and implementation of a blockchain-based energy trading platform for electric vehicles in smart campus parking lots. Energies, 12(24). https://doi.org/10.3390/en12244814

Sund, T., Lööf, C., Nadjm-Tehrani, S., & Asplund, M. (2020). Blockchain-based event processing in supply chains—A case study at IKEA. Robotics and Computer-Integrated Manufacturing, 65(March), 101971. https://doi.org/10.1016/j.rcim.2020.101971

Tanwar, S., Parekh, K., & Evans, R. (2020). Blockchain-based electronic healthcare record system for healthcare 4.0 applications. Journal of Information Security and Applications, 50. https://doi.org/10.1016/j.jisa.2019.102407

Truong, N. B., Sun, K., Member, S., Lee, G. M., & Member, S. (2019). GDPR-Compliant Personal Data Management : A Blockchain-Based Solution. 15(March), 1–13.

Tuan, T., Dinh, A., Liu, R., Zhang, M., Chen, G., Dinh, T. T. A., Liu, R., Zhang, M., Chen, G., Ooi, B. C., & Wang, J. (2018). Untangling blockchain: A data processing view of blockchain systems. IEEE Transactions on Knowledge and Data Engineering, 30(7), 1366–1385. https://doi.org/10.1109/TKDE.2017.2781227

Wang, H., Guo, C., & Cheng, S. (2019). LoC — A new financial loan management system based on smart contracts. Future Generation Computer Systems, 100, 648–655. https://doi.org/10.1016/j.future.2019.05.040

KNOWLEDGE RISKS IN DIGITAL SUPPLY CHAINS

PROPOSAL OF A DISSERTATION PROJECT AT THE SCHOOL OF BUSINESS, ECONOMICS AND SOCIAL SCIENCES UNIVERSITY OF GRAZ

JOHANNES P. ZEIRINGER

University of Graz, BANDAS-Center, Universitätsstraße 15 F3, 8010 Graz, Austria;
e-mail: johannes.zeiringer@uni-graz.at,

Abstract The digital transformation changes the way how organizations exchange data in supply chains (SC). Data traditionally shared, is enriched by detailed data sets captured by sensors in the production itself. Advanced data analytic approaches make it possible to extract knowledge from such data sets and thus increase the risk that competitive knowledge unintentionally spills over. From a knowledge management perspective, little attention is paid to such knowledge risks arising from data-centric collaborations. Hence, this proposed PhD project aims at investigating this, by using the overall method of Design Science Research. The project focuses on digital SC, as data-centric collaborations play a central role within them. To contribute to knowledge research, a framework is being sought. The elaborated framework should allow an assessment of knowledge risks and support the selection of suitable measures and it should contribute on how to support the management of knowledge risks in digital SC.

1   Introduction

The digital transformation offers many new opportunities to improve the operation of supply chains (SC) (Vial, 2019). This has led to innovations and changes in different industry sectors and equally affects knowledge management (KM) and supply chain management (SCM) (Schniederjans et al., 2019). Digitalization means the use of digital technologies to change or improve a business model and provide new revenue and value-producing opportunities (Mäkiö et al., 2018). It is not only penetrating SCM increasingly but also, more and more firms are inter-organizational connected and share data along the SC (Kazantsev et al., 2018), (North et al., 2019). From the perspective of knowledge protection, this increasing exchange of comprehensive data sets needs closer attention, because it is a possible gateway to new knowledge risks (Ilvonen et al., 2018), (Durst & Zieba, 2019).

Digitalization enhances the number of connected devices intensely. Implementation of advanced digital technologies (IoT, blockchain, predictive analytics, etc.) determine the digital SC. This results in each partner generating much more data which is shared with collaborators. Also, due to autonomous systems and affordable sensors, the amount of data which is being generated and shared has exploded in the past decade (Spanaki et al., 2018), (Brettel et al., 2014). Sensors in industrial ecosystems control and monitor processes of industrial production and, as part of it, generate and share data continuously (Chen et al., 2016). As a result digital SC emerge, which does not aim at the difference of physical or digital goods or services, but rather how processes within the SC are innovated and changed by modern technologies (Büyüközkan & Göçer, 2018). A digital SC includes a comprehensive exchange of data and is a multi-layered production network that can be flexibly and quickly optimized and (re)composed (Zeiringer J. P. & Thalmann S., 2020).

Knowledge is a key asset within organizations and a source of an organizations competitive advantage (Grant, 1996; Nonaka, 1994). With digital transformation going on, also knowledge management needs to be reopened as new issues arise. Sharing knowledge outside the company, in data-centric collaborations such as alliances, networks, joint ventures or SC partnerships, companies must take protective measures when transferring knowledge across companies, as knowledge risks arise (Krogh, 2012), (Durst & Zieba, 2017). As knowledge is mobile, it is difficult to protect. Especially in collaborations, different people have access to

valuable knowledge (Elliott et al., 2019). It is important that no unintentional outflow of knowledge should take place. Knowledge protection therefore concentrates on (1) preventing knowledge spill-over, (2) reducing the visibility of knowledge and (3) unwanted knowledge spill-over (Manhart & Thalmann, 2015).

Through the intensive exchange of data in inter-organizational collaborations and especially knowledge-intensive collaborations, companies need to find a suitable trade-off between the benefits and risks of collaborations. Research on this trade-off is rare and more research on inter-organizational knowledge transfer, respectively knowledge protection is urgently needed (Hernandez et al., 2015), (Loebbecke et al., 2016), (Manhart & Thalmann, 2015).

As collaboration involves the exchange of data, knowledge risks emerge, especially in data-centric collaborations. Unless these risks are eliminated or managed, they leave a company fragile. Nevertheless, data is a key asset to partners in SC and a source to support SC activities. The goal of data-centric collaborations is to minimize the manual intervention in production processes in order to improve safety, efficiency and sustainability of production through automation (Vyatkin, 2013). With modern data science approaches comprehensive data sets collected from industrial ecosystems, can be continuously analysed to gain useful knowledge for industrial automation (Chen et al., 2016). Hence, SC processes can be optimized, and quality improvements achieved (Kaiser et al., 2020).

2 Problem definition

Traditionally, data for order management and logistics management are exchanged in clearly specified and controllable ways (Min et al., 2019). Regarding digitalization, not only increasingly more data is being exchanged, but this exchange of data is becoming more important for the core operations areas of companies. Modern data analytics methods make it possible and affordable to analyse such data sets and to extract knowledge about these sensitive areas of operation (Schniederjans et al., 2019), (Birkel & Hartmann, 2019). Besides possible benefits of the increased sharing of comprehensive data sets, also risks of losing competitive advantage could arise. Therefore, the risk of losing competitive knowledge through data-centric collaborations in SC is needed to be researched. Furthermore, organizations should

carefully balance their activities to promote and control knowledge sharing, to protect their competitive knowledge (Ilvonen et al., 2018).

Referring to KM, inter-organizational knowledge sharing has a strategic dimension and requires a careful balancing of knowledge sharing and protection as otherwise a loss of competitive knowledge could arise (Loebbecke et al., 2016). Due to, among other things, digitalization, organizational and national boundaries become more blurred and knowledge can be diffused much easier. Openness and inter-organizational collaboration build the foundation of rich, contextualized and sustainable knowledge sharing activities among networked partners within and beyond organizational boundaries (Ilvonen et al., 2018). Referring to knowledge sharing, corporations increasingly rely on the know-how and expertise of external organizations in order to innovate, to remain competitive and to improve performance within the SC (Zacharia et al., 2019).

So far, research focuses mainly on knowledge sharing and protection between persons (representing organizations) in the form of implicit and explicit knowledge exchange (Loebbecke et al., 2016). Little is known about knowledge risks arising from knowledge discovery of huge and comprehensive data sets shared in the course of their digital SC (Ilvonen et al., 2018), (North et al., 2019). In addition, there are efforts to research data and information security, but knowledge protection received little attention so far (Manhart & Thalmann, 2015).

Based on the following observations within this proposal and the current state of research, the research question (RQ) below results:

How to support the management of knowledge risks in digital SC?

3   Methodology

This project makes use of a mixed methods approach. Design science research (DSR) is used as the overall method (Hevner et al., 2004). In the field of IS, the relevance of research is often directly related to the development of IT artefacts (Peffers et al., 2007). DSR is characterized by behavioural and design science. The basic principle in DSR is that knowledge about a real existing problem is gained through the design and evaluation of a solution (Hevner et al., 2004). The result of the research is not only a design-oriented solution, but also a scientific contribution

in the form of frameworks or models (March & Smith, 1995). In order to ensure this contribution to theory, all phases of design science must be rigorously carried out. This requires that both the design proposals and the cause-effect relationships must be empirically evaluated (Iivari, 1991). The research approach will be iterative, with each iteration having elements of (1) identifying and answering problem formulations from the relevant use case, (2) designing artefacts supporting decision making, and (3) elements of rigor, with behavioural theory, and support from IS to KM, SCM, and decision support systems research (A. R. Hevner, 2007). Referring to the stated research problem, the development of a framework, elaborated based on DSR, would be most suitable.

Figure 1: DSR timeline

Figure 1 shows which area the planned papers are assigned to and what their direct and indirect interaction is. The consecutive research papers are listed in the following.

3.1 Paper 1: Structured Literature Review

At first, a structured literature review by Webster and Watson (Webster & Watson, 2002) has been conducted (see chapter 4). In order to elaborate the state of the art in the research field, this is a common process in the information systems (IS) area (Webster & Watson, 2002). Furthermore, a literature review helps to identify the possible research gap. The RQ was regarding which kind of knowledge risks arise from data-centric collaborations and what suitable countermeasures are, see (Zeiringer J. P. & Thalmann S., 2020). The literature review is located at the rigor area within the DSR and an important knowledge base at the beginning of the dissertation project (A. R. Hevner, 2007).

3.2 Paper 2: Interview Study on Knowledge Risk Identification

For this work an interview study by (Patton, 2005) has been conducted (see chapter 4), which is part of the relevance cycle of the DSR (A. R. Hevner, 2007). It is planned to show a detailed requirement analysis for helping to develop the framework. The interview study tried, based on the literature review, to identify different approaches on how to handle knowledge risks in digital SC. Data-centric collaborations were focused, and the balancing of knowledge sharing and protection. There were two staged interviews held with 15 Experts and the elaborated paper has been submitted by now.

Based on the literature review, the risks were theoretically elaborated and analysed; with the interview study, the risks should become more tangible and comprehensible in organizational context. The RQ will be, which knowledge risks arise from data-centric collaborations and which current protection mechanisms are available in order to protect knowledge. Also, it will be shown, if there are already strategies on how to balance sharing and protection and if there are security action plans for what to do after an incident (Thalmann & Ilvonnen, 2020).

3.3   Paper 3: Technical and Organizational Countermeasures

First possible frameworks for technical and organizational countermeasures were deduced from literature and synthesized in a rigor paper. Insights from the first literature review and the interview study were used to develop actionable countermeasures. Also, it is helpful to gather and use theoretical sources to gain creative ideas for the design cycle (A. R. Hevner, 2007).

The RQ is about the possible prevention of unwanted knowledge incidents with help of technical and organizational countermeasures. It also tries to identify measurements that are suitable and easy actionable. The method was the structured literature review, according to (Vom Brocke et al., 2015).

3.4   Paper 4: Questionnaire Survey on Knowledge Risks

Based on the research paper on technical and organizational countermeasures and indirectly the case study, which were carried out in the previous steps, questions for the interviews and online survey can be clearly formulated and the interview study and survey can thus be carried out in a standardised form. The aim of the questionnaire study is to get more details on the problems to be investigated, regarding the identification of them and current protection mechanisms. The target group are experts: SC managers, risk managers or managing directors. After developing first countermeasures, the survey will cover the field of relevance within the DSR again (A. R. Hevner, 2007). The RQ will focus on how organizations are currently deal with arising knowledge risks, if knowledge risks in digital SC represent a barrier to digitalization and to what extent training can help identify knowledge risks in data-centric collaborations. Also, there will be a focus on how to support employees in recognizing knowledge risks in data sets.

Together with the insights of the reviews and the interview study, a first requirement analysis will be conducted. In order to construct a framework, it is necessary to focus on the design cycle after the case study (A. R. Hevner, 2007). It is important to note first intermediated findings and develop a first design concept for needed requirements. This will be processed in an internal working paper and helps to set focus on the fifth paper. The RQ will be to define first requirements for an effective knowledge protection management framework in digital SC. The method will be

user-centred design, according to (Chadia Abras et al., 2004). This will happen simultaneously to the questionnaire study and be a preparation before going into paper 5.

3.5 Paper 5: Framework for Balancing Knowledge Sharing and Knowledge Protection

Based on the findings by then, a framework for balancing knowledge sharing and knowledge protection will be developed. Within this design part of the project, all requirements gathered so far will be processed for this paper (A. R. Hevner, 2007). The main focus will be on the extent to which technical and organizational measures can be used to manage knowledge risks in digital SC.

The RQ will be on which technical or organizational measures can be used to manage knowledge sharing and protection in digital SC and how should a framework be designed to be successfully implemented. The method will be the user-centred design again, according to (Chadia Abras et al., 2004).

3.6 Paper 6: Evaluation of Framework

Finally, the evaluation of the framework will be conducted which, referring to the DSR, is assigned to the relevance circle again (A. R. Hevner, 2007). In DSR, it is important to test the developed artefact in the field, to see if it is appropriate. The results will show, if the artefact is suitable or another iteration is needed (A. R. Hevner, 2007). The evaluation should be executed by an evaluation study which is based directly on the developed framework and the help of the insights gained from the experts. The possible RQ and will potentially be, if the developed framework increases the decision quality of managing knowledge risks within digital SC. The method will be a two staged interview study followed by an (online) survey, according to (Bortz & Döring, 2006).

4 Preliminary/Expected results

As a first step, the state of the art had to be raised. Therefore, a literature review according to (Webster & Watson, 2002) was conducted and processed in a prime paper, see (Zeiringer J. P. & Thalmann S., 2020). In the paper, knowledge risks in data-centric collaborations as part of digital SC were dealt with. Traditional SC risk

management was used to identify causes of risks, risks themselves and potential countermeasures, which were then adapted to digital SC. One of the main insights of this review was that there is little research regarding the field of knowledge risks in data-centric collaborations, which indicated that there is a demand for further research on this main aspect of digital transformation. Furthermore, data-centric collaboration itself is not adequately dealt with so far, as there is still a focus on traditional risks and hardly on intangible risks. It was discovered that there is need for a knowledge risk management and that future research should investigate which kind of measures are meaningful to balance knowledge sharing and protection in data-centric collaborations (Zeiringer J. P. & Thalmann S., 2020). In addition, research shows that the resulting uncertainty creates a barrier to digitalization (North et al., 2019).

The Interview study showed that organizations use different approaches in data-centric collaborations to encounter knowledge risks. It is shown that all three approaches lead to different perspectives of sharing and protection of knowledge within the digital SC. The approaches can be viewed as steps of development, each as one step further in building awareness on knowledge risks and to balance knowledge sharing and protection more holistic. Furthermore, it is shown that minimizing risk can stifle innovation and there is a need for more research [being reviewed].

The second literature review deduced possible actions from literature, to show what is available and what is still missing in order to tackle knowledge risks in data-centric collaborations. In order to build on this and contribute to knowledge research, a framework will be sought after this. The elaborated framework should allow an assessment of knowledge risks and support the selection of suitable measures in practice. It should support the responsible person in the sense of decision support but should not automate the decision (Alter, 2004). With regard to DSR, several cycles of design, evaluation in practice and theoretical reflection should provide a solution to the problem rather than just explore it. Possible developed artefacts could be, e.g., selection lists, visualizations, algorithms or practices. Risks resulting from data exchange can be managed by organizational, technical and/or legal measures. The proposed research project uses this subdivision as a starting point and investigates the simultaneous management of knowledge sharing and knowledge protection in digital SC.

Regarding the theory of knowledge sharing, new categories of knowledge risks that emerge from the growing need to share larger and more comprehensive data sets from which competitive knowledge can be discovered, should be identified and investigated. Also, the data-centric perspective will provide new insights to knowledge sharing theory as well as knowledge risk management. An appropriate strategy to manage knowledge risks, taking data-centric collaborations into account, will be sought (Zeiringer J. P. & Thalmann S., 2020). The expected contribution should be a framework on how to support the management of knowledge risks in digital SC.

5   Future development

The whole project will be split into seven papers. The first paper was a literature review, which has already been accepted to the conference *Wirtschaftsinformatik 2020* (Zeiringer J. P. & Thalmann S., 2020).

The second paper was an interview study (Patton, 2005). Slightly delayed, the third paper, a literature research about technical countermeasures by (Vom Brocke et al., 2015), was written and is submitted in the begin of 2021. At the same time, the planned survey will be conducted and processed in a fourth paper by mid-2021. The elaboration on a working paper starts in Spring 2021, which will help to define design requirements. The final framework is planned to be processed in a paper by spring 2022. Finally, the evaluation of the framework starts in 2022 and ends in June 2022, by submitting the sixth paper.

References

Alter, S. (2004). A work system view of DSS in its fourth decade. Decision Support Systems, 38(3), 319–327. https://doi.org/10.1016/j.dss.2003.04.001

Birkel, H. S., & Hartmann, E. (2019). Impact of IoT challenges and risks for SCM. Supply Chain Management: An International Journal, 24(1), 39–61. https://doi.org/10.1108/SCM-03-2018-0142

Bortz, J., & Döring, N. (2006). Forschungsmethoden und Evaluation. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-33306-7

Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, M. (2014). How Virtualization, Decentralization and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective. International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering(8), 37–44.

Büyüközkan, G., & Göçer, F. (2018). Digital Supply Chain: Literature review and a proposed framework for future research. Computers in Industry, 97, 157–177. https://doi.org/10.1016/j.compind.2018.02.010

Chadia Abras, Diane Maloney-krichmar, & Jenny Preece (2004). User-Centered Design. In In Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications. Publications.

Chen, Y., Lee, G. M., Shu, L., & Crespi, N. (2016). Industrial Internet of Things-Based Collaborative Sensing Intelligence: Framework and Research Challenges. Sensors (Basel, Switzerland), 16(2), 215. https://doi.org/10.3390/s16020215

Durst, S., & Zieba, M. (2017). Knowledge risks - towards a taxonomy. International Journal of Business Environment, 9(1), Article 84705, 51. https://doi.org/10.1504/IJBE.2017.084705

Durst, S., & Zieba, M. (2019). Mapping knowledge risks: towards a better understanding of knowledge management. Knowledge Management Research & Practice, 17(1), 1–13. https://doi.org/10.1080/14778238.2018.1538603

Elliott, K., Patacconi, A., Swierzbinski, J., & Williams, J. (2019). Knowledge Protection in Firms: A Conceptual Framework and Evidence from HP Labs. European Management Review, 16(1), 179–193. https://doi.org/10.1111/emre.12336

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. Strategic Management Journal, 17(S2), 109–122. https://doi.org/10.1002/smj.4250171110

Hernandez, E., Sanders, W. G., & Tuschke, A. (2015). Network Defense: Pruning, Grafting, and Closing to Prevent Leakage of Strategic Knowledge to Rivals. Academy of Management Journal, 58(4), 1233–1260. https://doi.org/10.5465/amj.2012.0773

Hevner, March, Park, & Ram (2004). Design Science in Information Systems Research. MIS Quarterly, 28(1), 75. https://doi.org/10.2307/25148625

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems, 87–92.

Iivari, J. (1991). A paradigmatic analysis of contemporary schools of IS development. European Journal of Information Systems, 1(4), 249–272. https://doi.org/10.1057/ejis.1991.47

Ilvonen, I., Thalmann, S [Stefan], Manhart, M., & Sillaber, C. (2018). Reconciling digital transformation and knowledge protection: a research agenda. Knowledge Management Research & Practice, 16(2), 235–244. https://doi.org/10.1080/14778238.2018.1445427

Kaiser, R., Thalmann, S [Stefan], & Pammer-Schindler, V. (2020). An Investigation of Knowledge Protection Practices in Inter-organisational Collaboration. Protecting Specialised Engineering Knowledge with a Practice Based on Grey-box Modelling. VINE. Advance online publication. https://doi.org/10.1108/VJIKMS-11-2019-0180

Kazantsev, N., Pishchulov, G., Mehandjiev, N., Sampaio, P., & Zolkiewski, J. (2018). Formation of Demand-Driven Collaborations between Suppliers in Industry 4.0 Production Networks. 20th International Working Seminar on Production Economics.

Krogh, G. von (2012). How does social software change knowledge management? Toward a strategic research agenda. The Journal of Strategic Information Systems, 21(2), 154–164. https://doi.org/10.1016/j.jsis.2012.04.003

Loebbecke, C., van Fenema, P. C., & Powell, P. (2016). Managing inter-organizational knowledge sharing. The Journal of Strategic Information Systems, 25(1), 4–14. https://doi.org/10.1016/j.jsis.2015.12.002

Mäkiö, J., Miroliubov, A., & Zhgun, V. (2018). Digitalization – quo vadis? SHS Web of Conferences, 44, 56. https://doi.org/10.1051/shsconf/20184400056

Manhart, M., & Thalmann, S [Stefan] (2015). Protecting organizational knowledge: a structured literature review. Journal of Knowledge Management, 19(2), 190–211. https://doi.org/10.1108/JKM-05-2014-0198

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. Decision Support Systems, 15(4), 251–266. https://doi.org/10.1016/0167-9236(94)00041-2

Min, S., Zacharia, Z. G., & Smith, C. D. (2019). Defining Supply Chain Management: In the Past, Present, and Future. Journal of Business Logistics, 40(1), 44–55.

https://doi.org/10.1111/jbl.12201

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. Organization Science, 5(1), 14–37. https://doi.org/10.1287/orsc.5.1.14

North, K., Carvalho, A. de, Braccini, A., Durst, S., Carvalho, J., Gräslund, K., & Thalmann, S [S.] (2019). Information and knowledge risks in supply chain interactions of SMEs: Proceedings of the 10th International Conference on Practical Knowledge Management, Potsdam, Germany. Lecture notes on Informatics.

Patton, M. Q. (2005). Qualitative Research. In B. Everitt & D. C. Howell (Eds.), Encyclopedia of statistics in behavioral science. Wiley. https://doi.org/10.1002/0470013192.bsa514

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Schniederjans, D. G., Curado, C., & Khalajhedayati, M. (2019). Supply chain digitisation trends: An integration of knowledge management. International Journal of Production Economics, 107439. https://doi.org/10.1016/j.ijpe.2019.07.012

Spanaki, K., Gürgüç, Z., Adams, R., & Mulligan, C. (2018). Data supply chain (DSC): research synthesis and future directions. International Journal of Production Research, 56(13), 4447–4466. https://doi.org/10.1080/00207543.2017.1399222

Thalmann, S [S.], & Ilvonnen, I. (2020). Why should we investigate knowledge risks incidents? - Lessons from four cases. Proceedings of 53rd Hawaii International Conference on System Sciences.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

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. https://doi.org/10.17705/1CAIS.03709

Vyatkin, V. (2013). Software Engineering in Industrial Automation: State-of-the-Art Review. IEEE Transactions on Industrial Informatics, 9(3), 1234–1249. https://doi.org/10.1109/TII.2013.2258165

Webster, J., & Watson, R. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly, 26. https://doi.org/10.2307/4132319

Zacharia, Z., Plasch, M., Mohan, U., & Gerschberger, M. (2019). The emerging role of coopetition within inter-firm relationships. The International Journal of Logistics Management, 30(2), 414–437. https://doi.org/10.1108/IJLM-02-2018-0021

Zeiringer J. P., & Thalmann S. (2020). Knowledge Risks in Digital Supply Chains: A Literature Review. In: Proceedings of Wirtschaftsinformatik 2020 (WI 2020).

Conceptual Model for SMEs' Data Maturity Assessment

Blaž Gašperlin

University of Maribor, Faculty of Organizational Sciences, 4000 Kranj, Slovenia, e-mail: blaz.gasperlin1@um.si

Abstract Digital transformation has brought about a rapid shift towards a completely digital enterprise, generating a huge amount of data. Most small and medium-sized enterprises (SMEs) have data stored in different places, formats, and systems, or are unaware that it exists (Dark Data). While digital technologies are at the root of rapid data growth within and outside organizations, sharing and exchanging data between organizations presents an additional challenge. We argue that one of the barriers to the successful digital transformation of SMEs is data immaturity. The concept of data maturity has been addressed from different aspects (data quality, governance,...), in specific domains (supply chain management, manufacturing companies,...) and from the perspective of the Capability Maturity Model. However, there has been no study that has addressed a comprehensive assessment of data maturity for the SME sector as a multi-criteria problem. In this research, we propose to combine the ideas of maturity models and multi-criteria decision modeling by using a design science research approach. The developed model will help SMEs assess their data maturity level and help them understand what aspects of data maturity they need to advance, what steps they need to take, and how to evaluate their progress.

1 Introduction

Digital transformation has caused the widespread creation of digital products and services and initiated the transformation from traditional business towards the completely digital one. Digital transformation refers to changes in the way companies operate and create value (new business models, products, and services) and the way they communicate, using digital technologies (Morakanyane et al., 2017; Vial, 2019). To digitally transform, companies need to respond and adapt to these changes by creating an appropriate organizational climate (open communication, experimentation, and agility), investing in digital capabilities, developing employees, and engaging in effective knowledge management supported by informed and timely decision-making (Kljajić Borštnar & Pucihar, 2021).

One of the important elements of digital transformation is data (Mitra et al., 2019), which is becoming an important strategic resource in the organization. The amount of data collected and generated is rapidly increasing within and outside organizations. Data can be a competitive advantage, if managed properly. On the other hand, data and data management technologies, when not managed properly, present high costs. With no identified added value in the end, data can end up being a frustration for the organization (Ahlstrom, 2019; Sharma, 2020). The ability to utilize internal and external data to its fullest potential affects the ability to digitally transform as a whole (Kotsev et al., 2020). »Data-driven« is so intertwined with digital transformation, that some authors use the term data-driven digital transformation (Capgemini, 2018), or even data-driven transformation (Someh & Wixom, 2017). In our thesis paper, we understand that data is an essential part of the digitalization process, which results in digital transformation. Further, we define a data-driven organization as an organization that collects, generates, stores, manages, uses, and controls its data in a comprehensive manner, to support its daily operations and decision-making. Such organizations can be referred to as data mature organizations.

We derive our proposition from the idea, that data maturity is crucial to achieve broader goals of digital transformation.

2   Previous literature

Research on data-driven transformation is growing, however, we have noticed a lack of literature that is more oriented towards data maturity of organizations, with a particular focus on data maturity models in relation to SMEs. The previous literature is mainly oriented on data-driven business model innovation (Cheah & Wang, 2017; Marcinkowski & Gawin, 2020), data analytics and its capabilities (Carvalho et al., 2019; Dremel et al., 2017; O'Donovan et al., 2016), and guidelines or directions, how to become a data-driven organization (Anderson, 2015; Berntsson Svensson & Taghavianfar, 2020). The data maturity models, observed in the literature are either focused on a specific area of data maturity, too generic or focused on large organizations. Nevertheless, most elements of the available models can be also applied to small and medium-sized enterprises, to provide them the starting point for data maturity assessment.

(Sen et al., 2006) focused on the maturity of the data warehousing process and the identification of the influencing factors following the Capability-Maturity Model (CMM), which could help characterize the corresponding maturity levels. They found that the most prevalent factors were data architecture, supported by online analytical processing (OLAP), business analytics and its alignment with business strategy, data quality, organizational readiness and resources (human, financial and technical), change management, and data warehouse size. Change management relates to changes and adaptations of data mechanisms and technologies, and OLAP refers to the analytical tool that enables analysts to gain meaningful information and insights from a vast and diverse variety of data, stored in databases (Moon et al., 2007). The results also showed that analytical culture is one of the most important factors to consider when assessing data warehouse maturity.

(Rivera et al., 2017) developed the data governance maturity model for micro-sized organizations and validated it on the case of a financial organization. The results showed that organizational culture, data processing and analysis, data integration, and interoperability are the main issues in micro-sized organizations, that should be given more emphasis to achieve greater data governance maturity.

(Sternkopf & Mueller, 2018) developed a model to help organizations assess their level of data literacy based on 3 main dimensions - data culture, data ethics and security, and level of data manipulation (retrieval, verification, analysis, visualization, evaluation, and its interpretation). Others, (Loshin, 2011) proposed the maturity model for assessing data quality. The proposed model assesses data quality based on eight dimensions (data expectations, measures, the level of established data policies, standards and procedures, applied data governance mechanisms, the level of technology and tools, data auditing, and reporting). Data expectations indicate, how much the organization emphasizes the importance of data quality. Similarly, (Al-Sai et al., 2020) developed a classification framework for the factors that organizations can consider when implementing Big Data analytics.

3   Problem definition

Data has been an important driver of change since expert systems in decision making (Mandinach et al., 2006; Power, 2008; Provost & Fawcett, 2013), but less attention has been paid to the role of data in assessing data maturity. Every organization, including small and medium-sized enterprises (SMEs), needs quality data and needs to know, how to manage it comprehensively, if they want to be competitive in the market. Most SMEs have their data scattered in different information systems and thus have a structure with a low level of interoperability, a low level of data governance applied, and a lack of good auditing and traceability of the data they collect and generate internally or externally.

Even though the organizations began to exploit the use of data more effectively in their digital transformation process, small and medium-sized enterprises (SMEs) are lagging behind. This is evident from the data provided by (Eurostat, 2021), which shows that more than half of SMEs do not use data analytics to analyze the data they collect and further extract the value that the data they collect could provide. This shows that most SMEs are not aware of the important role of data, nor do they have a defined approach to data management and governance that would address this issue in a systematic and comprehensive manner. It is important for SMEs to first identify what data they collect and generate, what quality of data they have, and for what purpose they are currently using it. To help them do this, various digital maturity models and tools have been developed. In addition, data maturity frameworks and models have been developed to help companies analyze the current

state of the data they collect and generate, assess the level of their data usage and management, and serve as a guide for implementing the necessary steps towards becoming a data-driven organization and towards digital transformation in general.

We start from the proposition that it is possible to assess the data maturity of SMEs. This will help them to 1) better understand what to do with the data they have and collect and what steps to take to make better use of the data, and 2) to later assess their progress.

This dissertation will focus specifically on small and medium-sized enterprises and their assessment of data maturity. It will discuss the role of data as a key driver of change in digital transformation and as a foundation for quality-based data-driven decision-making.

The previous literature (Rivera et al., 2017; Sen et al., 2006; Sternkopf & Mueller, 2018) proposed a few data maturity models or frameworks, focusing on a specific area of data maturity (such as data warehouse maturity) or single aspects (data quality, data governance, data stewardship, etc.). None of the data maturity models examined, address the needs of SMEs. SMEs usually lack resources (financial, human, time, skills), so they need a comprehensive, systematic, and easy-to-use tool to help them assess the state of data and understand further steps towards data maturity. We need to consider that the problem of data maturity is multi-faceted. In order to assess the data maturity of an organization, we need to consider several criteria and address the problem by using multi-criteria decision modeling. The multi-criteria models are used to evaluate the alternatives (in this case, individual SMEs) (Mardani et al., 2015) and help us to set the proper criteria (in our case, criteria for data maturity assessment).

The aim of this paper is to develop a multi-criteria data maturity assessment model for small and medium-sized enterprises, to evaluate different levels of data maturity, and help SMEs to achieve a more systematic and comprehensive data management and governance, which will contribute to further digital transformation of SMEs. Based on the identified data maturity level, the proposed model will highlight the gaps in the areas they need to invest more, to raise their data maturity level, and start implementing the necessary steps to realize this issue. The results of the work will contribute to the modeling of decision knowledge in the field of data maturity

assessment and will be useful for creating policies and strategies in the field of data science.

4 Methodology

We will use a design science research (DSR) approach (Hevner et al., 2004), within which we will develop an IT artifact - a multi-criteria data maturity model. Design science research relates to the iterative sequence of expert activities, to produce an innovative product (artifact). The artifacts can be constructs, methods, instantiations, or models. The evaluation of the artifact provides feedback and a better understanding of the problem, which allows us to improve the quality and the design of the artifact (Hevner et al., 2004, p. 78).

First, we will conduct a literature review with a focus on a data-driven digital transformation field, data maturity, data-driven decision making, data-driven organization, data governance models, and related fields. From the selected literature, we will in addition review the references in the reference list and selected those relevant for our research. In the next step, we will conduct interviews with selected small and medium-sized enterprises, to identify the actual problems the SMEs have in practice and are related to the data maturity problem. This will give us a deeper insight into what they have already achieved in the data maturity field and how they currently address this issue and feedback to our existing knowledge and literature.

In doing so, we will follow the three main cycles rooted in the DSR research, represented by the rigor, relevance, and design cycle (Figure 1). Based on the reviewed literature and conducted interviews, we will define the criteria needed for data maturity assessment, design data maturity levels, and develop a multi-criteria data maturity model (the design cycle), using DEX methodology and the gathered insights from the interviews. DEX (Bohanec et al., 2013) is a qualitative multi-attribute decision-making method, implemented in freely available software for multi-attribute decision making, DEXi (Bohanec, 2021).
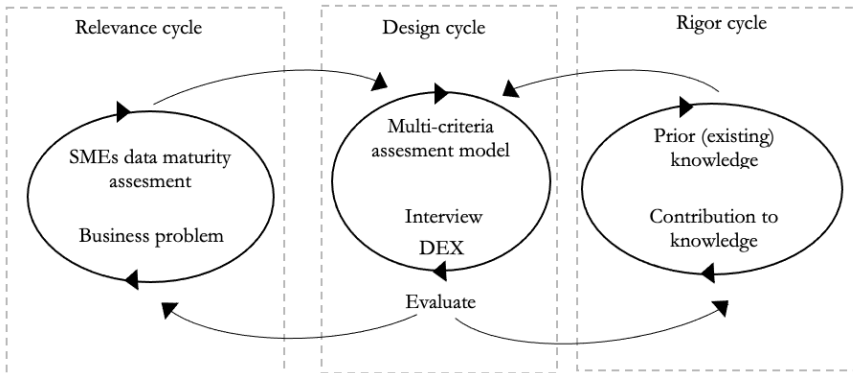
Figure 1: Methodology process - adapted from (Hevner, 2007)

The model will be validated on the real cases of SMEs. After the model validation, we will re-design the model, if needed, and extract the findings. Based on the current observation from the literature and the identification of a real (business) problem in practice, we will try to answer the following research questions: RQ1: "*What are the criteria that will allow us to assess the level of data maturity in SMEs?*" and RQ2: "*Is it possible to develop a sensitive multi-criteria model, that will well enough separate data maturity levels?*".

5   Preliminary/Expected results and future development

Future development in our research will focus on criteria identification, which will help us to develop the multi-criteria model to assess data maturity. We expect that the developed and validated model will be used in practice in the real environment and will thus represent a contribution to science and the real business environment. In addition, we expect that it will be possible to transfer the developed model to other businesses (not only to small and medium-sized enterprises) and to include the use of the questionnaire, which could increase the usability of the developed model for a larger number of organizations. In our next steps, we will validate the identified criteria from the literature with interviews in selected small and medium-sized enterprises. We expect that the findings from the interviews will contribute to a better understanding of the problem.

Acknowledgements

References

Ahlstrom, R. (2019). The Role Of Data In The Age Of Digital Transformation. https://www.forbes.com/sites/forbestechcouncil/2019/01/17/the-role-of-data-in-the-age-of-digital-transformation/?sh=68bad9994509

Al-Sai, Z. A., Abdullah, R., & Heikal Husin, M. (2020). Critical Success Factors for Big Data: A Systematic Literature Review. IEEE Access, 8, 118940–118956. https://doi.org/10.1109/ACCESS.2020.3005461

Anderson, C. (2015). Creating a Data-Driven Organization: Practical Advice from the Trenches. In T. McGovern (Ed.), O'Reily (First). O'Reilly Media, Inc. https://books.google.si/books?hl=en&lr=&id=MVpDCgAAQBAJ&oi=fnd&pg=PR2&dq=data-driven+organization&ots=Ij157j1YCV&sig=mdFzH1I90XUzGdBvtgH7SWhAFG0&redir_esc=y#v=onepage&q=data-driven organization&f=false

Berntsson Svensson, R., & Taghavianfar, M. (2020). Toward Becoming a Data-Driven Organization: Challenges and Benefits. In F. Dalpiaz, J. Zdravkovic, & P. Loucopoulos (Eds.), Research Challenges in Information Science (pp. 3–19). Springer International Publishing. https://doi.org/10.1007/978-3-030-50316-1_1

Bohanec, M. (2021). DEXi: A Program for Multi-Attribute Decision Making. https://kt.ijs.si/MarkoBohanec/dexi.html

Bohanec, M., Martin, Ž., Vladislav, R., Ivan, B., & Blaž, Z. (2013). DEX Methodology: Three Decades of Qualitative Multi-Attribute Modeling. Informatica, 37(1), 49–54. https://www.researchgate.net/publication/280805251_DEX_Methodology_Three_Decades_of_Qualitative_Multi-Attribute_Modeling

Capgemini. (2018). Data-driven digital transformation. https://www.capgemini.com/cn-zh/wp-content/uploads/sites/12/2018/06/Data-driven-digital-transformation-8.pdf

Carvalho, J. V., Rocha, Á., Vasconcelos, J., & Abreu, A. (2019). A health data analytics maturity model for hospitals information systems. International Journal of Information Management, 46, 278–285. https://doi.org/10.1016/j.ijinfomgt.2018.07.001

Cheah, S., & Wang, S. (2017). Big data-driven business model innovation by traditional industries in the Chinese economy. Journal of Chinese Economic and Foreign Trade Studies, 10(3), 229–251. https://doi.org/10.1108/JCEFTS-05-2017-0013

Dremel, C., Overhage, S., Schlauderer, S., & Wulf, J. (2017). Towards a Capability Model for Big Data Analytics. 13th International Conference on Wirtschaftsinformatik (WI 2017), 1141–1155. https://www.researchgate.net/publication/311510196_Towards_a_Capability_Model_for_Big_Data_Analytics

Eurostat. (2021). Big data analysis. http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems, 19(2), 1–6. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1017&context=sjis

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. MIS Quarterly, 28(1), 75–105. https://doi.org/10.2307/25148625

Kljajić Borštnar, M., & Pucihar, A. (2021). Multi-Attribute Assessment of Digital Maturity of SMEs. In Electronics (Vol. 10, Issue 8). https://doi.org/10.3390/electronics10080885

Kotsev, A., Minghini, M., Tomas, R., Cetl, V., & Lutz, M. (2020). From Spatial Data Infrastructures to Data Spaces—A Technological Perspective on the Evolution of European SDIs. In ISPRS International Journal of Geo-Information (Vol. 9, Issue 3). https://doi.org/10.3390/ijgi9030176

Loshin, D. (2011). Data quality maturity. In J. Niles & D. Bevans (Eds.), The Practitioner's Guide to Data Quality Improvement (pp. 44–51). Morgan Kaufman - Elsevier. https://books.google.si/books?hl=en&lr=&id=B3zd4GCAWeYC&oi=fnd&pg=PR3&dq=info:qMsybD7zSZUJ:scholar.google.com&ots=0SVyX2ZfGX&sig=Hux9Ymbl0emIu8MJTCqmY3CiheU&redir_esc=y#v=onepage&q&f=false

Mandinach, E. B., Honey, M., & Light, D. (2006, April 9). A Theoretical Framework for Data-Driven Decision Making. http://www.cct.edc.org/sites/cct.edc.org/files/publications/DataFrame_AERA06.pdf

Marcinkowski, B., & Gawin, B. (2020). Data-driven business model development-insights from the facility management industry. Journal of Facilities Management, 1–21. https://doi.org/10.1108/JFM-08-2020-0051

Mardani, A., Jusoh, A., Nor, K., Khalifah, Z., Zakwan, N., & Valipour, A. (2015). Multiple criteria decision-making techniques and their applications-a review of the literature from 2000 to 2014 Multiple criteria decision-making techniques and their applications-a review of the literature from 2000 to 2014. Economic Research-Ekonomska Istraživanja, 28(1), 516–571. https://doi.org/10.1080/1331677X.2015.1075139

Mitra, A., Gaur, S. S., & Giacosa, E. (2019). Combining organizational change management and organizational ambidexterity using data transformation. Management Decision, 57(8), 2069–2091. https://doi.org/10.1108/MD-07-2018-0841

Moon, S. W., Kim, J. S., & Kwon, K. N. (2007). Effectiveness of OLAP-based cost data management in construction cost estimate. Automation in Construction, 16(3), 336–344. https://doi.org/10.1016/j.autcon.2006.07.008

Morakanyane, R., Grace, A., & O'Reilly, P. (2017). Conceptualizing Digital Transformation in Business Organizations: A Systematic Review of Literature. 30TH BLED ECONFERENCE: DIGITAL TRANSFORMATION – FROM CONNECTING THINGS TO TRANSFORMING OUR LIVES, 427–444. https://doi.org/10.18690/978-961-286-043-1.30

O'Donovan, P., Bruton, K., & O'Sullivan, D. T. J. (2016). IAMM: A Maturity Model for Measuring Industrial Analytics Capabilities in Large-scale Manufacturing Facilities. International Journal of Prognostics and Health Management, 1–11. https://cora.ucc.ie/bitstream/handle/10468/5396/4005.pdf?sequence=3&isAllowed=y

Power, D. J. (2008). Understanding data-driven decision support systems. Information Systems Management, 25(2), 149–154. https://doi.org/10.1080/10580530801941124

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. Big Data, 1(1), 51–59. https://doi.org/10.1089/big.2013.1508

Rivera, S., Loarte, N., Raymundo, C., & Dominguez, F. (2017). Data Governance Maturity Model for Micro Financial Organizations in Peru. Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017), 203–214. https://doi.org/10.5220/0006149202030214

Sen, A., Sinha, A. P., & Ramamurthy, K. R. (2006). Data Warehousing Process Maturity: An Exploratory Study of Factors Influencing User Perceptions. IEEE Transactions on Engineering Management, 53(3), 440–455. https://doi.org/10.1109/TEM.2006.877460

Sharma, S. (2020, December 3). Data Is Essential To Digital Transformation. https://www.forbes.com/sites/forbestechcouncil/2020/12/03/data-is-essential-to-digital-transformation/?sh=527b1afa26c9

Someh, I. A., & Wixom, B. H. (2017, August 17). Data-Driven Transformation at Microsoft. https://cisr.mit.edu/publication/2017_0801_DataDrivenTransformation_SomehWixom#fn_2

Sternkopf, H., & Mueller, R. M. (2018). Doing Good with Data: Development of a Maturity Model for Data Literacy in Non-governmental Organizations. Proceedings of the 51st Hawaii International Conference on System Sciences, 5045–5054. https://core.ac.uk/download/pdf/143481465.pdf

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

IMPACT ASSESMENT OF OPEN GOVERNMENT DATA

ALJAŽ FERENCEK

University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia; e-mail: aljaz.ferencek1@student.um.si

Abstract Public sector organizations produce and process increasing amounts of data and the number of research and initiatives on open data is also increasing. Defining the true value of OGD is challenging without knowing how it impacts society and its economy. While the analysis of the economic benefits of open data is one way to describe the effect of government openness, the impact of open data is measured also in social and political context. Feedback mechanisms that are currently used are mostly surveys, while the number of OGD use cases is increasing. This paper proposes a preliminary model for research on assessing impact areas of OGD in an automated manner by using text mining techniques on existing use cases.

1 Introduction

Open Government Data (OGD) is one of the sources of open data published by the public sector in form of databases, with the aim of promoting transparency, accountability, and creating added value (Open Government Data, n.d.). Public sector organizations produce and process increasing amounts of data, and the number of research and initiatives on open data is also increasing (Attard et al., 2016; Attard et al., 2015; Safarov et al., 2017; Ubaldi, 2013; Yan & Weber, 2018). By making their databases accessible, public institutions are becoming more attractive for political participation of citizens (Ruijer & Martinius, 2017), the creation of companies and innovative services focused on citizens is encouraged (Pereira et al., 2016), and the long-term goal is to ensure overall transparency of government information (Jaeger & Bertot, 2010). As Ubaldi (2013) points out, one of the elements for creating added value of OGD is high-impact data for the public. Currently, there are many open data policies at different levels of government, while very few systematic and structured studies have been conducted on their actual impact (Roa et al., 2019; Ruijer & Martinius, 2017; Zuiderwijk & Janssen, 2014). Many other authors (Afful-Dadzie E. & Afful-Dadzie A., 2017; Crusoe et al., 2019; Safarov et al., 2017; Wilson & Cong, 2020) also recognize the problem of measuring the impact of open data, and the reasons for that are mostly in the availability and low quality of data, costs and legal barriers and in users. Although not much attention has been paid to the impact of open data in the past, many useful solutions have emerged in recent years. As stated by Kalampokis et al. (2013), companies use open data and methods of business intelligence which help them survive in the global economy, open data helps designers to make better policies and academics to analyze data for hypotheses testing, understanding of patterns and predictions. For this reason, we recognized an opportunity to identify areas of open government data impacts from the actual use cases of open government data through text mining, and thus enable governments to create targeted and better-quality data sets.

The use of text mining is quite common in the literature, but not for the purpose of identifying impact areas of OGD. Common applications of text mining are topic modeling using Latent Dirichlet Allocation (LDA) which can be used to identify areas of OGD (Afful-Dadzie, E. and Afful-Dadzie A., 2017), to identify opportunities and design market strategies for the private sector (Gottfried et al., 2021) and to analyze relationship between citizens and government (Bagozi et al.,

2021). Classification and clustering algorithms, regression models and feature selection were used to predict taxpayer groups (Cha, 2020), to formulate an environmental management strategy (Kang et al., 2021), to classify government expenditure records (De Oliviera, 2021) or to make analysis of open data judgments (Metsker, 2019). For this research, knowledge extraction would be made by using Natural Language Processing (NLP). Knowledge extraction or knowledge discovery is extraction of previously unknown, and potentially useful information from data (Frawley et al., 1992). NLP models can use, for example, part-of-speech tagging, chunking and parsing to describe syntactic information or use word-sense disambiguation, semantic role labeling or named entity extraction for gaining semantic information (Collobert et al., 2011). Statistical technique used behind NLP is Probabilistic Latent Semantic Analysis (PLSA), the purpose of which is to identify or distinguish contexts of words without a need to recourse to a dictionary or thesaurus (Hofmann, 2001). In other words, we can use NLP to reveal similarities of topics by grouping together words of a common context. The use of this method for identifying the impacts of OGD could not be detected in the literature, while the use of the method for various purposes is quite common. Among other things, NLP is used to develop smart linked open government data (Sinif and Bounabat, 2019), to develop methods for classifying government documents (Peña et al., 2018; Song et al., 2019), or to predict the emergence of civic activism and protests (Kallus, 2014).

As there is not much research with the chosen methodology for the selected purpose in the literature, the same is true for research on measuring the impact of OGD. Quantifying the economic impact of open data is relatively complex, as the most important benefits are indirect (Huyer and van Knippenberg, 2020). Although the analysis of the economic benefits of open data is a good way to describe the effect of government openness, the impact of open data is not connected to economic field only, as public sector openness brings other benefits to society by increasing government or institutional responsiveness (Keserű and James Kin-sing, 2015). As Keserű & James Kin-sing (2015) further point out, evidence of the social and political impact of OGD is extremely rare. Further, Carrara et al. (2015) found out that the majority of studies conducted on the impact of OGD are preliminary assessments, which are given on the basis of classical mechanisms for obtaining feedback - surveys. According to the guidelines of the Organisation for Economic Co-operation and Development (OECD), member states of the European Union (EU) are required to submit annual surveys to review the state of open data policies

(OECD, 2018). While surveys can provide important feedback, governments face constraints on staffing and funding in the regular collection, maintenance, and exchange of data as they meet other priorities in their work (Zuiderwijk & Janssen, 2014). On the other hand, Young & Verhulst (2016) report that case studies of individual initiatives can help to better understand the impact of open data, which we see as an opportunity to relieve public sector employees and use existing use cases to identify effects in an automated manner.

2   Problem definition

The problem we are solving in this research is therefore the identification of impact areas of OGD with the help of use cases, which would be compared with the existing impact areas, defined in the OECD surveys. The outcome of this research will be validated in cooperation with Ministry of Public Administration of the Republic of Slovenia who also provided us with their surveys.

Current impact areas, identified by OECD are policy, impact, portals and data quality. These are intertwined with the identified impacts from the literature, which are mainly grouped into three categories - operational and technical, social and political, and economic (Janssen et al., 2012; Zuiderwijk et al., 2018). Problem debated in this paper is also recognized in the literature, as many authors cite monitoring of the feedback and assessment of the actual impact of OGD for further research (Attard et al., 2015; Johnson, 2016; Lourenço, 2016; Ruijer & Martinius, 2017; Wilson & Cong, 2020, Zuiderwijk et al., 2018, Zuiderwijk & Janssen, 2014).

Research questions that we aim to address are the following:

- Can we use text mining methods on open data use cases to determine an objective assessment of OGD impact by automatic extraction of semantic structure?
- Can we use the proposed methodology to validate impact areas defined by OECD?

This section of the paper can best be summed up by Janssen et al. (2012, p. 260) "The main challenge is that open data has no value in itself; it only becomes valuable when used".

3   Methodology

The methodological approach for this research fits under the Design Science Research (DSR) approach, where an IT artefact, rooted in real-world problem is designed (Hevner et al., 2004). For understanding of the data, several authors (Azevedo & Santos, 2008; Bosnjak et al., 2009; Nadali et al., 2011; Schafer et al., 2018) recommend the use of CRISP-DM methodology.

CRISP-DM aims, as described by Wirth & Hipp (2000), to make larger data mining projects reliable, faster and cost efficient. This is achieved by following six phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment, nevertheless, the phases are usually intertwined and not linear (Bohanec et al., 2017).

Following CRISP-DM methodology, we will first start by analyzing OGD impact ecosystem, evaluation surveys and use cases from European Data portal (Data Europa EU, n. d.). By analyzing nearly 1000 use cases, we should be able to get a better understanding of the research area and the data. Since data understanding is the second phase of CRISP-DM we will then start by preprocessing the data. As preprocessing of the data is one of the first and most critical phases in data mining (Xiang-Wei & Yian-Fang, 2012), we suspect it will take the most effort. Next, modelling of the data will be made using text mining techniques, among which we will use NLP with its basic components such as word tokenization, stop words removal, part of speech tagging and stemming/lemmatization (Collobert et al., 2011). Finally, evaluation of impact areas with the help of Ministry of Public Administration will be made and a model for indexing, analyzing and searching heterogeneous document collections in order to extract knowledge from textual contents through NLP will be presented. A brief idea of the preliminary model is displayed in Figure 1 where keyword extraction from case studies is first made on the right side of the figure. Keywords from the documents are then grouped together and formed into collections or impact areas. Groups of variables are also created in order to better distinguish between groups of keywords and collections. Administration dashboard is the central entity, where inputs (input documents for impact assessment analysis) are uploaded and then classified based on similarity of specific impact area or impact area variables.
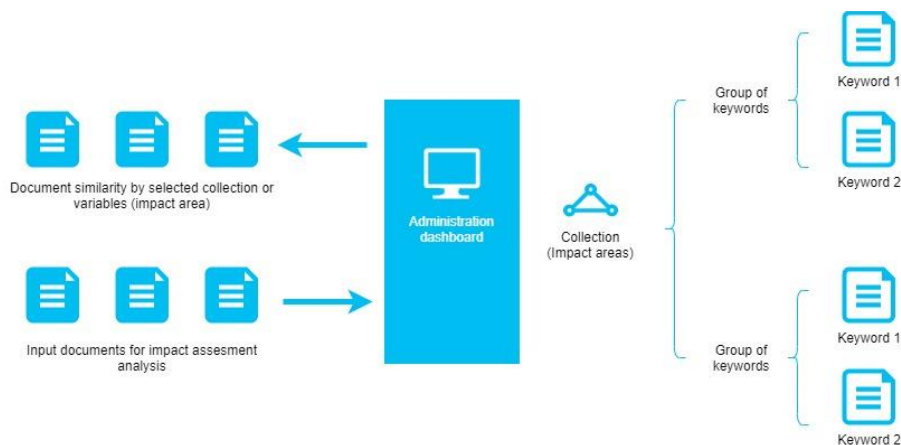
Figure 1: Preliminary schema of semantic structure for artefact development.

4   Expected results

Through our research, we aim to assess impact areas of OGD in an automated manner through use case analysis. We aim to cross-validate impact areas by comparing them with OECD surveys and further develop a model that will help governments with classification of impact areas based on provided inputs.

Based on research gaps we defined in this paper, we believe that assessing impact areas of OGD is important and will play an important role when benefits of open data initiatives are evaluated and summarized. Results of our research will contribute to better understanding of the actual impact of OGD and will help governments to prepare beneficial and focused datasets for providing even more value for citizens based on DSR's practical aspect of this research. By understanding the actual impact of OGD, more focused approach for opening the data and feedback mechanisms will be introduced from the governments. By designing a model for OGD impact assesment with the proposed methodology that hasn't been yet used for the addressed problem, we are filling up research gaps from the literature and that is essential for further academic research and for advancing knowledge in the field of open data.

References

Afful-Dadzie, E., & Afful-Dadzie, A. (2017). Liberation of public data: Exploring central themes in open government data and freedom of information research. International Journal of Information Management, 37(6), 664–672. doi:10.1016/j.ijinfomgt.2017.05.009

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. Government Information Quarterly, 32(4), 399–418. https://doi.org/10.1016/j.giq.2015.07.006

Attard, J., Orlandi, F., & Auer, S. (2016). Value Creation on Open Government Data. 2016 49th Hawaii International Conference on System Sciences (HICSS). doi:10.1109/hicss.2016.326

Azevedo, A., & Santos, M. (2008). KDD, SEMMA and CRISP-DM: a parallel overview. IADIS European Conf. Data Mining.

Bagozzi, B. E., Berliner, D., & Almquist, Z. W. (2019). When does open government shut? Predicting government responses to citizen information requests. Regulation & Governance. doi:10.1111/rego.12282

Bohanec, M., Robnik-Šikonja, M., & Kljajić Borštnar, M. (2017). Decision-making framework with double-loop learning through interpretable black-box machine learning models. Industrial Management & Data Systems, 117(7), in print (July, 2017b), http://dx.doi.org/10.1108/IMDS-09-2016-0409

Bosnjak, Z., Grljevic, O., & Bosnjak, S. (2009). CRISP-DM as a framework for discovering knowledge in small and medium sized enterprises' data. 2009 5th International Symposium on Applied Computational Intelligence and Informatics, 509-514. doi: 10.1109/SACI.2009.5136302

Carrara, W., Chan, W. S., Fischer, S., & van Steenbergen, E. (2015). Creating value through Open Data. European Commission. https://doi.org/10.2759/328101

Cha, T. (2020). Open Government Data for Machine Learning Tax Recommendation. The 21st Annual International Conference on Digital Government Research.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K. & Kuksa, P. (2011). Natural Language Processing (Almost) from Scratch. Journal of Machine Learning.

Crusoe, J., Simonofski, A., Clarinval, A., & Gebka, E. (2019). The Impact of Impediments on Open Government Data Use: Insights from Users. 2019 13th International Conference on Research Challenges in Information Science (RCIS). doi:10.1109/rcis.2019.8877055

De Oliviera Almeida, G., Revoredo, K., Cappelli, C. & Maciel, C. (2021). Method for Improvement of Transparency: Use of Text Mining Techniques for Reclassification of Governmental Expenditures Records in Brazil. International Journal of Business Intelligence and Data Mining, 1, 1. https://doi.org/10.1504/IJBIDM.2021.112989

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview. AI Magazine, 13(3), 57. https://doi.org/10.1609/aimag.v13i3.1011

Gottfried A, Hartmann C & Yates D. (2021). Mining Open Government Data for Business Intelligence Using Data Visualization: A Two-Industry Case Study. Journal of Theoretical and Applied Electronic Commerce Research, 16(4):1042-1065. https://doi.org/10.3390/jtaer16040059

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. MIS Quarterly, 28(1), 75-105. doi:10.2307/25148625

Hofmann, T. (2001). Unsupervised Learning by Probabilistic Latent Semantic Analysis. Machine Learning 42, 177–196. https://doi.org/10.1023/A:1007617005950

Huyer, E., & van Knippenberg, L., (2020) The Economic Impact of Open Data: Opportunities for value creation in Europe, Capgemini Invent. European, Data Portal. Doi:10.2830/63132.

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. Information Systems Management, 29(4), 258–268.

Jaeger, P. T., & Bertot, J. C. (2010). Transparency and technological change: Ensuring equal and sustained public access to government information. Government Information Quarterly, 27(4), 371–376. doi:10.1016/j.giq.2010.05.003

Johnson, P. A. (2016). Reflecting on the success of open data: How municipal government evaluates their open data programs. International Journal of E-Planning Research, 5(3), 1-12.

Kalampokis, E., Tambouris, E., & Tarabanis, K. (2013). Linked Open Government Data Analytics. Electronic Government, 99–110. doi:10.1007/978-3-642-40358-3_9

Kallus, N. (2014). Predicting crowd behavior with big public data. Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion. doi:10.1145/2567948.2579233

Kang, A., Ren, L., Hua, C., Song, H., Dong, M., Fang, Z. & Zhu, M. (2021). Environmental management strategy in response to COVID-19 in China: Based on text mining of government open information. Science of The Total Environment, Vol. 769. https://doi.org/10.1016/j.scitotenv.2021.145158.

Keserű, J. & James Kin-sing C. (2015). The Social Impact of Open Data." Sunlight Foundation. http://www.opendataresearch.org/dl/symposium2015/odrs2015-paper20.pdf

Lourenço, R. P. (2016). Evidence of an Open Government Data Portal Impact on the Public Sphere. International Journal of Electronic Government Research, 12(3), 21–36. doi:10.4018/ijegr.2016070102

Metsker O., Trofimov E., Sikorsky S. & Kovalchuk S. (2019) Text and Data Mining Techniques in Judgment Open Data Analysis for Administrative Practice Control. Communications in Computer and Information Science, Vol 947. https://doi.org/10.1007/978-3-030-13283-5_13

Nadali, A., Kakhky, E. N., & Nosratabadi, H. E. (2011). Evaluating the success level of data mining projects based on CRISP-DM methodology by a Fuzzy expert system. 2011 3rd International Conference on Electronics Computer Technology. doi:10.1109/icectech.2011.5942073

Open Government Data. (b. d.). https://www.oecd.org/gov/digital-government/open-government-data.htm

OECD (27. 9. 2018). Open Government Data Report. https://www.oecd.org/gov/open-government-data-report-9789264305847-en.htm

Peña, P., Aznar, R., Montañés, R. & Hoyo, R. (2018). Open Data for Public Administration: Exploitation and semantic organization of institutional web content. Sociedad Española para el Procesamiento del Lenguaje Natural, 61: 155-158. doi:10.26342/2018-61-21

Pereira, G. V., Macadar, M. A., Luciano, E. M., & Testa, M. G. (2016). Delivering public value through open government data initiatives in a Smart City context. Information Systems Frontiers, 19(2), 213–229. doi:10.1007/s10796-016-9673-7

Roa, H. N., Loza-Aguirre, E., & Flores, P. (2019). A Survey on the Problems Affecting the Development of Open Government Data Initiatives. 2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG). https://doi.org/10.1109/icedeg.2019.8734452

Ruijer, E. H. J. M., & Martinius, E. (2017). Researching the democratic impact of open government data: A systematic literature review. Information Polity, 22(4), 233–250. doi:10.3233/ip-170413

Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. Information Polity, 22(1), 1–24. https://doi.org/10.3233/ip-160012

Schafer, F., Zeiselmair, C., Becker, J., & Otten, H. (2018). Synthesizing CRISP-DM and Quality Management: A Data Mining Approach for Production Processes. 2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD). doi:10.1109/itmc.2018.8691266

Sinif, L., & Bounabat, B. (2019). A general framework of smart Open Linked Government Data. Proceedings of the 2019 2nd International Conference on Geoinformatics and Data Analysis - ICGDA 2019. doi:10.1145/3318236.3318243

Song, Y., Li, Z., He, J., Li, Z., Fang, X., & Chen, D. (2019). Employing Auto-Annotated Data for Government Document Classification. Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence - ICIAI 2019. doi:10.1145/3319921.3319970

Ubaldi, B. (2013), "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", OECD Working Papers on Public Governance, No. 22, OECD Publishing, Paris, https://doi.org/10.1787/5k46bj4f03s7-en.

Data Europa EU. (n. d.). Use Cases. Retrieved May 4, 2021, from https://data.europa.eu/en/impact-studies/use-cases

Wilson, B., & Cong, C. (2020). Beyond the Supply Side: Use and Impact of Municipal Open Data in the U.S. Telematics and Informatics, 101526. doi:10.1016/j.tele.2020.101526

Wirth, R. & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining, pg: 29 – 39.

Xiang-wei, L., & Yian-fang, Q. (2012). A Data Preprocessing Algorithm for Classification Model Based On Rough Sets. Physics Procedia, 25, 2025–2029. doi:10.1016/j.phpro.2012.03.345

Yan, A., & Weber, N. (2018). Mining Open Government Data Used in Scientific Research. Lecture Notes in Computer Science, 303–313. doi:10.1007/978-3-319-78105-1_34

Young, A. & Verhulst, S. (2016). The Global Impact of Open Data: Key Findings from Detailed Case Studies Around the World. O'Reilly Media.

Zuiderwijk, A. & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. Government Information Quarterly, 31(1), 17–29. doi:10.1016/j.giq.2013.04.003

Zuiderwijk, A., Shinde, R., & Janssen, M. (2018). Investigating the attainment of open government data objectives: Is there a mismatch between objectives and results?. International Review of Administrative Sciences, 85(4), 645-672.

Development of Prediction Model for Support in Decision-Making Process in Football Academies – Literature Review

Rok Vrban

Unvieristy of Maribor, Faculty of Organizational Sciences, 4000 Kranj, Slovenia, e-mail: rok.vrban@student.um.si

Abstract Talent development process in football is considered as a process of providing the optimal environment for identifying and realising the maximum ability of young athletes. A multidimensional approach to analysing factors that influence junior to senior transition can produce much better support for coaches and management to distinguish elite and non-elite players. With development of digital technologies and artificial intelligence, more clubs are able to perform detailed analysis of their youth development programme. In this paper, we focus on identfying good practices in connecting digital technologies with talent development process in sports. Based on established methods and techniques used by experts in a field of data mining within sports, we want to select an appropriate methodology and approach in discovering knowledge from the data for the doctoral dissertation. Literature review presents a first step in a hollistic process of identifying key attributes in junior to senior transition. The findings suggest that the comprehensive approach towards analysing data in sports, results in better identification of skills and attributes of young athletes. Consequently, data mining in sports is becoming more and more important in assessing important characteristics on every level within talent development process.

1   Introduction

European football market size has doubled over the past decades. Many private investors from Asia, Middle East and Russia have joined the European football landscape in that period. This inflow of funding raised transfer activity and many clubs started searching for better options on market rather than in their own youth academy. Financial capabilities allow big clubs to invest money in the more developed players, while clubs with limited financial abilities need to go through the entire junior-to-senior transition (JST) process in order to transition them to the first squad or sell them. As most of the clubs depend on selling players, they must adapt to current trends and development of new technologies. Among these are comprehensive statistical analysis, which, combined with use of sophisticated technology and artificial intelligence, create a new understanding of the development process and evaluation of the players. Measuring performance of individuals during training sessions and league games has increased in recent years together with technological improvements. Artificial intelligence and modern statistical approaches in discovering knowledge in data (data mining) have brought new ways of understanding key factors in measuring performance, skill, attributes, talent development and other important aspects, which play an important role in the world of football. Most of the small and medium sized European football clubs evaluate their youth prospects solely qualitatively or by using basic statistical methods, which turn out to be insufficient in accurate monitoring of talents' professional development. Moreover, the collected data has small impact on decision-making process as coaches and managers have limited or no professional support from data analysts. Consequently, the data present a burden rather than the advantage for the club.

2   Methodology

The central question of the research project is: Which athlete's attributes play a key role in successful transition from youth to professional football player?

In order to answer the central question, the purpose of the study is to create a model, which will help us recognize importance of factors in youth development in football. By creating a model, which will be based on quantitative research of sub-elite and elite youth academies across various European football clubs, we want to encourage

R. Vrban:
Development of Prediction Model for Support in Decision-Making Process in Football Academies – Literature Review

791

sport clubs around Europe to start using advanced data methods and modern technological equipment to improve their youth academies and to understand the value behind the work they put in raising youth to become professionals. First step is literature review where we will overview the topic, compare different views of authors about the existing problem and examine the gap in the research. Articles in the literature review will be obtained through various online databases, such as Web of Science (WoS) and other bibliographic databases, mainly connected to sports, talent development and digital technology. The key words used for literature review included "talent development in sports", "talent development and artificial intelligence" and "data mining in sports". The data for the research will mostly be collected through interaction with high profile subjects within management of the clubs and the players of the observed club and through different research methods. We plan on using a mixed method of design science research (DSR) and multiple case study (MCS). The main methodological approach will be DSR, while we intend on using MCS for collecting and analysing the data for model development. Quantitative data will be collected through observation and measurement tests of pre-defined mental, physical, tactical and technical skills. The data will also include other attributes and characteristics of each individual. Qualitative data will be collected through interaction with managers and coaches within youth academy and by interviewing observed players (semi-structured interviews). While some data are publicly available, the rest is to be collected in agreement with the selected clubs and individuals who are willing to provide their personal opinions on requested topics. The proposed duration of collecting the data is a year and a half. Each of the selected clubs will be evaluated once every 6 months in a span of a week. Clubs will be selected according to groups and sub-groups classified by the ECA evaluation system. Proposed sample involves up to 3 football clubs in each division (i.e. max. 12 clubs). Each club will be requested to assign their youth teams (between 13 – 17 years old) for the tests. Tests will be performed by coaches and training specialists according to instructions given by the researchers. Altogether the sample should involve between 600 - 1000 talents. The data will be analysed with modern statistical methods and tools. Model will be based on data mining process and evaluated through CRISP-DM process model. Anticipated methods include algorithms such as KNN, AdaBoost, Naïve Bayes and Random forest.

3   Existing literature

As football is the most popular and played sport in the world with estimated 3,5 Billion fans all around the world, much research has been done in the academic and non-academic field. With more and more technology involved in sports, many other professions, traditionally less connected to football, find their common points of research with football. A lot of research has been done connecting football and medicine (such as Hawkings et al., 2001, Krustrup & Krustrup, 2018, Kramer, 2010) or football and tracking systems (e.g. Link et al., 2016, Schütz et al., 2019). With development of football related topics, clubs learn and implement new business models to their youth academies. On the other hand, it is important to understand the perspectives of youth players themselves. Chamorro et al. (2016) suggest, that youth football players, who are not only interested in becoming professionals but also give importance to education and private life, appear to be more resourceful to cope with the transition to professional football. According to Franck (2018, p.55) JST is very complex and therefore, she suggests the following: "I recommend that future research should continue combining quantitative methods with in-depth interviews in a mixed-methods approach when exploring the JST. This type of approach provides detailed pictures of different pathways and life experiences. In the future, research exploring how social agents in the environment perceive their roles in assisting JST athletes could also be beneficial.« Beside athletic predispositions, it is also crucial for youth to compete on high level before moving to senior level. Hollings (2013) claims, that there is a higher probability for youth athletes, who have already competed on high level, to become elite senior players. Lorenzo et al. (2009) came to the same conclusion in basketball. Interviewed youth players pointed out tougher demand and seriousness when they train with the senior team. They required more mental and physical effort. In order to train more, they had to sacrifice their leisure time and consequently they wasted more time for studying. The transition from amateur to professional can differ from sport to sport. One of the most prominent researchers in the youth development process in sports, Stambulova (2010), presents a 5-step career planning strategy as a guideline for each individual to estimate different spheres in his/her current life and structure the future by thinking forward and adjusting priorities to the wished result. Bennie & O'Connor (2006) claim, that transition from youth to senior stage is also influenced by other forces, such as psychological, social, economic and political forces. This claim is consistent with the prediction, that development of youth players is not only

R. Vrban:
Development of Prediction Model for Support in Decision-Making Process in Football Academies –
Literature Review

793

connected to one's motivation and existing sportive components, but also challenges and problems of a daily life. Another important factor in the youth development process are talent's parents. They are important for young athletes, but can often be too pushy (Stambulova et al., 2009). To check intensiveness of pushing young talents by their parents, coaches or other authorities, authors many times rely on the push pull anti-push anti-pull model. Ekstrőm & Sundqvist (2009) use the model to investigate the reasons of continuing or dropping out of the sport at Swedish youth hockey players in transition. In order to understand why and how much an athlete improves, we must create a continuous process of measuring and evaluating individual's physical, tactical and technical attributes. Growing sport industry has become more demanding and consequently, professional sport teams hardly compete on an international level without a full support of advanced technology. Yokesh & Kumar (2015) state, that technology changes the way we see sports. It influences what we play and how we play it. It improves ways to predict and treat injuries. De Koning (2010) agrees with that by claiming, that technological improvements could be the key to future world records. Author based his claims on comparing world record times in 1500-m speed skating for men and development of technology. Cortsen & Rascher (2018) focus on technology and data within football. They claim that data inside sport have much higher value than many would think and are much more difficult to use effectively. The problem is not just technical but lack of financial and human resources as well. They also argue the ownership of the data, which is crucial for profitability. Yet, profitability is only possible if the quality of the data is sufficient and if the owner knows how to commercialize it properly. Rein & Memmert (2016) agree with these claims and add that big amount of data may pose a problem if the club does not have a sufficient organizational structure and knowledge behind using modern technology. Machine learning process in football can be valuable, but it must be in collaboration with sport scientists as the process of understanding big data within tactical and technical progress is extremely complex. With implementation of machine learning process Musa et. al. (2020) present multiple performance-related parameters in beach football and sepak takraw. Key performance indicators that define performance were found in beach football, while results from analysis of sepak takraw turned out inconclusive. However, it was demonstrated in the findings, that a number of performance indicators are essential in distinguishing between losing and winning. It was also concluded, that anthropometric indexes (such as standing and sitting

height, leg length, skinfold measurements and other) have an impact in performance of sepak takraw. Deep statistical analysis of the team and individuals performing in team sports have also shown correlation between behavioural characteristics and overall result of the team. Excellent players, who are hated in the dressing room, who are prone to injuries or homesick, are less likely to bring out the best in their colleagues (Tunaru & Viney, 2010). Another important characteristic of a successful talent development is motivation. Forsman et. al. (2016) identify motivation as one of the key reasons of successful talent development among football players in Finland. Motivation does not only help a player maximize his abilities, furthermore, it helps the team boost their confidence and rises morality. Authors also recognized agility, passing and centering skills as most important ones in becoming elite player. Pappalardo et. al. (2019) show that elite players do not always play excellence, they just achieve it more often than the other players. Murr et al. (2018) add, that elite players perform significantly better in dribbling tests and ball control, while shooting ability had little impact on future development success. Authors have also found significant positive correlation between motivation and future performance, which confirms Forsman et. al. assumptions. Players with higher self-determination were more likely to get selected to a higher performance level compared to the players with lower self-determination. The results are supported by Kelly et. al. (2020) analysis, which identified higher potentials had significantly better technical skills such as lob pass ability, possession reliability, pass completion and average total touches. Another important factor in talent identification at early age is relative age effect (RAE). Doncaster et. al. (2020) demonstrate a persistent bias toward selecting individuals born earlier in year in male football and basketball.

4   Conclusion

Based on this paper, the findings regarding talent development process and data analysis show significant correlation. This study systematically analysed literature review within connections between sports, talent development and use of data. It is suggested that an appropriate use of data can determine and identify future abilities and skills of young athletes. Moreover, it is required for professional sport teams to use modern machine learning processes in order to compete on a top level. However, it is important to understand the value of the ownership of the data. Furthermore, data need to be addressed professionally, as too much data might bring more confusion rather than positive effect if treated wrongly. Conclusions regarding

the importance of social factors show that parents, coaches and management of the club have an important role in talent development. Another important feature in JST is personality. Studies have shown that individuals with professional approach and teamwork mindset are more likely to influence positive on the other individuals within the squad. They are also more likely to advance faster in development process as they spend more leisure time for individual practice. Future research is required to explore other attributes and factors which play an important role in elite and sub-elite player assessment. Additionally, data driven decision-making in football is still relatively young process, which offers many opportunities to research and explore.

References

Bennie A., O'Connor D. (2006). Athletic transition: An investigation of elite track and field participation in the post-high school years. Change: Transformations in Education. Volume 9.1, November 2006.
https://www.researchgate.net/publication/277786513_Athletic_transition_An_investigation _of_elite_track_and_field_participation_in_the_post-high_school_years. Accessed: 05/04/2020

Chamorro J. L., Torregrosa M., Oliva Sánchez D., Calvo Garciá T., León B. (2016). Future Achievements, Passion and Motivation in the Transition from Junior-to-Senior Sport in Spanish Young Elite Soccer Players. Cambridge University Press: 20 October 2016. https://www.ncbi.nlm.nih.gov/pubmed/27762186 Accessed: 10/04/2020.

De Koning J. J. (2010). World Records: How Much Athlete? How Much Technology? International Journal of Sports Physiology and Performance, 2010, 5, 262-267.
https://www.researchgate.net/publication/45167594_World_Records_How_Much_Athlete _How_Much_Technology Accessed: 07/01/2021

Doncaster G., Medina D., Drobnic F., Gomez-Diaz, A. J., Unnithan, V. (2020). Appreciating Factors Beyond the Physical in Talent Identification and Development: Insights From the FC Barcelona Sporting Model. Front. Sports Act. Living 2:91.

Ekström, C., & Sundqvist, C. (2009). Karriärövergången från nybörjarstadiet till utvecklingsstadiet inom ishockey utifrån spelare, tränare och föräldrars perspektiv. (C-uppsats i psykologi inriktning idrott, 61-90p) Högskolan i Halmstad: Sektionen för Hälsa och Samhälle. https://www.diva-portal.org/smash/get/diva2:292741/FULLTEXT01.pdf. Accessed: 26/03/2020

Forsman H., Blomqvist M., Davids K., Liukkonen J., Konttinen N. (2016). Identifying technical, physiological, tactical and psychological characteristics that contribute to career progression in soccer. International Journal of Sports Science & Coaching 2016, Vol. 11(4) 505–513

Franck A. (2018). THE JUNIOR-TO-SENIOR TRANSITION IN SWEDISH ATHLETES. A longitudinal study. Linnaeus University Dissertations Nr 305/2018. https://www.diva-portal.org/smash/get/diva2:1168706/FULLTEXT01.pdf Accessed: 10/04/2020

Hawkings R.D., Hulse M. A., Wilkinson C., Hodson A., Gibson M. (2001). The association football medical research programme: an audit of injuries in professional football. https://bjsm.bmj.com/content/35/1/43.short Accessed: 27/03/2020

Hollings, S.C. (2013). The Transition from Elite Junior Track-and-Field Athlete to Successful Senior Athlete: Implications for Athletics High Performance Programmes. https://core.ac.uk/download/pdf/56364504.pdf Accessed: 06/04/2020

Kelly, A., Wilson M. R., Jackson, D.T., Williams, C. A. (2020). Technical testing and match analysis statistics as part of the talent development process in an English football academy, International Journal of Performance Analysis in Sport, 20:6, 1035-1051.

Kramer E.B. (2010). Football emergency medicine. Emergency medicine in football is now regarded as a sub-speciality of emergency medicine. CME: Your SA Journal of CPD (Vol. 28, Issue 5).

Krustrup P., Krustrup B.R. (2010). Football is medicine: it is time for patients to play! British Journal of Sports Medicine 2018;52:1412-1414. https://bjsm.bmj.com/content/52/22/1412. Accessed: 30/03/2020

Link D., Lang S., Seidenschwarz P. (2016). Real Time Quantification of Dangerousity in Football Using Spatiotemporal Tracking Data. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5201291/. Accessed: 25/03/2020

Lorenzo, A., Borrás, P. J., Sánchez, J. M., Jiménez, S., Sampedro, J. (2009). CAREER TRANSITION FROM JUNIOR TO SENIOR IN BASKETBALL PLAYERSRevista de Psicología del Deporte, vol. 18, núm. 3, 2009, pp. 309-312. https://www.researchgate.net/publication/237030588_Career_transition_from_junior_to_senior_in_basketball_players. Accessed: 25/03/2020

Murr, D., Feichtinger, P., Larkin P., O'Connor D., Höner, O. (2018). Psychological talent predictors in youth soccer: A systematic review of the prognostic relevance of psychomotor, perceptual-cognitive and personality-related factors. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0205337. Accessed: 07/05/2021

Musa, R. M., Majeed A. P. P. A., Kosni N. A., Abdullah M. R. (2020). Machine Learning in Team Sports Performance Analysis and Talent Identification in Beach Soccer & Sepak-takraw. SPRINGER BRIEFS IN APPLIED SCIENCES AND TECHNOLOGY.

Pappalardo, L., Cintia P., Ferragina, P., Massucco, E., Pedreschi, D., Giannotti, F. (2019). PlayeRank: data-driven performance evaluation and player ranking in soccer via a machine learning approach. https://arxiv.org/pdf/1802.04987.pdf. Accessed: 07/05/2021

Rein R., Memmert, D. (2016). Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. SpringerPlus (2016) 5:1410

Schütz N.S., Pedersen B.B., Rerup B. (2019). VALUE CREATION THROUGH DATA TRACKING TECHNOLOGIES IN THE FOOTBALL INDUSTRY OBSTACLES DYSFUNCTIONAL EFFECTS. MCIS 2019 Proceedings. 35. https://pure.au.dk/portal/files/173292768/VALUE_CREATION_THROUGH_DATA_TRACKING_TECHNOLOGIES_IN_THE_FOOTBALL.pdf. Accessed: 06/04/2020

Stambulova N. (2010). Counseling Athletes in Career Transitions: The Five-Step Career Planning Strategy. https://www.researchgate.net/publication/232844112_Counseling_Athletes_in_Career_Transitions_The_Five-Step_Career_Planning_Strategy Accessed: 06/04/2020

Stambulova N., Alfermann D., Statler T., Côté J. (2009). ISSP Position Stand: Career Development and Transitions of Athletes. IJSEP, 2009, 7, 395-412. https://www.issponline.org/images/isspdata/position_stands/Career_Development_and_Transitions_of_Athletes.pdf Accessed: 07/04/2020

Tunaru R. S., Viney H. P. (2010). Valuations of Soccer Players from Statistical Performance Data. Journal of Quantitative Analysis in Sports. Volume 6, Issue 2, Article 10.

Yokesh T.P., Kumar S. (2015). VARIOUS SPORTS TECHNOLOGIES ARE USED IN DEVELOPING THE SPORTS PERFORMANCE. ISSN 2231-3265 International Journal of Health, Physical Education and Computer Science in Sports Volume No.29, No.2 pp 5-7 https://www.researchgate.net/publication/339642888_Various_Sports_Technologies_are_used_in_Developing_the_Sports_Performance. Accessed: 07/01/2021

34ᴛʜ Bled eConference Digital Support from Crisis to Progressive Change

Andreja Pucihar (et al.)

Program Committee Chair, 34th Bled eConference, Professor, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia, e-mail: andreja.pucihar@um.si

Abstract Bled eConference, organized by University of Maribor, Faculty of Organizational Sciences, has been shaping electronic interaction since 1988. In 2021, Bled eConference was held online for the 2nd time due to the Covid-19 pandemic. The role of digital technologies has never been more important than during the pandemic, allowing people to interact, collaborate and find new opportunities and ways to overcome challenges. The theme of the 34th conference is "Digital Support from Crisis to Progressive Change". Society is beginning to envision what the post-crisis world will look like and is calling for different economic models for the well-being and sustainable development of society. Digital technologies play an important role in achieving these goals. We address the opportunities and challenges of digital transformation and provide guidance for organizations. Topics of the conference proceedings include Digital Transformation, Digital Business, Business Models, Data Science, Digital Health, Digital Wellness, Digital Ethics, Digital Education, Sustainable Cities and Digital Consumers.

Keywords:
digital support,
crisis,
progressive change

University of Maribor

Faculty of Organizational Sciences

BLED
eCONFERENCE
2021

UNI BIZ TECH GOV