Isabelle Fries, Michael Grabatin, Manfred Hofmeier (eds.)

Sovereign by Design

The LIONS Approach to Digital Sovereignty

λογος

Sovereign by Design
The LIONS Approach to Digital Sovereignty

Sovereign by Design

The LIONS Approach to Digital Sovereignty

Isabelle Fries, Michael Grabatin, Manfred Hofmeier (eds.)

Logos Verlag Berlin

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de .

ISBN 978-3-8325-5834-5

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10,
12681 Berlin, Germany

Tel.: +49 (0)30 / 42 85 10 90
Fax: +49 (0)30 / 42 85 10 92
https://www.logos-verlag.com

Table of Contents

Digital Sovereignty as a Field of Learning

LIONS Research Project

Ledger Innovation and Operation Network for Sovereignty

The transdisciplinary LIONS research project established a research platform for exploring Distributed Ledger Technology (DLT) as a digitalization technology to increase resilience and digital sovereignty.

LIONS developed technical and analytical competencies, provides a laboratory environment with infrastructure for DLT of realistic size, and is building a community comprising the Bundeswehr, government agencies, and the private sector. Indicators and tools for the analysis, design, and implementation of DLT-based information systems and their contribution to resilience and digital sovereignty are developed, taking into account three perspectives of analysis: (1) individual, (2) supply chain, and (3) society.

The project is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

For more information about the research project, visit the LIONS website at https://www.unibw.de/lions.

dtec.bw
Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr

Funded by
the European Union
NextGenerationEU

Foreword

Dear Reader,

Take a minute to think about your future, how digital technology will shape it, and how much it may change in a handful of years. Resilience and sovereignty are two future ideas that are new on the political agenda. What needs to be done to achieve resilience and digital sovereignty? What needs to change in societies' digitalization strategy and the way they design, build, and operate their digital infrastructure? This book and the ideas it presents hopefully inspire readers to think about important topics and novel design ideas.

This book is a result of a unique four-year journey, the LIONS research project. Led by a consortium of researchers from the University of the Bundeswehr Munich and the Helmut-Schmidt-University in Hamburg, the project took a novel and interdisciplinary approach in collaboration with industry partners and the Bundeswehr. The researchers explored new ideas and design concepts from the diverse perspectives of ethics, psychology, media pedagogy, information systems, and computer science. The Digitalization and Technology Research Center of the Bundeswehr, dtec.bw, funded by the European Union in the NextGenerationEU initiative, provided a framework for a unique, transdisciplinary setting for exploring pioneering ideas.

The LIONS research project addressed two starting points: approaches to designing the digital future after the Covid pandemic, with its impact on the supply chain; and the new, more digital normality. Digitalization, decarbonization, and geopolitical shifts shape public discourse about the future of civil and democratic society. It seems that interdependent crisis situations dominate the debate and that society is weary of the game of hare and hedgehog. As digitalization is both one of the crisis themes and a way to mitigate risks, LIONS has embarked on the research with a focus on technology design.

LIONS emphasizes the individual in shaping resilience and sovereignty: individuals make decisions about the future, and therefore, awareness of risks, strategic options, and efficacy of measures are key to empowering individuals to become active in shaping the digital future through investments in novel technology. Organizations and the state set the frame for implementing these decisions. The research of LIONS addresses the resilience of the individual as a virtue; awareness about digital infrastructure and its risks; design options for future digital infrastructures; and identity management to empower individuals to shape their individual role in the digital infrastructures. Distributed Ledger Technology is employed to demonstrate future designs of novel infrastructure with information flows across the supply chain, to

enable the resilience of the supply chain and the individual sovereignty of all actors. Serious games, simulations, and demonstrators are tangible research results alongside ethical considerations, guidelines, and models from the consultancy perspective.

It was a pleasure to lead this initiative, and I wish the reader inspiration and energy to strive for a more digital, resilient, and sovereign future!

Prof. Dr. Ulrike Lechner

Preface

Dear Reader,

Many people want it, most find it difficult to define, and the LIONS project has been researching it from 2021 to 2024: the answer is digital sovereignty, which we address in this book. Researchers from computer science, information systems, psychology, pedagogy, and ethics have contributed their respective expertise to demonstrate transdisciplinary approaches to a socially desired goal that is more than just a buzzword in the digital age.

Digital sovereignty is demanded for individuals, society, organizations, and businesses, not just politically on national and EU levels. Questions of cybersecurity, especially concerning the protection of critical infrastructure, are directly related. Decentralization and interoperability on the one hand and legal frameworks such as data protection on the other are just some of the relevant reference points. If digital sovereignty means technological independence, the focus is on reducing dependencies on foreign providers. When the digital sovereignty of individual citizens is in focus, questions of empowerment (pedagogy), individual resilience (psychology), and responsibility (ethics) follow.

In a unique way, the LIONS research group has mapped the field of digital sovereignty. With "Sovereign by Design – The LIONS Approach," we present a collective culmination of intensive project work. All researchers on the LIONS project were invited to present their perspectives on digital sovereignty, viewing digital sovereignty from a macro perspective or dealing with a specific aspect or a concrete technical implementation to enhance digital sovereignty.

The result is a book that is divided into three parts. The first part, "The Challenge of Digital Sovereignty," approaches the topic of digital sovereignty. It begins by introducing the multi-layered model of sovereignty developed by the LIONS project. Further exploration includes defining the relationship between sovereignty and resilience, identifying success factors for digital sovereignty, and drawing up ethical guidelines for the use of DLT, culminating in an advocacy for ethical sovereignty in the digital sphere.

The second part, "Designing Sovereign Information Systems," builds on the ethical perspective but locates it very concretely in practical application within the supply chain, which remains relevant in the further contributions of this section. Those are equally application-oriented and partly based on design studies, focusing on the governance of a traceability system, semi-autonomous workforces, electronic signatures using blockchain technology and self-sovereign identity, or a reputation evaluation system for more resilient IT supply chains.

In the third part, “Digital Sovereignty as a Learning Field,” contributions from pedagogical and psychological perspectives focusing on the digital competence of individuals are included. Important parts of the LIONS project, from serious games to escape room challenges, are also featured among the contributions in this section.

Our goal is to present a nuanced examination of digital sovereignty, to acknowledge the achievements of our project work, to make a significant contribution to the ongoing scholarly discourse, and to provide impetus for public discussion. By weaving together diverse insights and rigorous research, this book offers perspectives for scholars, practitioners, and policymakers.

We would like to take this opportunity to thank dtec.bw – Digitalization and Technology Research Center of the Bundeswehr and the European Union – NextGenerationEU for funding the LIONS project and the publication of this book. We also thank the other project members and partners for their contributions to the LIONS project: Prof. Dr. Florian Alt, Sandra Bayer, John Bechara, Yasemin Dolu, Judith Fingerle, Prof. Dr. Andreas Fink, Dr. Tiago Gasiba, Maximilian Greiner, Lilian Grimmeisen, Benjamin Gröschel, Isabelle Haunschild, Rene Hennen, Michael Hofmeier, Tim Hoiß, Dr. Jens Holtmannspötter, Prof. Dr. Wolfgang Hommel, Razvan Hrestic, Andrei-Cristian Iosif, Martha Klare, Lisa Kolb, Matias Meno, Prof. Dr. Ulrike Lechner, Prof. Dr. Bernhard Leipold, Prof. Dr. Friedrich Lohmann, Prof. Dr. Manuela Pietraß, Prof. Maria Pinto-Albuquerque, Markus Rebhan, Prof. Dr. Burkhard Schäffer, Thomas Schuhrke, Karl Seidenfad, Ersin Söbütay, Thomas Stöger, Kai Weeber, Prof. Dr. Thomas Wendeborn, Tiange Zhao.

We hope that this volume not only provides insights into the complex topic of digital sovereignty but also inspires future research and discussions. May this work contribute to deepening and enriching our understanding and practices in the digital age.

Dr. Isabelle Fries, Dr. Michael Grabatin, and Dr. Manfred Hofmeier

The Challenge of Digital Sovereignty

Layered Model of Digital Sovereignty

Isabelle Fries[1], Maximilian Greiner[2], Manfred Hofmeier[3], Razvan Hrestic[4], Ulrike Lechner[5], and Thomas Wendeborn[6]

Abstract: This paper addresses the concept of "digital sovereignty" from various academic disciplines in a holistic approach: in the discussion of digital sovereignty, questions arise, whose digital sovereignty is being addressed, what digital sovereignty means for the respective entities, how to increase digital sovereignty, and how to build a digital sovereign civil society and its critical infrastructures. We present a layered model that conceptualizes the meaning of digital sovereignty on three layers: (1) state or supranational institution, (2) organization, (3) individual, as well as the relationships between the three layers. This model provides guidance for research and practice – including policy and decision-making. An earlier version of this model was published at the CRITIS 2022 conference (Fries et al., 2023).

Keywords: Digital Sovereignty, Model, Interdisciplinary

1 Introduction

While digital sovereignty has become an increasingly important topic in the political field and in public discourse, it is now also gaining more interest in scientific research. When discussing digital sovereignty, in most cases the focus is on a state or supranational entity level (Seidel & Bös, 2009). But "many other meanings are emphasized when talking about sovereignty" (Couture, 2020). Sovereign entities could be states or supranational institutions, organizations such as companies, or individuals. Security efforts are generally related to "the economic lifeblood, social stability, and public interests of countries, and even national security, and involves common concerns of all countries" as Fang rightly points out (Fang, 2018, p. 196). In German politics this is made clear by the fact that the 2021 coalition agreement enshrines digital sovereignty at both the state and individual level (Scholz, 2021, p. 12 f.). In this context, the term "digital sovereignty" seems to be used frequently, yet lacks a clear definition (Steinbach, 2019, p. 82).

Disentangling the notion of digital sovereignty from informal and political discourse is a challenging task. There are other terms describing the topic, often in the context of cybersecurity: while in the discourse at the level of the European Union (EU) there is talk of "cyber sovereignty" and this seems to refer to a nationally restrictive idea with a view to the policies of China or Russia, there is talk both nationally and supranationally

1 University of the Bundeswehr Munich, Neubiberg, isabelle.fries@unibw.de
2 University of the Bundeswehr Munich, Neubiberg, maximilian.greiner@unibw.de
3 University of the Bundeswehr Munich, Neubiberg, manfred.hofmeier@unibw.de
4 University of the Bundeswehr Munich, Neubiberg, razvan.hrestic@unibw.de
5 University of the Bundeswehr Munich, Neubiberg, ulrike.lechner@unibw.de
6 Leipzig University, Leipzig, thomas.wendeborn@uni-leipzig.de

of "data sovereignty", which more strongly emphasizes data handling at the level of the individual and the organization (Chin & Li, 2021). The concept of data sovereignty is also found in particular in the context of the GAIA-X initiative (AISBL & Cloud, 2021), an European attempt to create an alternative to provider-specific commercial cloud platforms. Pertaining to this, the term "strategic autonomy" is used in the EU context (European Political Strategy Centre of the European Commission, 2019).

However, one crucial point seems to uniformly presuppose today's discussion of digital sovereignty: digital sovereignty is perceived as a desirable good, which at the same time seems to exist no longer, not yet, or not to a satisfactory degree. In the context of political efforts in Germany, for example, the 2013 coalition agreement already states that the "technological sovereignty" referred to here is to be "regained" (CDU, CSU, & SPD, 2013, p. 103). At the same time, the pursuit of digital sovereignty is linked to the pursuit of independence and competitiveness. There is discussion about the dependence of states and organizations on external hardware and software, which must be overcome, or of the digital literacy of citizens, which has not yet been achieved to a satisfactory degree.

When looking for a concept-historical approach to understanding, one quickly comes across state theory. Reference is often made to Thomas Hobbes's "Leviathan" and to a social contract between an absolute sovereign and the people. Although this is initially far from an intuitive understanding of what seems to be meant by digital sovereignty today, it does make one thing very clear: the term "sovereignty" is accompanied by a relational determination that remains intact even in the digital space. An entity is always sovereign in relation to another entity or thing. It is therefore useful to look at various relational constellations in the following in order to make the talk of digital sovereignty tangible. Ultimately, the same applies to the still rather new discussion of digital sovereignty as it does to language in general: language changes, is a fluid mass rather than a static construct, and gains its shape in social interaction.

In the interaction of social actors and in the interdisciplinary discourse, open questions regarding digital sovereignty also emerge. Some of them already appear in research. If one defines digital sovereignty as "the ability to act and make decisions in the digital space" (Beyerer, Müller-Quade, & Reussner, 2018, p. 6), it is a construct contextualized from the state's power of disposal. But is the state digitally sovereign? When is an individual able to act as digitally sovereign? Who or which institution has digital sovereignty? All these and other questions cannot be answered unequivocally (Couture, 2020). On the other hand, there are different aspects to the effective implementation of digital sovereignty. According to Beyerer (Beyerer et al., 2018), these are the sovereignty of infrastructures, sovereignty of data, sovereignty of decision-making, and sovereignty platforms.

These aspects accompany our approach to understanding digital sovereignty presented below. To conceptualize the term "digital sovereignty" for researchers and practice (such as policy and decision making), we developed a layered model using creative techniques. We start from the categories of state, organization, and individual, and

present both, them and their respective relational structures. We also include the above-mentioned related idioms, such as of cyber sovereignty or data sovereignty, to sharpen understanding.

Our approach is holistic in that we assume that the various layers of digital sovereignty cannot be viewed in isolation from one another. We assume that the political aspiration to increase digital sovereignty can be met if the latter is thought of as a systemic overall process that encompasses state, organization, and the individual in a reciprocal, interactive, and interrelational context. To approach such an understanding of digital sovereignty, we deliberately work in an interdisciplinary way. Given the breadth of the concept and the different views of the disciplines, we do not believe that an all-encompassing definition of digital sovereignty can (or should) be given. Thus, we deliberately avoid a normative framework for the term. Our goal with the resulting model is to provide a flexible frame by using few, but widely defined categories that are interconnected by factors either reducing or supporting digital sovereignty. This structure can be scaled by adding depth, e.g., when considering the individual layer; one could state that it composes the aspects of culture and formal education, which themselves could be interconnected in the above sense. This creates a self-similar but consistent basis to build upon when further structuring future knowledge regarding digital sovereignty is needed.

An earlier version of the model as well as the research design were published at the CRITIS 2022 conference (Fries et al., 2023).

2 Layered Model of Digital Sovereignty

As mentioned in the research design (Fries et al., 2023), we found suggestions in the domain literature that a layered approach to digital sovereignty would depict its wide-reaching implications better than a one-dimensional approach. Wittpahl (Institut für Innovation und Technik, 2018), for instance, structures his book in three sections: citizen, company, and state. Pohle (2020) suggests that there are individual, state, and economic dimensions, the intersection of which falls into the notion of digital sovereignty. The organization BITKOM, representing the IT industry in Germany, also takes a layered approach to describing digital sovereignty, but positions it as a stakeholder relationship between consumers, enterprise users, and the state (BITKOM, 2018).

Our proposed model introduces a novel element in placing the organizational layer on the same level as those of the individual and the state, both of which represent the main focus of the current literature on the topic of digital sovereignty. We also expand the state dimension to include supranational entities such as the EU, since the issue of sovereignty always entails the question of an international determination of relations and the EU in particular can have a great influence on the policies of individual member states.

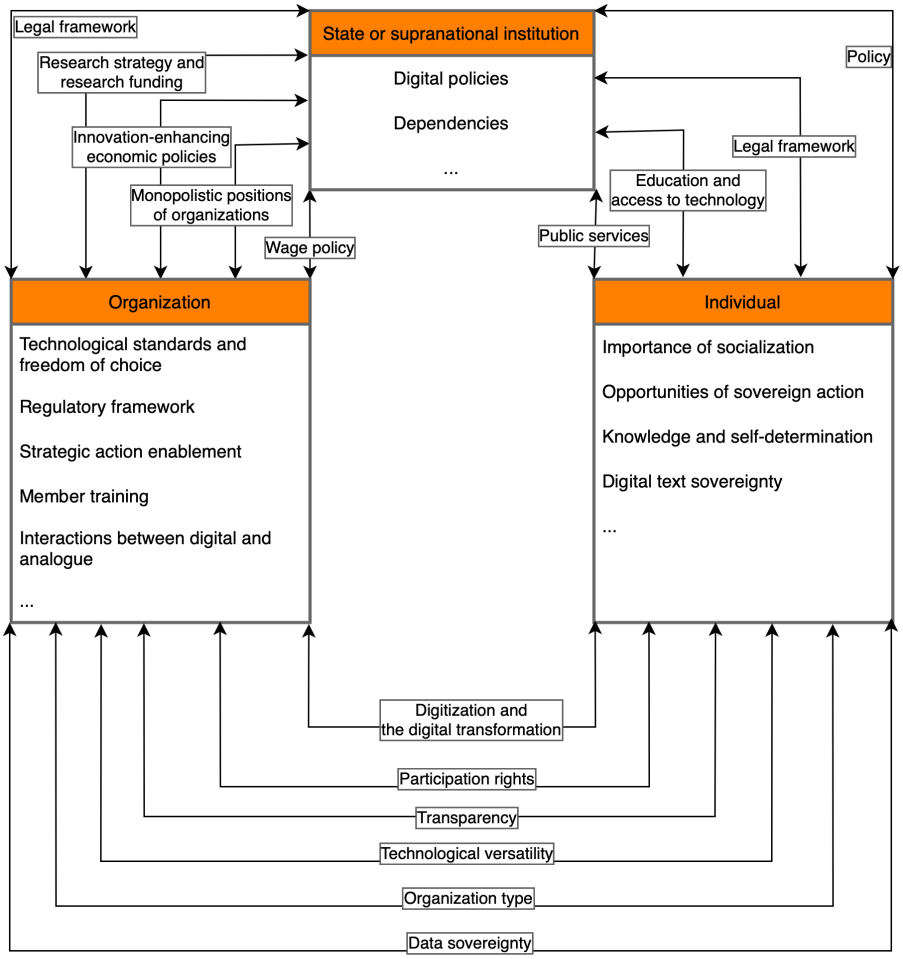


Fig. 1: Layered Model of Digital Sovereignty

2.1 Model Layers

Our model for digital sovereignty presented in this paper consists of the following three layers: the first layer addresses states and supranational institutions (such as the EU), the second layer addresses organizations including both companies and non-profit organizations, private or governmental, and the third layer addresses the individual citizen.

In this section, we describe each of the three levels using statements made in the world cafés which we had used as a method in previous research (Fries et al., 2023) as well as broader literature. The following section then proceeds to describe the relationships between layer elements in terms of how they influence each other in terms of digital sovereignty. In this second section, we have also included statements from the world cafés as well as relevant literature.

2.1.1 State or Supranational Institution Layer

Digital sovereignty on state or supranational level is closely connected to the underlying layers of organization and individuals, as the state can be seen as a superordinate entity. The reference points of the state layer can be divided into three areas: international, EU and national. The characteristics differ accordingly from one layer to another, but on the other hand they are strongly branched together, which poses a challenge of classification and demarcation. Moerel and Timmers offer a different view regarding the state perspective by relating the term "sovereignty" to the dimensions of cyber resilience of critical systems, processes, and data; control of economic ecosystems; as well as trust in the rules of law and democratic processes (Moerel & Timmers, 2021).

The results of the world café focused on the digital policy of a state entity layer on national, EU, and international level and on the national education system, as well as on dependencies on other states. Workshop participants linked the term "digital sovereignty" to autarky, interoperability, portability, and open standards.

Digital policy. In 2021, a letter from ministers of four European countries was written to the president of the European Commission explicitly on the subject of digital sovereignty. In addition to concrete deficits and inferential measures, an understanding of digital sovereignty on European level has been established. In this context, digital sovereignty is outlined as a guiding theme of the digital transformation and digital policy to encompass the interests of society (individuals), business (organizations), and the state (international, EU, national) in equal measure (Merkel, Frederiksen, Marin, & Kallas, 2021). Considering the political texts above, when referring to technological innovation one can posit that being a top innovator (as a state) makes it easier to remain a top innovator or even increase the advantage. As innovation can be considered a network effect, one may consider investigating preferential attachment (as suggested originally by Barabási and Albert (1999)) in innovation networks in the public and private sector. Furthermore, the education system plays a key role in the context of

digital sovereignty on both the state and individual layer. This is confirmed by an expert opinion based on Blossfeld, who describes the education system as a model of digital sovereignty (Blossfeld et al., 2019).

Dependencies. According to the results above, this is about building on strengths and reducing strategic weaknesses, not excluding other states or taking a protectionist approach. In the context of supranational striving for digital sovereignty, research literature thus also refers to "the creation of a sovereign common good protected by a border and administered by a common rule imposed on incoming actors" (St.-Hilaire, 2020, p. 144). It is also about developing a global supply chain in the interests of all participants (Merkel et al., 2021). This is in accordance with workshop results where, besides the terms mentioned above, the balance of power and digital access on a mental and material level, were outlined as characteristic in the field of digital sovereignty. Based on the dimensions of Moerel and Timmers, the description of digital sovereignty is outlined on the basis of case studies from various research projects, including deficits and challenges for the Dutch state (Moerel & Timmers, 2021). Similar views can also be found in the current German coalition agreement, which characterizes digital sovereignty as the self-determined development of digital innovations and the use of digital infrastructure and further discusses the topics of digital civil rights and IT security as well as the use of data and data law (Scholz, 2021). This socio-political demand for digital human rights has long been reflected in scientific discourse (Benedek, 2019). Compared to the results of the workshop, characteristic patterns can be identified as reference was also made to digital rights such as interoperability and portability, as well as open standards and European ecosystems, such as 5G or artificial intelligence.

2.1.2 Organization Layer

Switching perspectives towards organizations, we provide an overview of the topic clusters which came up during our world café. This is supplemented by a discussion again correlating these statements to findings in the literature.

Technological standards and freedom of choice. This includes quite frequently occuring concerns regarding e.g. the dependency of an organization upon specific technologies, data formats, or standards. Supporting and developing standards is a discussion that has been going on for decades. For instance, the question of flexible standards and how some healthcare information systems (HMIS) support them has been raised in (Braa, Hanseth, Heywood, Mohammed, & Shaw, 2007).

Regulatory framework. This refers to governance issues, norms, and laws within which any organization exists. Our participants mentioned these issues less frequently and in rather generic fashion, e.g., *What are the legal requirements?* Ranchordas takes note in (Ranchordas, 2014, p. 202) that legislation still has trouble keeping up with technological advancement. This is highly apparent in the issue of autonomous driving, the vision of which goes back to 1935 (in the General Motors film *The Safest Place* by

Jam Handy). Even after relevant technologies appeared in the late 1960s, a significant period of time passed until the issue was first embedded in a legal framework in Germany (2017) and later in other countries.

Strategic action enablement. Points included here refer to the longer-term ability of organizations to act strategically and where technology is a key enabler, e.g., which financial resources are allotted to technology purchasing. We also discuss this below with respect to small and medium enterprises (SMEs), which arguably have more limited strategic resources, as was painfully demonstrated by their performance during the COVID-19 pandemic. This has caused multiple weak points in the organizations' strategies (or lack thereof) in dealing with remote work to surface.

Member training. A core pillar of the organization is its members. As will be mentioned in the section concerning the individual, they need to be supported and trained in new technologies introduced by the organization. We will not specify this point further as it is covered in more depth in section 2.1.3.

Interactions between digital and analog. Even though this cluster was smaller, there have been questions about the way the digital and the analog interact and coexist within organizations. This opens up the field to further questions, for instance, the question of whether such interactions can be intentionally modeled or whether this should, or even can be, a binary choice.

Another interesting fact we encountered while doing literature research was that there are very few results which refer both to digital sovereignty and to organizations in general. An exception with regard to scientific reflection is the work of Hartmann (Hartmann, 2020, 2022). He understands digital sovereignty of organizations as analogous with digital sovereignty of individuals and seeks to make it measurable in terms of the degree of agency and control in the digital space. In this sense, Lehmann and Dörr also understand digital sovereignty with regard to SMEs as being able to inform oneself independently about relevant technologies and new technical possibilities, in order to be able to choose between several options – so that the questions "What does digitalization mean for my company, and how do I set the strategic course?" can be answered (Lehmann & Dörr, 2022, p. 14). In doing so, the authors conceptually include the training of digital skills of managers and employees (Lehmann & Dörr, 2022, p. 23). Nevertheless, the vast majority of literature results have either an individual or political focus. This may be interpreted as a sign that organizations do not think of their work in this direction in terms of "digital sovereignty" or prefer not to use the term in public discourse. It may also be that very few organizations see this as an overarching strategic goal, and thus simply do not frame their work in these terms. When used in a political context, digital sovereignty is usually limited to the contexts of data sovereignty or technological sovereignty. In the literature, too little focus is set on the fact that organizations are part of ecosystems and thus digital sovereignty discussions should be focused on the entire chain. We argue that the chain perspective is a more consistent one and thus the single-organization perspective, while valid, has to be evaluated in the context of its chain. The type of chain is here a criterion which researchers should use

at their own discretion. For instance, in a logistics context, the supply chain perspective would be relevant and one would thus choose to look at the suppliers, customers, and other partners of a chosen organization.

Our search suggests that organizations currently do not approach digital sovereignty at a strategic level. A switch in perspective is needed from the strategic to the more tactical or operative layers. Our workshop results suggest that technological and data sovereignty are topics of interest for organizations (e.g., the question of how dependent an organization is on software or hardware providers, or the question of proprietary digital formats and standards). When comparing these with the literature, we see mixed results.

Like Hartmann, Dörr, and Lehmann, which have already been mentioned, Pentenrieder, Bertini, and Künzel (2021) also frame digital sovereignty in the context of SMEs as the ability to maintain an overview of new technical capabilities in order to be able to choose the best among different digital options. They further argue that some technologies can serve as enablers more than others for continued self-determination of business processes and name artificial intelligence (AI) technologies as an example in the context of the machine tool industry. A frequent association between the companies' ability to innovate and digital sovereignty is also suggested in previous work. For example, Bogenstahl and Zinke (2017) focus on trends such as intelligent algorithms, big data, artificial intelligence (machine learning), or the Internet of Things. Competence development in these areas has enjoyed a wider attention, for instance regarding machine learning in (Panusch, Büscher, Wöstmann, & Deuse, 2022).

2.1.3 Individual Layer

The concept of sovereignty is also a way of comprehending the human individual – to be understood as an open system – in its entirety. From this perspective, sovereignty is a term that reflects the real wealth of human beings. It can be linked and supplemented with terms from a wide variety of sciences. These include terms such as subject and subjectivity, individual and individuality, personality and personal development, autonomy and maturity. These terms are closely related to issues and problems of society, communities, families and parents, as well as issues of education and formation. The associated assumptions allow people to use the possibilities for self-realization and individual development. In the discourse on digital sovereignty, these fundamental thoughts are particularly reflected in the topic of self-sovereign identity (SSI), which incorporates the idea of sovereignty in its very name. This is followed, especially in the German discourse, by a striving for free self-determination in the digital space that is seen as worthy of protection and that is also subject to corresponding legal regulations.

Importance of socialization. The socialization process is particularly important for the quality of development opportunities. However, socialization can only take place within the framework of opportunities given to an individual. Even under the assumption of different social and cultural conditions in large parts of the world, the political and

economic conditions in many countries of the world differ considerably, and thus also the development opportunities for individuals.

Opportunities of sovereign action. In increasingly digital societies, this becomes particularly relevant when individuals of all ages are confronted with very specific requirements. This includes the acquisition of mechanisms and techniques for accessing and using digital technologies (Pohle, Thiel, et al., 2021). Above all, it creates associations with questions about the acquisition of knowledge and skills. Yet it concerns more than just consuming content. In particular, the focus is on the means available to individuals for self-actualization in an increasingly digital world, for the pursuit of personal goals and the fulfillment of tasks. Regarding this, it is important to identify the opportunities that individuals have to assert themselves in a world shaped by digitalization and to get involved (or not) in social reproduction processes. The concept of digital sovereignty – so far almost exclusively used in legal terms and anchored especially in political theory, as mentioned before – has become an important concept in politics and is more and more the subject of scientific analysis. It attempts to describe a very ambivalent structure of state control mechanisms, economic and political interests, and personal development of individuals at different levels of action. It is evident that these developments are associated with qualitatively new challenges for people of all ages in the 21st century.

Knowledge and self-determination. In our world cafés, we discussed in more detail the extent to which digitalization affects the sovereignty of the individual – in a positive or possibly negative way. Indeed, it is less evident whether the digital separates people from sovereignty or connects them to it. Can sovereignty be digital at all? Can people act and make sovereign decisions in the digital world? What about data sovereignty and data ownership (Hummel, Braun, & Dabrock, 2021): does data belong to the person to whom it refers, or to those who collect it? It does not seem to make sense to try to find out how individuals can maintain sovereignty over their digital traces. Rather, the question arises as to how individuals can be put in a position of knowing and sovereign self-determination (Friedrichsen & Bisa, 2016). Two fundamental perspectives emerge. The first is an individual, very human perspective, because first of all, each individual independently decides whether or not to use digital technologies. For example, in our world café, the question arose: *Can digital sovereignty also mean that an individual exercises a sovereign choice to remain analog, and should there be a legal provision for this?* In practice, individuals are free to use an analog watch or a smartwatch. During the COVID-19 pandemic, for example, individuals were able to request a digital vaccination certificate or use a printed version. Other examples can certainly be found. However, in an increasingly digitized world, it is also more difficult to evade digital technologies. This brings us to the second perspective, which, on the other hand, concerns the design of the framework conditions in a digital world in which individuals act. These framework conditions refer to a digital text world. In this digital text world, skills are required that enable the individual to work with texts that were generated under the conditions of **digitization** (Frederking & Krommer, 2019). These texts differ significantly from those texts that are typographically and scriptographically generated.

Digital text sovereignty. "Digital text sovereignty" (Frederking & Krommer, 2019) as a terminus technicus is an essential variable for coping with these changes at the individual layer. As mentioned before regarding SSI, digital text sovereignty already addresses the issue of sovereignty by name. In our workshop, we discussed the various implications of this skill. Following the approach of digital text sovereignty has various advantages in this context, because in international educational policy discourse and international educational research reference is clearly made to the digital world. But the associated aspects of ability are described in very different terminology. In international discourse, the main focus is on "skills" (OECD, 2016) and "digital literacy" (OECD, 2011). In Germany, for example, the terms "digital competencies" and "digital education" are used. In this context, the construct of digital sovereignty has gained importance in educational policy, educational research, and educational theory. Closely related to this is digital text sovereignty, which is differentiated by Frederking and Krommer (2019) in an 8-level model (e.g., the semantic level of digital texts, the level of the source code of digital texts, the level of intentionally of digital texts). The challenges for the individual can be summarized with the sovereign functional-technical use of digital texts and a sovereign self-reflective attitude compared to digital texts.

2.2 Layer Relationships

While digital sovereignty is characterized differently on three layers, the layers are not isolated from each other. Rather, there are various connections between them such as dependencies, empowerment, or even conflicts. The characteristics of the relationships are described in the following.

2.2.1 Relationship Between State and Organization

According to the characteristics of the previous chapters, the relationship between the state institution layer and organizations can be seen from two opposing perspectives. On the one hand, the positive or negative influence of the state on organizations must be taken into account. On the other hand, one should consider how the digital sovereignty of organizations can be influenced by the state: The state cannot make society digitally sovereign – it is the organizations that build the digital infrastructure which may facilitate digital sovereignty.

However, the workshop results suggest that in discourse, the participants differentiated between organizations within a state (jurisdiction) and organizations outside it (such as supranational organizations, e.g., the UN). While international legal frameworks exist (such as the International Trade Law), they are often difficult to navigate, and some organizations then tend to adhere to local laws when uncertain. In addition, the following five potential mutual relationship aspects were identified:

Legal frameworks. Central guidelines and standards can be defined and passed by the state institution layer in consultation with organizations, and ideally provide support

the introduction of transparency in the information base or decision-making basis for companies. Within the workshop, data protection and data security specifications were also given as examples in this regard: they determine how personal data can be processed and why cybersecurity investments and technologies are necessary for achieving compliance with regulations. Regulations shape competition in the markets and the competitiveness of organizations in a global market. This contributes to shaping the digital sovereignty of organizations, as these will not need additional assistance in identifying or implementing standards, but instead use the embedded guidelines provided by the state. An example of this is basic IT security guidance, which in some states is implemented as a certifiable standard. On the other hand, such regulation also tends to be implemented as required, no more and no less. Hence, if the regulation itself is watered down in a political compromise or is loosely enforced, its implementation will tend to have the same quality.

Research strategy and research funding. Research programs are designed to add value to the digital sovereignty of both the state and organization layer. Close cooperation between research and practice, i.e., modern ecosystems facilitates novel technologies, e.g., blockchains, and enables the resulting business models onto the market to create value. This effect of public research and development (R&D) funding has been the object of public evaluation under general additionality principles, leading to mixed results according to Fantino and Cannone (2014). The results also vary by state, so it would be interesting to investigate potential correlations of positive additionality and digital sovereignty in the future.

Innovation-enhancing economic policies. On the one hand, this can promote the availability of alternative products or services and thus positively influence self-determined action at national, EU, and international level. On the other hand, these types of regulations can act as barriers to innovation. Organizations are thus restricted in their right of co-determination and cannot contribute to technological diversity. Furthermore, economic policies support the state layer in understanding and expanding its own options for action by means of innovations by organizations. Social networks have been voiced as an example. Due to the strong increase in users, the state has become aware of a need for action with regard to regulation and control.

Monopolistic positions of organizations. A balance of power defines the relation between state and organization. Our participants expressed concerns that modern platforms or digital organizations may yield monopolistic positions. Examples include social media platforms and tech companies. Through technologies, standards, and patents, large corporations can encroach upon the sovereignty of the state layer, especially when regulation lags behind technical advances.

Wage policy. Both the relationship between the individual and the state and between the state and organization become clear within this topic. As a rule, IT professionals are more highly paid in the private sector than in the public sector. This can lead to an imbalance of digitally sovereign competencies that culminates in the business sector on national, EU, and international level.

2.2.2 Relationship Between State and Individual

The Karlsruhe theses on the digital sovereignty of Europe (Beyerer et al., 2018) place digital sovereignty in the context of individuals, companies, and the European community of states in the digital space. The theses show very clearly that the perspective of the individual on digital sovereignty must always be linked to the perspective of the state. However, democratic states and democratically elected governments are made up of the people. The state represents the people. The constitutions of democratic states are constituted by the basic rights of the citizens. The people can assert these rights against the state. Therefore, the individual development of digital sovereignty is strongly linked with the development of democracy in the states. This relates to the following fields of interaction between state and individual.

Policy. As part of the political decision-making process and political control, the state creates a framework for action in which the people can make digitally sovereign decisions. Political decisions are intended to protect the rights of citizens in the digital space (Pohle et al., 2021). This includes the prevention of comprehensive communication surveillance of citizens, surveillance of government members (e.g., by the opposition), spying on companies and media, up to infiltration of the entire network communication structure. Without political decisions, there is a high risk of an attack on the fundamental rights and freedoms of citizens and a threat is posed to an open, free and democratic society.

Legal framework. In an open and democratic society, it is presupposed that there are spaces in which individuals can move and communicate without being observed (Pohle et al., 2021). In this context, probably the most important aspect for securing or developing digital sovereignty is the adoption of the European General Data Protection Regulation (GDPR). The GDPR ensures a high level of data protection throughout Europe. All providers who offer their services in Europe are subject to it (GDPR, 2018). In addition, concepts such as data security, data protection, patent and copyright law, cloud computing, the regulation of e-commerce and the consideration of general contractual and business conditions in the digital economy can be described as part of the legal framework. However, recent years have shown that the dynamic development of technology, media, and communication is far ahead of the development of appropriate regulatory concepts. This means that digital sovereignty must always be linked to the underlying legal framework.

Education and access to technologies. Education is one of the key parameters for achieving digital sovereignty. In particular, institutionalized, compulsory general education with a focus on acquiring digital skills is a basic requirement for sovereign participation in society. Training and adult education is also important (Gegenfurtner, Schmidt-Hertha, & Lewis, 2020). A certain level of digital competence is the prerequisite for digital sovereignty (Blossfeld et al., 2019). It is the duty of the state to create framework conditions in and outside of educational institutions (e.g. schools, colleges, universities) in which citizens can acquire digital skills. This refers not only to the

linking of education and digitization in general. In fact, it requires a profound and sustainable implementation of this topic in the curricula of general school education, as well as university study programs. However, access to digital technologies must also be made possible. For example, achieving data sovereignty requires knowledge of different media, relevant security aspects, and potential risks of their use. Furthermore, certified IT products, systems, and network infrastructures that guarantee secure data transmission are also required.

Public services. Public services can be seen as an interface between the obligations of the state and the interests of the citizens. This shows that citizens are increasingly demanding transparency, efficiency, and responsiveness from public authorities, public administration, and public organizations in the context of digitization. Thus, the increasing adoption of digital technologies represents a key element of governments' response to such requirements. The use of digital technologies to edit and process sensitive citizen data also requires completely new concepts in terms of data security, availability, and communication. This not only relates to the technical perspective of the IT systems, but also to the ICT competencies (e.g., knowledge of current IT technologies and support processes, understanding of technology, knowledge of software architecture) of the individuals involved. In the interaction of political decisions, legal framework conditions, the design of educational processes in a digital world, and the granting of access to digital technologies, public services are an important instance of sovereignly acting citizens in democratic states.

2.2.3 Relationship Between Organization and Individual

As we have seen in section 2.1.2, the political and scientific discourse on digital sovereignty appears to primarily revolve around the state and the individual. Our workshop findings seem to confirm this gap in the area of relationships between organizations and individuals: we had markedly less input on that level compared to any of the others discussed in sections 2.2.1 and 2.2.2.

In structuring future research paths, research questions regarding how organizations and individuals influence each other's digital sovereignty need to be phrased better.

Emerging from our workshop are six candidates for criteria, which we present in the following paragraphs. These criteria constitute early results, and therefore we also propose relevant research question for follow-up work.

Digitization and the digital transformation are terms often mentioned in the context of digital sovereignty of organizations. Low digital competencies of workers may mean fewer possibilities for organizations to become more digitally sovereign. The converse is thus also possible when the organization itself is more digitally sovereign as a result of external or internal pressure, and thus influences workers or collaborators to also become more digitally sovereign as a result. There are hints of these in the efforts of governments to improve planning outcomes for local regions in Australia (Alam,

Erdiaw-Kwasie, Shahiduzzaman, & Ryan, 2018). A recent Italian study for public organizations suggests that there may be a significant downside for these organizations when workers do not possess e-skills required for digital transformation (Casalino, Saso, Borin, Massella, & Lancioni, 2020). In assessing the role of digital competencies in work engagement, Oberländer and Bipp (2022) suggest that digital communication and collaboration are central competencies at work, both in theory and practice, and that more work is needed in this area in order to understand the exact implications of said digital competencies.

Participation rights Another question arising from the workshop is whether or not there is a correlation among workers or associates having a say in the organization decisions and the digital sovereignty of said organization. In other words: Do more participation rights result in more digital sovereignty? This question may prove to be very challenging to answer. A literature review on the topic of participation suggests that there are vast amounts of scholarly material and views available on this topic (Lee, 2015). Thus it is necessary to be very careful and specific in choosing an interpretative framework for this in order to frame it for the context of digital sovereignty. In any case, the question of the extent to which individuals want to act with digital sovereignty at all is also highly relevant to security. It is possible that individuals would rather hand over responsibility to a technical system than take responsibility themselves and thus become a security risk themselves (as considered in Fries (2022)).

Transparency is also interesting from multiple points of view and on different levels. At the level of enterprise architecture, for instance, we would posit that a transparent application landscape plays an important role when a decision needs to be made as to which digital tool to use for a given task. On a higher level, market transparency seems to play a role. There are, for instance, companies as well as governmental agencies that help consumers get better overviews of several different offers in markets with high information asymmetry. This suggests that organizations can positively influence the digital sovereignty of individuals if no other non-trivial downsides are involved. A platform can use its positioning to take advantage of the information asymmetry by placing sponsored offers or by manipulating prices.

Technological versatility represents the experience and know-how in a wide spectrum of technologies and offerings (e.g., open source software or multiple potential providers for cloud computing for employees or associates could enable organizations themselves to become more digitally sovereign as this capability may enable better strategic choices to be made. The converse may also hold true when organizations apply anti vendor lock-in strategies and thus themselves opt to use open source software. In turn, the members of the organization are trained and can hence also apply these strategies in their personal software choices. An interesting question here is how such organizations perform in the long term, compared with organizations without these strategies.

Organization type may be an important factor in how digital sovereignty is represented. We should differentiate between commercial and nonprofit companies, between NGOs,

clubs, and government agencies. Also, their internal structure – centralized versus decentralized – may affect how they perceive and act with respect to digital sovereignty.

Data sovereignty reflects the ability of organization to control their data. An organization with a high degree of control may mean more data protection for individuals. Low capabilities in organizations, however, could lead to choosing providers with low quality or obscure data protection practices. We also consider good IT security practices to be part of data sovereignty.

3 Discussion and Limitations

Particularly in interdisciplinary approaches to digital sovereignty, the question arises as to whose digital sovereignty is being addressed and what digital sovereignty means for the respective entities. We developed a layered model to conceptualize the meaning of the term "digital sovereignty" on three layers (state or supranational institutions, organization, individual) as well as the relationships between the layers. This model attempts to provide guidance for research and practice – including policy- and decision-making – on the complex subject of digital sovereignty.

Nevertheless, the proposed layered model does not claim to be fully exhaustive, but is seen as a suitable way to conceptualize digital sovereignty. Other characteristics are possible that have not been included here. The model is a theoretical approach and does not directly serve the purpose of measuring digital sovereignty, although, it does enable operationalization.

4 Conclusion

The goal of the presented model for digital sovereignty is to provide guidance and orientation on how to achieve the political goal of digital sovereignty. We employ the well-established model as a kernel with three layers: state (or supranational institution), organization, and individual. Furthermore, we emphasize the relationships between the levels in favor of an increase in digital sovereignty when located in an overall systemic process. We explored it from various academic perspectives, drawing on expertise in cybersecurity, security of critical infrastructures, and resilience, as well as educational, psychological, and philosophical expertise. The world café method allowed us to capture the ideas and concepts which various disciplines associate, justify, and emphasize with sovereignty or digital sovereignty. The resulting model bridges political discourse with design-oriented, empirical, and hermeneutic disciplines and provides an idea of what various academic disciplines can offer for achieving digital sovereignty.

5 Acknowledgments and Contributions

The LIONS research project is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, which we gratefully acknowledge. dtec.bw is funded

by the European Union – NextGenerationEU. We also thank the LIONS consortium and our research partners for their contributions to this research activity.

Authors have been listed in alphabetical order. The contributions can be attributed to the coauthors as described below:

- Isabelle Fries contributed content from a philosophical point of view, ensured overall consistence, and wrote most of the introduction.
- Maximilian Greiner moderated world café tables and mainly contributed on the relationship between state and organization and the state layer.
- Manfred Hofmeier prepared and conducted the world café and contributed research design, discussion, and outlines of limitations.
- Razvan Hrestic co-organized the world café, and mainly wrote on the organization layer and relationship between the organization and the individual.
- Ulrike Lechner contributed to the research design, data collection, data interpretation, and the revisions of the article.
- Thomas Wendeborn contributed on the individual layer, the state of the art of digital sovereignty, and definitions of digital sovereignty.

References

AISBL, G.-X. E. A. f. D., & Cloud. (2021). *Gaia-X Dataspaces.* Retrieved from `https://www.gaia-x.eu/what-is-gaia-x/data-spaces` (last accessed 2022/04/20)

Alam, K., Erdiaw-Kwasie, M. O., Shahiduzzaman, M., & Ryan, B. (2018). Assessing regional digital competence: Digital futures and strategic planning implications. *Journal of Rural Studies*, *60*, 60–69.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *science*, *286*(5439), 509–512.

Benedek, W. (2019). International organizations and digital human rights. In B. Wagner, M. C. Kettemann, & K. Vieth (Eds.), *Research Handbook in Human Rights and Digital Technology. Global Politics, Law and International Relations* (pp. 364–375). Edward Elgar Publishing.

Beyerer, J., Müller-Quade, J., & Reussner, R. (2018). Karlsruher Thesen zur Digitalen Souveränität Europas. *Datenschutz und Datensicherheit - DuD*, *42*(5), 277–280. doi: 10.1007/s11623-018-0940-2

BITKOM. (2018). Digitale Souveränität. *Datenschutz und Datensicherheit - DuD*, *42*(5), 294–300. Retrieved from `https://doi.org/10.1007/s11623-018-0944-y` doi: 10.1007/s11623-018-0944-y

Blossfeld, H.-P., Wilfried, B., Hans-Dieter, D., Bettina, H., Olaf, K., Dieter, L., ... vbw – Vereinigung der Bayerischen Wirtschaft e.V. [Ed.] (2019, January). Digitale Souveränität und Bildung. Gutachten.
doi: 10.25656/01:16569

Bogenstahl, C., & Zinke, G. (2017). Digitale Souveränität – ein mehrdimensionales Handlungskonzept für die deutsche Wirtschaft. *Digitale Souveränität*, 65.

Braa, J., Hanseth, O., Heywood, A., Mohammed, W., & Shaw, V. (2007). Developing health information systems in developing countries: the flexible standards strategy. *Mis Quarterly*, 381–402.

Casalino, N., Saso, T., Borin, B., Massella, E., & Lancioni, F. (2020). Digital competences for civil servants and digital ecosystems for more effective working processes in public organizations. In *Digital Business Transformation* (pp. 315–326). Springer.

CDU, CSU, & SPD. (2013). *Deutschlands Zukunft gestalten: Koaltionsvertrag zwischen CDU, CSU und SPD.* Union Betriebs-GmbH.

Chin, Y. C., & Li, K. (2021). A Comparative Analysis of Cyber Sovereignty Policies in China and the EU.
doi: 10.2139/ssrn.3900752

Couture, S. (2020). *The Diverse Meanings of Digital Sovereignty.* Retrieved from `https://bit.ly/3kCDVYa` (last accessed 2022/04/13)

European Political Strategy Centre of the European Commission. (2019). *Rethinking strategic autonomy in the digital age.* Brussels. Retrieved from `https://bit.ly/3MSarkW` (last accessed 2022/04/23)

Fang, B. (2018). *Cyberspace Sovereignty: Reflections on Building a Community of Common Future in Cyberspace*. Springer.

Fantino, D., & Cannone, G. (2014). Evaluating the efficacy of European regional funds for R&D. *Evaluating the efficacy of European regional funds for R&D.*, 165–196.

Frederking, V., & Krommer, A. (2019, Mar). *Digitale Textkompetenz: Ein theoretisches wie empirisches Forschungsdesiderat im deutschdidaktischen Fokus.* Retrieved from `https://bit.ly/38fJoRU`

Friedrichsen, M., & Bisa, P.-J. (2016). Digitale Souveränität. *Vertrauen in der Netzwerkgesellschaft*.

Fries, I. (2022). “In Code We Trust"? Zur Vertrauens-Verheißung der Blockchain-Technologie. *Zeitschrift für Evangelische Ethik*, *66*(4), 264-276.

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., & Wendeborn, T. (2023). Towards a Layer Model for Digital Sovereignty: A Holistic Approach. In B. Hämmerli, U. Helmbrecht, W. Hommel, L. Kunczik, & S. Pickl (Eds.), *Critical Information Infrastructures Security* (pp. 119–139). Cham: Springer Nature Switzerland.

GDPR. (2018). General Data Protection Regulation (GDPR). *Intersoft Consulting, Accessed in October*.

Gegenfurtner, A., Schmidt-Hertha, B., & Lewis, P. (2020). Digital technologies in training and adult education. *International Journal of Training and Development*, *24*, 1–4. doi: 10.1111/ijtd.12172

Hartmann, A. E. (2020). Digitale Souveränität in der Wirtschaft – Gegenstandsbereiche, Konzepte und Merkmale. In A. E. Hartmann (Ed.), *Digitalisierung souverän gestalten: Innovative Impulse im Maschinenbau* (pp. 1–16). Springer. doi: 10.1007/978-3-662-62377-0{_}1

Hartmann, A. E. (2022). Digitale Souveränität: Soziotechnische Bewertung und Gestaltung von Anwendungen algorithmischer Systeme. In A. E. Hartmann (Ed.), *Digitalisierung souverän gestalten II: Handlungsspielräume in digitalen Wertschöpfungsnetzwerken* (pp. 1–13). Springer. doi: 10.1007/978-3-662-64408-9{_}1

Hummel, P., Braun, M., & Dabrock, P. (2021). Own Data? Ethical Reflections on Data Ownership. *Philosophy & Technology*(34), 545–572. doi: 10.1007/s13347-020-00404-9

Institut für Innovation und Technik. (2018). *Digitale Souveränität: Bürger, Unternehmen, Staat* (V. Wittpahl, Ed.) [BOOK]. Berlin: Springer Vieweg. doi: 10.1007/978-3-662-55796-9

Lee, C. W. (2015). Participatory practices in organizations. *Sociology Compass*, *9*(4), 272–288.

Lehmann, C., & Dörr, L. (2022). Digital souveräne Gestaltung von Services – ein marktfähiger Mehrwert? In A. E. Hartmann (Ed.), *Digitalisierung souverän gestalten II: Handlungsspielräume in digitalen Wertschöpfungsnetzwerken* (pp. 14–24). Springer. doi: 10.1007/978-3-662-64408-9{_}2

Merkel, A., Frederiksen, M., Marin, S., & Kallas, K. (2021, March). *DE-DK-FI-EE: Letter to COM President on Digital Sovereignty.* Retrieved 2022-04-08, from `https://politi.co/3FfWdrs`

Moerel, L., & Timmers, P. (2021, January). *Reflections on Digital Sovereignty* (SSRN Scholarly Paper No. ID 3772777). Rochester, NY: Social Science Research Network. Retrieved 2022-04-07, from `https://papers.ssrn.com/abstract=3772777`

Oberländer, M., & Bipp, T. (2022). Do digital competencies and social support boost work engagement during the covid-19 pandemic? *Computers in Human Behavior*, *130*, 107172.

OECD. (2011). *PISA 2009 Results: Students On Line*. doi: 10.1787/9789264112995-en

OECD. (2016). *Skills for a Digital World* (Tech. Rep.). Author.

Panusch, T., Büscher, J., Wöstmann, R., & Deuse, J. (2022). Konzept zur zielgerichteten Kompetenzentwicklung für Initiativen des Maschinellen Lernens. BT - Digitalisierung souverän gestalten II. In E. A. Hartmann (Ed.), (pp. 93–109). Berlin, Heidelberg: Springer Berlin Heidelberg.

Pentenrieder, A., Bertini, A., & Künzel, M. (2021). Digitale Souveränität als Trend? [PDF]. *Digitalisierung souverän gestalten*. (last accessed 2022/03/24) doi: 10.1007/978-3-662-62377-0{_}2

Pohle, J. (2020). Digitale Souveränität. *Handbuch Digitalisierung in Staat und Verwaltung*. doi: 10.1007/978-3-658-23669-4{_}21-1

Pohle, J., Thiel, T., et al. (2021). Digital sovereignty. In *Practicing Sovereignty: Digital Involvement in Times of Crises* (pp. 47–67). Bielefeld: transcript Verlag.

Ranchordas, S. (2014). Innovation-friendly Regulation: The Sunset of Regulation, the Sunrise of Innovation. *Jurimetrics*, *55*, 201.

Scholz, O. (2021). *Koalitionsvertrag 2021-2025: Mehr Fortschritt Wagen.* Sozialdemokratische Partei Deutschlands. Retrieved from `https://bit.ly/3seGljP` (last accessed 2022/04/09)

Seidel, I., & Bös, P. K. (2009). Grundlagen , Inhalte und Implikationen. , *18*(November).

Steinbach, J. (2019). *Souveränitätsfragmente. Ein Beitrag zur Literaturgeschichte der Souveränität und gegenwärtigen Herausforderungen der Rechtswissenschaften im Spiegel der Digitalisierung*. Mohr Siebeck.

St.-Hilaire, W. A. (2020). *Digital Risk Governance: Security Strategies for the Public and Private Sectors*. Springer.

Singing in the Rain

The False Promise of Sovereign Independence

Friedrich Lohmann[1]

Abstract: The chapter focuses on the call for (more) digital sovereignty that has frequently been made in the last years. The core feature of sovereignty that is particularly emphasized in politics is independence as a means to self-emancipation and better control. Such a claim for independence is, however, a false promise that is neither desirable nor realistic. The digital transformation goes along with a high number of threats and vulnerabilities which cannot be simply shut down by the proclamation of sovereign independence. It would, therefore, be better to emphasize resilience than independence, when speaking of digital sovereignty. In these times, sovereignty must be understood far more in terms of resilience than as a call for independence.

Keywords: Sovereignty, Resilience, Vulnerability

1 Digital Sovereignty as a Multi-Layered, All-Encompassing Concept of Self-Emancipation

"Digital sovereignty" is a frequently used term these days. It is employed (1) on the level of the individual, (2) for organizations, including both companies and non-profit organizations, (3) for states and for supranational institutions (Gesellschaft für Informatik e.V. 2020, p. 4; Fries et al. 2023). This multi-layered use of the word "sovereignty" is new. Even if there are facets of sovereignty on the individual level that have been used here and there in the past, such as "sovereignty of the consumer" or "sovereignty of the patient", the traditional understanding of the term "sovereignty" refers to the state, both in its internal structure (sovereignty of the king vs. sovereignty of the people) and its external relations (national sovereignty). We can, therefore, speak of an expansion of sovereignty discourse from the political domain to other layers of society. Sovereignty has become a goal for all, well beyond the affairs of the state.

This expansion is an interesting development I would like to reflect upon in the following paragraphs. It is remarkable with regard to the history of ideas. More than that, it says a lot about our current society. To put it in a nutshell: sovereignty is aimed for on all levels of action because there is a common feeling that we are no longer masters of our actions, that we are more and more dependent on others, and that we have to win back our freedom. In politics in particular, "digital sovereignty" is promoted as a means of self-emancipation and independence, which is very much in line with the

[1] University of the Bundeswehr Munich, Neubiberg, friedrich.lohmann@unibw.de

traditional use of the sovereignty concept in politics, now expanded to all dimensions of society.

Following this reconstruction which will be presented in more detail in the following two sections, I will add two critical sections that emphasize both parts of the notion of "digital sovereignty": (1) Is it recommendable to react to the challenges of the twenty-first century by proclaiming a concept of freedom as *sovereign independence*? (2) Can the *digital* sphere keep the promise of self-emancipation? My answers to both questions will be rather negative, and, therefore, the last section of my paper will propose to aim rather for resilience than for sovereign independence, given the phenomenon of human vulnerability, which cannot be simply erased by proclaiming independence and which is predominant in the digital sphere as well.

2 Sovereignty as Independence

Some proponents of sovereignty in our days use the concept as an equivalent of autonomy. One example is Paul Timmers, who explains sovereignty by referring to strategic autonomy (Timmers, 2019). Likewise, digital sovereignty and strategic or technological autonomy have been closely related to each other in a 2020 ideas paper edited by the Research Service of the European Parliament (EPRS, 2020). This is interesting in view of the expansion mentioned above, because the autonomy concept has undergone a history in its use that went in the exactly opposite direction than the notion of sovereignty: autonomy used to be a quality of a person and was only recently – as "strategic" autonomy – applied to an institution like the European Union.

There is, indeed, an overlap between both concepts: autonomy and sovereignty both emphasize the capability of self-determination. However, there are differences, too, which can be traced back to the original sense of both words. "Sovereignty" is the English translation of the French word "souverain" which in turn is the translation of the Latin "superanus", "above the others". It is very much in accordance with this original meaning when Jean Bodin, who in the sixteenth century coined the term for the political realm, defines sovereignty in the following way: "La souveraineté est la puissance absoluë & perpetuelle" (Bodin, 1576, I/9). It has no limits: "Or la souveraineté n'est limitee, ny en puissance, ny en charge, ny à certain temps" (Bodin, 1576, I/8). With these qualities "above the others", political sovereignty is thought to be an image of God's power and might (Deppisch, 2015, p. 26–27), and Carl Schmitt is certainly right when he takes Bodin's notion of sovereignty as a key example of the adaptation of a theological concept to politics (Schmitt, 1996 [1922]).

"Autonomy", on the other hand, is a combination of the two Greek words "autos" and "nomos," meaning "self-legislation." Its main proponent in the history of ideas is Immanuel Kant, for whom the notion of "law" in autonomy is crucial: to live an autonomous life does not mean to live without limits and dependencies, but to live according to obligations and laws as they are given by reason. Therefore, beyond the overlap in terms of a general notion of self-determination, there is a clear distinction to be made between the two concepts of sovereignty and autonomy, and this in particular with regard to their relationship to the idea of independence. "Autonomie bedeutet

also vielmehr Selbstbestimmung als Unabhängigkeit, wie dies – zumindest implizit – oft gedacht wird. So verstanden schließt Autonomie den Bedarf nach Hilfe durch andere keineswegs aus" (Dierks et al., 2001, p. 1–2).

This distinction seems to be seen by most of the proponents of digital sovereignty, because most of them avoid speaking of autonomy, while employing a notion of sovereignty as independence, thereby – often tacitly – emphasizing the quality that distinguishes sovereignty from autonomy (at least in the Kantian understanding of the latter). A good example is the definition of "digital sovereignty" as proposed in a study edited by the German Ministry of Economics: "Souveränität bezeichnet die Möglichkeit zur unabhängigen Selbstbestimmung von Staaten, Organisationen oder Individuen. Digitale Souveränität ist heute ein wichtiger Teilaspekt allgemeiner Souveränität, der die Fähigkeit zur unabhängigen Selbstbestimmung in Bezug auf die Nutzung und Gestaltung digitaler Systeme selbst, der darin erzeugten und gespeicherten Daten sowie der damit abgebildeten Prozesse umfasst" (BMWi, 2021, p. 11). The double mention of "unabhängig"/"independent" is key: it shows that the proclamation of sovereignty is driven by an interest in independence. As early as 2015, in one of the first programmatic papers on digital sovereignty, it was defined prominently as "independence" by Bitkom e.V., a federation of German enterprises: "Unter dem Begriff 'Souveränität' versteht man allgemein die Fähigkeit zu ausschließlicher Selbstbestimmung. Diese Selbstbestimmungsfähigkeit wird durch Eigenständigkeit und Unabhängigkeit gekennzeichnet. Sie grenzt sich einerseits von Fremdbestimmung und andererseits von Autarkie ab" (Bitkom e.V., 2015, p. 7). The same is true for the already mentionned EPRS ideas paper, to give a third example: "In this context, 'digital sovereignty' refers to *Europe's ability to act independently in the digital world*" (EPRS, 2020, p. 11; emphasis in the original).

3 The Struggle for Independence as a Struggle for Values and Self-Emancipation

The emphasis on independence (and, henceforth, the recourse to the notion of sovereignty) in the last quote gets its explanation in its immediate context within the EPRS ideas paper: "Strong concerns have been raised over the economic and social influence of non-EU technology companies, which threatens EU citizens' control over their personal data, and constrains both the growth of EU high-technology companies and the ability of national and EU rule-makers to enforce their laws" (EPRS, 2020, p. 11). In this quote, we not only find a good example for the three-layer approach to digital sovereignty (individual person – companies – states and supranational institutions), but also a common threat of control as the over-arching reason for envisioning it. Without expanding on the idea of "control" in the context of digitalization (see Tretter, 2022), it is obvious that the political striving for "digital sovereignty" is put forward as a reaction to a situation of crisis in which self-determination is threatened by forces from outside. The perceived levels of threat are succinctly described in a letter from four European prime ministers that was sent to the head of the European Commission in 2021: "The dependencies and shortcomings in European digital capacities, skills and technologies have become more apparent. A significant amount of

digital value-added and innovation takes place outside Europe. Data has become a new currency that is mainly collected and stored outside Europe. And fundamental democratic values are under severe pressure in the global digital sphere" (Merkel et al., 2021).

It is, therefore, more than a vague feeling of loss of control that stimulates the call for a renewed European sovereignty in the digital age. In a globalized society and economy, a lack of enforcement power and a growing dependence on others who do not adhere to the same moral principles put fundamental values at stake. Digital sovereignty is seen as a way to restore Europe's political identity, economic welfare, and cultural-moral vision in a process of self-emancipation.

On the national German level, the Grand Digital Strategy issued by the German government in 2022 takes a similar approach from a perception of threat to a concept of digital sovereignty. Threats are identified on all three layers: the individual (Bundesregierung, 2023, p. 24: "Eine moderne und den europäischen Grundwerten entsprechende Datenökonomie kann daher nur zusammen mit einem starken Schutz der Rechte der Verbraucherinnen und Verbraucher im digitalen Raum gedacht und umgesetzt werden."), companies ("Wir sichern funktionierenden Wettbewerb durch zeitgemäße digitale Ordnungspolitik für digitale Märkte. Dabei geht es um die Verhinderung wettbewerbsschädlicher Konzentration von Marktmacht genauso wie um die effektive Verhinderung konkreter Praktiken, die faire und bestreitbare Märkte besonders gefährden", ibid., p. 33), and the state ("Um die Kontrolle über die eigene IT sicherzustellen und insbesondere Informations- und Datenschutz gewährleisten zu können, muss die öffentliche Verwaltung unabhängiger von einzelnen Anbietern und Produkten werden", ibid., p. 48). The proposed answer to these threats is the transformation to a digital sovereign society with reduced dependencies ("Technologische und digitale Souveränität sind notwendig, um Handlungsfähigkeit zu stärken und Abhängigkeiten zu reduzieren", ibid., p. 2). These efforts for more digital sovereignty are presented as preconditions for an implementation of core values in times of digitalization ("Mit dieser Strategie wollen wir die Rahmenbedingungen verbessern und dazu beitragen, dass der digitale Wandel im Sinne einer nachhaltigen, vielfältigen, inklusiven und demokratischen Gesellschaft geschlechtergerecht und diskriminierungsfrei gestaltet werden kann und insbesondere Zivilgesellschaft, Wirtschaft, Bildung und Wissenschaft die Chancen der Digitalisierung und die Gestaltungsmöglichkeiten des digitalen Wandels im Sinne der Menschen nutzen können", ibid., p. 4–5).

4 The False Promise of Sovereign Independence in the Twenty first Century

The analysis of threats that was briefly presented in the last section seems evidence-based and correct. But is the proclamation of more independence and the struggle for sovereignty the right way to deal with these threats to self-determination? Bodin developed his concept of sovereignty in the post-Reformation period, which was shaped by conflicts all around at both national and European level. It was also a time in which an idea of absolute power ("puissance absoluë") was thriving, both in theology and in

politics. A concept of self-determination with an emphasis on powerful independence fitted well in this intellectual and historical context.

This is no longer the case in our times, and the presented threat analysis confirms this, somehow inadvertently. Bodin's sovereign republic does not thrive on threats, but on its own power. As an entity which claims to represent the almighty God on earth, it is above all possible threats, rather threatening others with its use of force (Loick, 2012). On the other hand, by enumerating threats to self-determination on all levels, the quoted strategic papers start from a completely different view of the preconditions of political action. In my analysis, they presuppose notions of weakness, individual rights, and interconnectedness that shape postmodern culture and are alien to Bodin's vision of politics.

(1) Weakness. Even if the emphasis on power and the presumed need of God-like strongmen has regained attraction and relevance in the recent past, modern societies have learnt first to accept and then to cherish a weaker posture in the centuries since Bodin, with the first half of the twentieth century and its two world wars perceived as an ultimate call for change. Modernity has become reflexive (Beck at al., 1994), acknowledging the mistakes of the past and the threats in the present and future times. Violence is no longer the preferred way to resolve conflict (Pinker, 2012). Threats and suffering, the "times of crisis" (Lodge & Wegrich,2012), are more emphasized than the promises of human progress (Horkheimer & Adorno, 2007). The postmodern society perceives itself as overcharged (Nassehi, 2021) and vulnerable (Rostalski, 2024). This experience of vulnerability on all levels of society (Lenz, 2009) is very much the backdrop for all threat perception that marks the current talk on restoring sovereignty. However, a vulnerable society that acknowledges its weakness never can be sovereign and independent in the proper sense. In our times, if talk of sovereignty is to be maintained, it must be acknowledged there can be nothing more than a "fragile" sovereignty (Essen, 2024).

(2) Individual rights. The notion of individual, human rights was from its inception very much guided by acknowledging vulnerability and weakness as essential parts of the human condition. "Empathy" was a key word for the first human rights activists in the eighteenth century (Hunt, 2008), and it has since then been a continuous motive for strengthening a moral and legal framework that starts with the individual and its need for protection. There have been efforts to rethink sovereignty as responsibility for the protection of human rights (Deng et al., 1996; Annan 1999), which is a valuable development in reconnecting the concepts of sovereignty and human rights, even if it has to be acknowledged that there has been, since the 2000's, a "sovereignty backlash" in the sense of "sovereign absolutism" in politics (Traub, 2009; quotes: p. 80 and p. 76). The understanding of sovereignty as independence, in any case, is not reconcilable with the feeling of mutual and universal responsibility for the other members of the human family and their protection that is inherent to the idea of human rights.

(3) Interconnectedness. The universal human rights culture is the moral answer to the experience of connectedness that shapes modernity and its continuous efforts of globalization, both on the economic and cultural level. It upends the one-way understand-

ing of this connectedness – as a means for imperial domination and economic exploitation – and changes it into a notion of mutual interconnectedness in which the same rights are given to all (Linklater, 2010). In the times of ecological crisis, the circle of interconnectedness must be expanded to a vision of being in which everything is related to and dependent on one another (Latour, 1993). This vision of being is irreconcilable with an idea of sovereignty as independence. Or, in the words of Judith Butler: one has to get rid of sovereignty in order to become human (Butler, 2007, p. 11; cf. Lieb 2022). In particular, an autarkic understanding of sovereignty implies a denial of interconnectedness. The fact that many proponents of digital sovereignty emphasize the difference between sovereignty and autarky confirms that such a (mis-) understanding is real and obvious (see, e.g., Bitkom e.V., 2015, p. 7: "Autarkie ist in einer global vernetzten digitalen Welt weder zu erreichen noch anzustreben."; BMWi, 2021, p. 10: "Jedoch ist Autarkie im Sinne einer (vollständigen) Unabhängigkeit nicht zwingend mit Souveränität gleichzusetzen."). It is very much present in populist political discourse from left to right, like in Vladimir Putin who recently pledged to make "sovereignty" one of the key aims of his fifth term in the Kremlin: "We must remember and never forget and tell our children: Russia will be either a sovereign, self-sufficient state, or it will not be there at all" (Voice of America, 2023). The same vision of sovereignty was promoted by Boris Johnson when he declared that the United Kingdom had "recaptured sovereignty" by Brexit: "We will rediscover muscles that we have not used for decades – the power of independent thought and action" (Politico, 2020).

5 Singing in the Rain

If we analyze the programmatic calls for "more" digital sovereignty in our days, we find two intentions behind those calls, echoing the very ambivalent view of the ongoing digital transformation of society by the broad public: (1) prophetic: we need more sovereignty and independence, and digitalization is the best way to fulfil this goal; (2) apologetic: digitalization is perceived by many people as more of a threat than a promise, so we must show that it actually increases freedom instead of limiting it.

(1) The prophetic use: sovereignty is possible *because of* digitalization. Digitalization and the digital transformation of society are often presented as a "must" by their proponents (see, e.g., the many "musts" in the "Aktionsplan Digitale Souveränität" issued by Bitkom, Bitkom e.V., 2015, pp. 16–18). One of the reasons for this presumed necessity is the contention that freedom and sovereignty are enhanced by digitalization. Some of the proponents of blockchain technology, for example, promote it as a way to be liberated from human intermediates in transactions (Nakamoto 2008, cf. Fries 2022). The digitalization of patients' records is introduced by the German government as a way of enhancing the sovereignty of the patient ("Elektronische Patientenakte für mehr Patientensouveränität" is one of the headings in Bundesministerium für Gesundheit, 2024). On the level of companies, the potential of the digital transformation to increase productivity and to enable more choice in the supply and demand chains are emphasized over and over again. There is a feeling of "technological solutionism" (Corballis & Soar, 2022, p. 2) that characterizes all these efforts to promote digitalization as a means to gain freedom and independence.

(2) The apologetic use: sovereignty is possible *in spite of* digitalization. These prophetic visions are, however, overshadowed by a deep and somehow vague feeling of distrust with regard to digitalization in broad parts of the public. They are like singing in the rain, pretending that everything is fine while it is not. Therefore, many of the statements in favor of digital sovereignty have an apologetic intention: they admit the threat to sovereignty that comes with digitalization and contend that these threats should, and can, be overcome. This approach is clearly evident in the German Grand Digital Strategy: the threats and dependencies it presents in order to make the claim for digital sovereignty (see above section 3) are, for the most part, threats and dependencies that stem from the digital transformation of politics, economics, and the society as a whole. The concentration of market power with its inherent growing dependencies is, for example, mentioned as a characteristic particularly of *digital* markets. On the level of the individual, the notion of "data sovereignty" (Hummel et al., 2021) addresses the widespread fear that digitalization (e.g., of patient data) actually implies a loss, and not a gain, in the protection of data, understood as self-determination with regard to the use of data (see, e.g., Bundesregierung, 2023, p. 29: "Wir werden verschiedene Datenräume domänenübergreifend miteinander vernetzen. Ziel ist ein sektorübergreifendes digitales Datenökosystem, in dem Daten unter Wahrung der Datensouveränität und des Datenschutzes zwischen Akteuren geteilt werden können. Hierzu unterstützen wir die Entwicklung eines universalen globalen Datenstandards und etablieren dafür strategische internationale Partnerschaften."). The apologetic posture implies that digital sovereignty is not an automatic outcome of digitalization; rather, it has to be secured by particular measures against risks that are growing with the digital transformation of society.

However, if these threats, risks, and dependencies are real and caused by digitalization, it hardly seems recommendable to go for the long haul and call for digital *sovereignty*, which is, for memory, defined as the "capability of exclusive [!] self-determination" (Bitkom e.V., 2015, p. 7, see above section 2). In the following section, I will plead for a more prudent and modest approach which emphasizes the capability to resist threats to security and freedom as a goal of political governance, instead of the false promises connected with the call for sovereignty as independence. The term which has become familiar as a label for this capability is "resilience."

6 Conclusion: Digital Sovereignty as Resilience

The term "resilience" has become even more frequent in recent years than "sovereignty". The main reason for this popularity seems to be that it fits very well with the widespread feeling of vulnerability that was already mentioned above (see section 4): to ask for resilience signifies not to minimize risks and vulnerabilities, but to ask for a better way to cope with them. It implies the notion of power (the power to resist), but not with the claim of absolute power as it is connected with the original idea of sovereignty. Power is conceived as a capacity to resist. The more complex the system, the more vulnerable it is. Therefore, resilience is a key factor when it comes to digitalization (Hiermaier & Scharte, 2018), given, for example, its dependence on the electrical grid (Knauf, 2020).

Given its popularity, it is no surprise that the papers on digital sovereignty which have been quoted above in this contribution mention "resilience" as well. "Critical infrastructures and technologies need to become resilient and secure" (Merkel et al., 2021). In the Grand Digital Strategy of the German government, resilience is often mentioned in the same context as digital sovereignty, stating that more digital sovereignty would bring more resilience to the German society (e.g. Bundesregierung 2023, p. 2: "Technologische und digitale Souveränität sind notwendig, um Handlungsfähigkeit zu stärken und Abhängigkeiten zu reduzieren. Dies wiederum sind Bedingungen für Wettbewerbs- und Innovationsfähigkeit sowie Resilienz.").

However, the relationship between sovereignty and resilience is far more complex than this. It is certainly true that a reduction of outward dependencies can increase resilience, but in the current situation digitalization is shaped by dependencies (Pohle & Thiel, 2021, p. 326–327), and getting rid of them does not seem easy. In addition, the sovereign lone rider is much more vulnerable than the one who looks for shelter in federated efforts in order to increase cybersecurity and to decrease uncomfortable dependencies (Autolitano, 2023). Sovereignty, when it is understood in the sense of independence and self-sufficiency, would actually stand in the way of vigilance, which is a key factor of a resilient society. This understanding comes close to autarky, which is – and rightly so – not seen as the right pathway by those who emphasize the difference between sovereignty and autarky. However, if we reject autarky, the idea that resilience is a follow-up to sovereignty, falls apart. It is rather the other way around: a society must first develop resilience before it can strive for self-determination. In our times, sovereignty must be understood much more in terms of resilience than as a call for independence.

The digital transformation offers enormous potential for all three layers of society: individuals, organizations, and state and supranational institutions. However, it comes with a lot of threats and vulnerabilities that must be coped with. In this situation, the call for digital sovereign independence in politics and in the public discourse projects false promises that are neither desirable nor realistic. On a rainy day, singing hymns of praise doesn't offer protection. It is better to strengthen resilience by putting on a coat.

References

Annan, K. A. (1999). Two concepts of sovereignty. *The Economist*, 352(8137), 49–50.

Autolitano, S. (2023). Why the EU should stop talking about digital sovereignty. https://www.cfr.org/blog/why-eu-should-stop-talking-about-digital-sovereignty

Beck, U., Giddens, A., & Lash, S. (1994). *Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order*. Polity.

Bitkom e.V. (2015). *Digitale Souveränität: Positionsbestimmung und erste Handlungsempfehlungen für Deutschland und Europa*. https://www.bitkom.org/sites/main/files/file/import/BITKOM-Position-Digitale-Souveraenitaet.pdf

Bodin, J. (1576). *Les Six Livres de la République.* Jacques du Puy.

Bundesministerium für Gesundheit (2024). *Elektronische Gesundheitskarte*. https://www.bundesgesundheitsministerium.de/themen/digitalisierung/elektronische-gesundheitskarte

Bundesministerium für Wirtschaft und Energie (BMWi) (2021). *Schwerpunktstudie Digitale Souveränität: Bestandsaufnahme und Handlungsfelder.* https://www.de.digital/DIGITAL/Redaktion/DE/Digitalisierungsindex/Publikationen/publikation-download-schwerpunkt-digitale-souveraenitaet-2021.pdf

Bundesregierung (2023 [2022]). *Digitalstrategie: Gemeinsam digitale Werte schöpfen.* Aktualisierung, Stand: 25.04.2023. https://digitalstrategie-deutschland.de/static/fcf23bbf9736d543d02b79ccad34b729/Digitalstrategie_Aktualisierung_25.04.2023.pdf

Butler, J. (2007). *Kritik der ethischen Gewalt.* Suhrkamp.

Corballis, T., & Soar, M. (2022). Utopia of abstraction: Digital organizations and the promise of sovereignty. *Big Data & Society,* 1–12. DOI: 10.1177/20539517221084587

Deng, F. M., Kimaro, S., Lyons, T., Rothchild, D., & Zartman, I. W. (1996). *Sovereignty as Responsibility.* The Brookings Institution.

Deppisch, A. (2015). *Die Religion in den Werken von Jean Bodin und Michel de Montaigne: Ein Vergleich.* Dissertation Würzburg. https://opus.bibliothek.uni-wuerzburg.de/opus4-wuerzburg/frontdoor/deliver/index/docId/12041/file/deppisch_aaron_die+religion.pdf

Dierks, M.-L., Bitzer, E.-M., Lerch, M., Martin, S., Röseler, S., Schienkiewitz, A., Siebeneick, S., & Schwartz, F.-W. (2001). *Patientensouveränität: Der autonome Patient im Mittelpunkt.* https://elib.uni-stuttgart.de/bitstream/11682/8693/1/AB195.pdf

Essen, G. (2024). *Fragile Souveränität: Eine Politische Theologie der Freiheit.* Mohr Siebeck.

European Parliamentary Research Service (EPRS) (2020). *Digital sovereignty for Europe.* https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf

Fries, I. (2022). "In Code We Trust?" – Zur Vertrauens-Verheißung der Blockchain-Technologie. *Zeitschrift für Evangelische Ethik* 66(4), 264–276. https://doi.org/10.14315/zee-2022-660405

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., & Wendeborn, T. (2023). Towards a Layer Model for Digital Sovereignty: A Holistic Approach. In B. Hämmerli et al. (Eds.): *CRITIS 2022, LNCS 13723*, 119–139.

Gesellschaft für Informatik e.V. (2020). *Schlüsselaspekte digitaler Souveränität.* https://gi.de/fileadmin/GI/Allgemein/PDF/Arbeitspapier_Digitale_Souveraenitaet.pdf

Hiermaier, S., & Scharte, B. (2018). Ausfallsichere Systeme: Resilienz als Sicherheitskonzept im Zeitalter der Digitalisierung. In R. Neugebauer (Ed.): *Digitalisierung: Schlüsseltechnologien für Wirtschaft & Gesellschaft* (pp. 295–310). Springer Vieweg.

Horkheimer, M., & Adorno, T. W. (2007). *Dialectic of Enlightenment: Philosophical Fragments*. Stanford University Press.

Hummel, P., Braun, M., Augsberg, S., Ulmenstein, U. v., & Dabrock, P. (2021). *Datensouveränität. Governance-Ansätze für den Gesundheitsbereich*. Springer VS. https://link.springer.com/book/10.1007/978-3-658-33755-1

Hunt, L. (2008). *Inventing Human Rights: A History*. Norton & Company.

Knauf, A. (2020). *Urbane Resilienz gegenüber Stromausfällen in deutschen Großstädten*. Springer VS.

Latour, B. (1993). *We Have Never Been Modern*. Harvard University Press.

Lenz, S. (2009). *Vulnerabilität Kritischer Infrastrukturen.* Bundesamt für Bevölkerungsschutz und Katastrophenhilfe.

Lieb, D. (2022). Souveränität. In M. Feldmann et al. (Eds.), *Schlüsselbegriffe der Allgemeinen Erziehungswissenschaft: Pädagogisches Vokabular in Bewegung* (pp. 406–413). Beltz Juventa.

Linklater, A. (2010). Global civilizing processes and the ambiguities of human interconnectedness. *European Journal of International Relations,* 16(2), 155–178. https://doi.org/10.1177/1354066109350796

Lodge, M., & Wegrich, K. (Eds.) (2012). *Executive Politics in Times of Crisis.* Palgrave Macmillan.

Loick, Daniel (2012). *Kritik der Souveränität.* Campus.

Merkel, A., Frederiksen, M., Marin, S., & Kallas, K. (2021). *Joint letter to European Commission President Ursula von der Leyen.* https://news.err.ee/1608127618/estonia-eu-countries-propose-faster-european-digital-sovereignty

Nakamoto, S. (2008, October 31). Bitcoin: A Peer-to-Peer-Electronic Cash System. Satoshi Nakamoto Institute. https://nakamotoinstitute.org/bitcoin

Nassehi, A. (2021). *Unbehagen: Theorie der überforderten Gesellschaft.* C.H.Beck.

Pinker, S. (2012). *The Better Angels of Our Nature: Why Violence Has Declined.* Penguin Books.

Pohle, J., & Thiel, T. (2021). Digitale Souveränität – Von der Karriere eines einenden und doch problematischen Konzepts. In C. Piallat (Ed.): *Der Wert der Digitalisierung: Gemeinwohl in der digitalen Welt* (pp. 319–340). transcript. https://doi.org/10.14361/9783839456590-014

Politico (2020). *Boris Johnson heralds 'recaptured sovereignty' after Brexit.* https://www.politico.eu/article/boris-johnson-heralds-recaptured-sovereignty-after-brexit

Rostalski, F. (2024). *Die vulnerable Gesellschaft: Die neue Verletzlichkeit als Herausforderung der Freiheit.* C.H.Beck.

Schmitt, C. (1996 [1922]). *Politische Theologie: Vier Kapitel zur Lehre von der Souveränität.* Duncker & Humblot.

Timmers, P. (2019). Ethics of AI and Cybersecurity When Sovereignty is at Stake. *Minds and Machines*, 29, 635–645 https://doi.org/10.1007/s11023-019-09508-4

Traub, J. (2009). Absolute Fiction: The Perversion of Sovereignty. *World Affairs,* 171(3), 73–83.

Tretter, M. (2022). "Digitale Souveränität" als Kontrolle: Ihre zentralen Formen und ihr Verhältnis zueinander. In G. Glasze, E. Odzuck, & R. Staples (Eds.), *Was heißt digitale Souveränität? Diskurse, Praktiken und Voraussetzungen "individueller" und "staatlicher Souveränität" im digitalen Zeitalter* (pp. 89–125). transcript.

Voice of America (2023). *Putin Vows to Make Russia 'Self-Sufficient' in Fifth Term*. https://www.voanews.com/a/putin-vows-to-make-russia-self-sufficient-in-fifth-term/7402117.html

German Digital Sovereignty

Success Factors in a National Defense Scenario from the Standpoint of IT Consultancy

Martha Klare[1] and Ulrike Lechner[2]

Abstract: We show success factors of Digital Sovereignty from the perspective of IT consultancy. To do this, we compare success factors in the case of peace with the case of national defense in Germany. Our findings show that the following factors are most important in the case of national defense: (1) developing a strategy for Digital Sovereignty, (2) considering IT security, and (3) relying on European digital skills. Moreover, the following factors gain significant weight comparing to the peace scenario: adapting the IT infrastructure in terms of security and resilience, development of IT systems and data sovereignty in companies. We have transferred the findings into a model.

Keywords: Digital Sovereignty, German Sovereignty, Defense Scenario

1 Introduction

The current scientific discourse shows that the term 'digital sovereignty' is subject to a certain degree of vagueness (Glasze et al., 2023; Ruohonen, 2021). While Moerel & Timmers (2021) rather describe the dependency on suppliers as a problem, Avila Pinto (2018) points out that not only does digital dependency have to be reduced to strengthen digital sovereignty, but in fact public policies need to be adapted in order to promote the digital capabilities of the European Union. Thus, the question arises of which factors are important to focus on in order to strengthen digital sovereignty.

This paper offers a contribution to the factors that are important in the context of digital sovereignty from the perspective of IT consulting. We have noticed that IT consultants can provide an essential perspective, as they combine economic and technical contexts with real-life experience. By doing this, we differentiate between two scenarios: 1) peace and 2) national defense. We address the following research questions:

- RQ1: What factors are important in gaining more digital sovereignty in preparation of a defense scenario in Germany from the perspective of IT consultancy?

- RQ2: How do these factors change when compared with a peace scenario?

In a previous work, we have already provided a model with weighted factors for digital sovereignty from the perspective of IT consultancy (Klare & Lechner, 2023). We try to

1 University of the Bundeswehr Munich, Neubiberg, martha.klare@unibw.de
2 University of the Bundeswehr Munich, Neubiberg, ulrike.lechner@unibw.de

answer RQ1 with the help of a series of expert interviews. For RQ2, we take the model by Klare & Lechner (2023) as the object of research and compare the results from RQ1 with the given model by Klare & Lechner (2023). The aim of this work is therefore to obtain a model for the case of national defense in Germany.

This paper starts with an introduction (see chapter 1), continues with the research design (see chapter 2), and thirdly gives theoretical background information such as important definitions in chapter 3. Chapter 4 presents the results, and chapter 5 provides an outlook for the future as well as a conclusion of the paper as a whole.

2 Research Design

The methodology used here is based on two case studies (see figure 1).

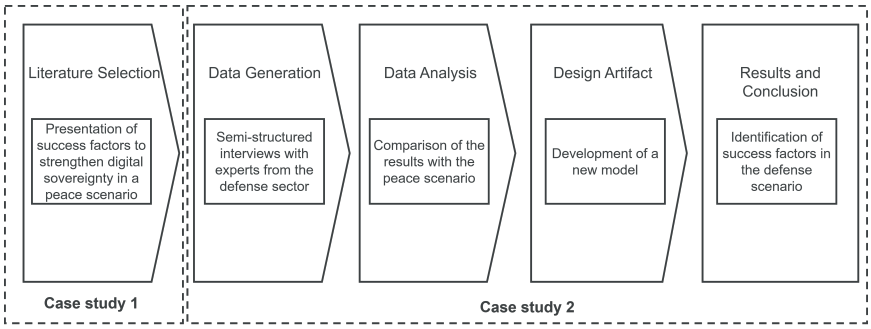


Fig. 1: Overview of research steps

The first case study focuses on a peace scenario, and the second case study on a defense scenario in Germany. Case study 1 provides a model with factors for digital sovereignty from the perspective of IT consultancy. To further develop the model in case study 2, eight semi-structured interviews with specialists in the field of defense were held. This study uses qualitative data fundamentally. Based on a review of the literature and the results from the previous model of Klare & Lechner (2023), the interview guide was divided into three types of open questions: 1) What does such a scenario mean for the digital sovereignty of companies? 2) How would you weight the factors from the peace case in the defense case? 3) Are there factors that become obsolete compared to the model in the peacetime situation? The interviews lasted between 0:49h and 1:28h. The data was collected between October and December 2023. The data was analyzed by comparing the defense scenario with the peace scenario. Afterwards, we designed a new model in which critical success factors were highlighted.

3 Theoretical Background

3.1 Digital Sovereignty: Understanding from the Standpoint of Politics

In an opening speech at the Internet Governance Forum 2019, the German chancellor described how Germany is dependent on software products from the USA (Merkel, 2019). In terms of hardware, there is a dependence on China and Taiwan (Weber et al., 2018). It became clear that more independence in terms of software and hardware was a topic driving the initial debate around digital sovereignty (Lambach & Oppermann, 2023). The beginning of the debate originated in politics and asked how Europe could digitally isolate isolate from other major powers (Pohle & Thiel, 2020). This view has changed over the years. The European perspective on digital sovereignty is nowadays much more about creating alternatives, regulations, and space for more self-determination (Pohle et al., 2022). At the same time, cooperation with other major powers and companies should not be ruled out (Broeders et al., 2023). Researchers distinguish between three perspectives when investigating the topic of digital sovereignty: 1) the state, 2) the economy, and 3) the individual (Pohle & Thiel, 2020). In this work, we adopt the perspective of IT consulting, and thus one that is subordinate to the economic perspective.

3.2 Digital Sovereignty: Understanding from the Standpoint of IT Consultancy

The DSMIC model in figure 2 shows key factors of digital sovereignty from the perspective of IT consulting (Klare & Lechner, 2023). It offers a vocabulary that companies can use when speaking about digital sovereignty. The factors are assigned to three different clusters: 1) internal company topics, 2) external company topics, and 3) cross-cutting topics. The factors in each cluster are arranged on the respective axis according to their weighting. A weighting of 4 stands for a ‘very important’ topic, 3 stands for ‘somewhat important’ and 2 for ‘somewhat unimportant’.

The most important factor according to the model is IT security. To strengthen digital sovereignty, companies should address the question of the extent to which they already consider IT security measures (such as crime protection). The cluster of strategy asks about the extent to which companies have already created a strategy regarding digital sovereignty. In the cluster of ‘fields of action’ the question then arises of the extent to which companies have already written down concrete actions for digital sovereignty. Regarding IT-infrastructure, companies should take a sovereign, secure, and self-determined IT-infrastructure (cloud or data center) into account. Behind the cluster of ‘IT systems’, the authors discuss a sourcing, development and customizing strategy for applications. On the one hand, standardization means that companies are caring for standard-based, compatible technology and applications inside their IT systems. On the other hand, interoperability means that companies have already considered the possibility of communication between two or more IT systems. When IT systems are discussed in more detail, the question of usability arises, i.e., how easy the

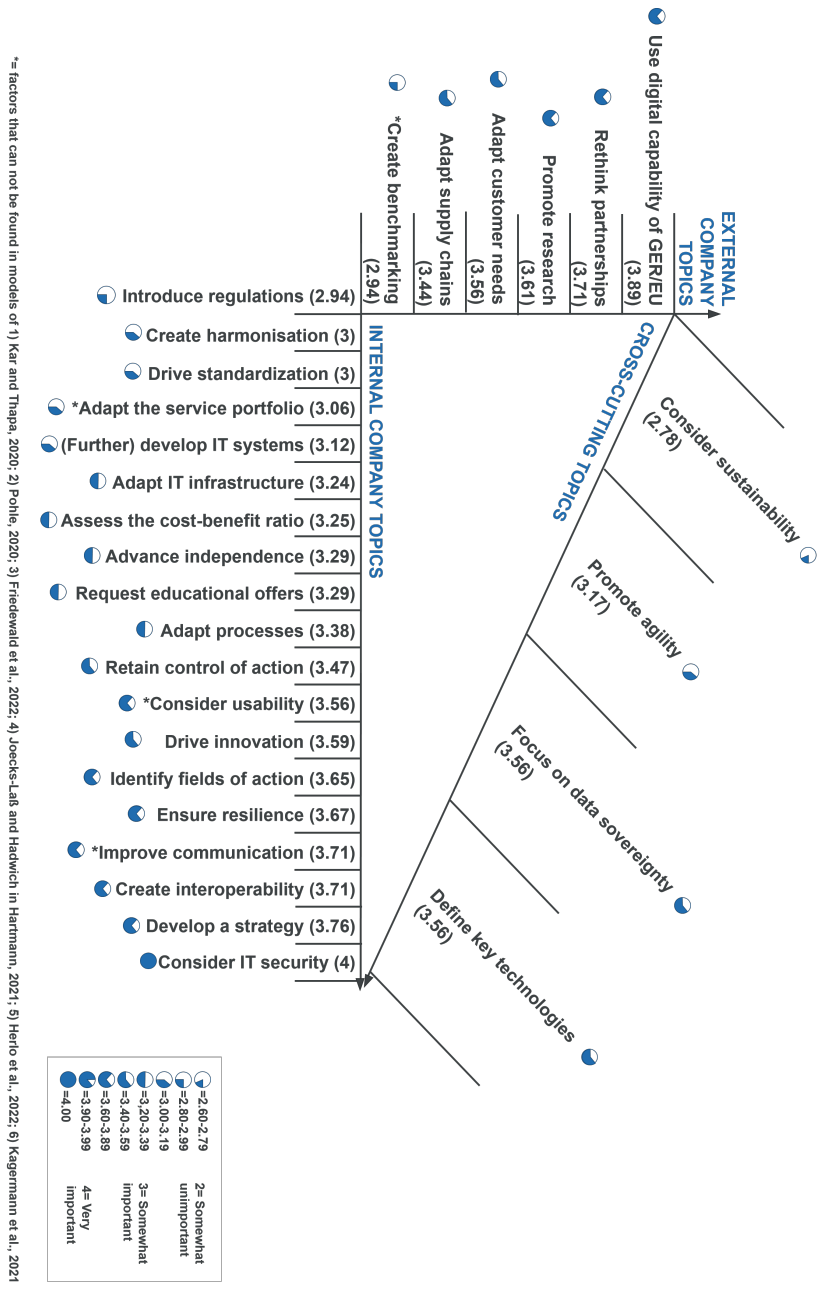


Fig. 2: DSMIC Model (Digital Sovereignty Model from the perspective of IT consultancy in a peace scenario (Klare and Lechner, 2023)

application is to use. Communication regarding digital sovereignty within a company and education in terms of trainings are further factors. Another factor considered to be important in the model is the cost-benefit ratio. According to Klare & Lechner (2023), companies should calculate the cost-benefit ratio as part of a cost-benefit analysis. The following questions are assigned to the factors of 1. process, 2. harmonization, and 3. regulation:

1. To what extent have you already identified procedural adjustments through changes regarding digital sovereignty?

2. How far have you already brought processes and applications into line?

3. To what extent have you already introduced company rules to deal with digital sovereignty?

Moreover, companies should check the extent to which they are already striving to achieve more independence (reducing dependence on external partners) and how far successful they have already been in gaining control over action inside their company. Measures for more resilience can provide success, i.e., the ability of a company to recover quickly from disruptions or unplanned changes. As for innovation, companies should ask themselves whether if they have already taken innovations into account when trying to strengthen digital sovereignty. Lastly, it might be helpful to create a complete list of the given applications (IT products and services) inside a company, because the measures regarding digital sovereignty may lead to a change in the service portfolio.

External factors in the model are the digital capabilities of Germany (GER) and the European Union (EU), partnerships, research, customer needs, supply chain, and benchmarking. Thus, in the 'external' cluster companies should review the extent to which they have already checked their digital possibilities based on German and European goods and services and the extent to which they have already questioned partnerships and cooperations. It also seems important to consider the influence of customers, supply chains, and suppliers regarding more digital independence. Research can help to penetrate the topic of digital sovereignty as well as benchmarking.

And finally, with respect to the cross-cutting factors, the following questions mainly explain the factor: Sustainability: To what extent have you already considered the ecological balance goals by the European Union for your company? Data sovereignty: To what extent have you already considered regulations for data that is stored or analysed by your company (f.e. GDPR)? Key technologies: To what extent have you already considered the key enabling technologies (KETs) of the European Union? And finally, agility: To what extent have you already strengthened your business agility, i.e. the ability of your company to respond quickly and positively to change?

The clusters marked with an asterisk are factors that are not reflected in the models of Friedewald et al. (2022); Hartmann (2021); Herlo et al. (2022); Kagermann et al. (2021); Kar & Thapa (2020), or "Digital sovereignty" (2021).

4 Case Studies

Case Study 1 analyzes success factors to strengthen digital sovereignty in a peace scenario. We are guided by Articles 1–9 from the 107th plenary session of Unesco when we talk about peace. This states that, among other things, the principles of territorial independence, sovereignty, and political respect must be observed in order to be able to speak of a culture of peace (Berliner Kommitee für UNESCO-Arbeit e.V., 2017).

In Case Study 2, we assume that Germany is preparing for a national defense scenario. In the scenario of national defense, Germany is attacked by another country. Article 115a GG describes what a scenario of defense includes. It states that an attack must be carried out with weapons and the approval of the Federal Council is needed. Subsequently, in accordance with Article 87a German Basic Law (Grundgesetz), military forces are mobilized for defense (Grundgesetz, 2022).

In both case studies, the research framework is deliberately limited. We examine critical success factors for strengthening digital sovereignty.

5 Results

5.1 Factors for Digital Sovereignty from the Standpoint of IT Consultancy in a Defense Scenario

Table 1 gives an overview of the weightings in a peace scenario and the defense scenario, and highlights differences.

STRATEGY FOR DIGITAL SOVEREIGNTY As already described in the peace case, the development of a strategy for the topic of digital sovereignty is necessary before a national defense case occurs. The IT consultants consider it very important to develop a plan in the form of a strategy for such a crisis situation (M=4). The weighting thus increases in the crisis situation compared to the peace situation (delta=0.24).

GUARANTEEING THE HIGH PRIORITY OF IT SECURITY IT security is particularly important in the context of digital sovereignty. Companies are increasingly concerned with the question of how they can move securely in cyberspace, but also how they can procure secure hardware and software or embed security-conscious partners in their digital ecosystem in crisis situations. IT consultants rate IT security as very important (M=4). Compared to the peacetime situation, the weighting does not changed (delta=0).

DRAWING ON EUROPEAN DIGITAL SKILLS Another facet that IT consultants consider to be important is European skills. Where possible, companies should draw

Factors with regard to digital sovereignty	M1 (mean value peace scenario)	M2 (mean value defense scenario)	Δ (delta)	σ (standard deviation)	Q1 (assesment M1)	Q2 (assesment M2)
Develop a strategy	3.76	4	+0.24	0	Somewhat important	Very important
Consider IT security	4	4	0	0	Very important	Very important
Use digital capability of GER/EU	3.89	3.88	-0.01	0.35	Somewhat important	Somewhat important
Adapt supply chains	3.71	3.75	+0.04	0.46	Somewhat important	Somewhat important
Consider sustainability	3.44	3.75	+0.31	1.28	Somewhat important	Somewhat important
Adapt IT infrastructure	3.24	3.75	+0.51	0.71	Somewhat important	Somewhat important
Create harmonisation	3	2.63	-0.37	0.91	Somewhat important	Somewhat unimportant
Ensure resilience	3.56	3.63	+0.07	0.35	Somewhat important	Somewhat important
Promote research	3.59	3.63	+0.04	0.99	Somewhat important	Somewhat important
Adapt customer needs	3.56	3.63	+0.07	0.52	Somewhat important	Somewhat important
Retain control of action	3.71	3.57	-0.14	0.53	Somewhat important	Somewhat important
Create interoperability	3.71	3.57	-0.14	0.78	Somewhat important	Somewhat important
Advance independence	3.29	3.5	+0.21	0.53	Somewhat important	Somewhat important
(Further) develop IT systems	3.12	3.5	+0.38	0.53	Somewhat important	Somewhat important
Driving innovation	3.47	3.5	+0.03	0.52	Somewhat important	Somewhat important
Improve communication	3.71	3.38	-0.33	0.74	Somewhat important	Somewhat important
Adapt the service portfolio	3.06	3.38	+0.32	0.52	Somewhat important	Somewhat important
Identify fields of action	3.38	3.38	0	1.60	Somewhat important	Somewhat important
Promote agility	3.17	3.38	+0.21	0.92	Somewhat important	Somewhat important
Focus on data sovereignty	3.67	3.28	-0.39	0.35	Somewhat important	Somewhat important
Consider usability	3.56	3.25	-0.31	0.89	Somewhat important	Somewhat important
Drive standardization	3	3.13	+0.13	0.64	Somewhat important	Very important
Request educational offers	3.29	3.13	-0.16	0.83	Somewhat important	Very important
Assess the cost-benefit ratio	3.25	2	-1.25	1.10	Somewhat important	Somewhat important
Adapt processes	3.38	2.88	-1.10	0.64	Somewhat important	Somewhat unimportant
Introduce regulations	2.94	2.43	-0.49	1.27	Somewhat unimportant	Somewhat unimportant
Create benchmarking	2.94	2.13	-0.81	1.46	Somewhat unimportant	Somewhat unimportant
Advance independence	3.29	3.5	-0.21	0.53	Somewhat important	Somewhat important
Define key technologies	2.78	1.25	-1.53	0.46	Somewhat unimportant	Very unimportant

Tab. 1: Overview of the weightings in the peace scenario, defense scenario, and their differences

on European skills. From the perspective of IT consultants, it is important (M=3.88) to investigate the European opportunities. The weighting of this facet barely changes in comparison to the peace case (delta=-0.01).

ENSURING DATA SOVEREIGNTY Data sovereignty is even more important in this scenario, although already classified as important. Only with the help of access to data can a decision be derived, and thus the actions of the actor be well supported (M=3.88, delta=0.32).

ENSURE RESILIENCE Resilience is slightly more important in the defense scenario than before (M=3.88, delta=0.21). 37.5% of respondents understand resilience part of the cluster of IT infrastructure. 25% of respondents assigned resilience to the IT systems cluster. These respondents therefore understand resilient IT systems to be systems that still function to a certain degree in the case of partial failures.

ENSURING A SECURE, RESILIENT IT INFRASTRUCTURE The IT infrastructure used by the company must be secure and resilient to attacks. This is even more important in the national defense scenario (delta=0.51). IT consultants consider it important (M=3.75) to reduce the number of failures as much as possible, to have fast response times with crisis plans, and to protect from unauthorized hackers.

ENTERING INTO SECURE, STRATEGIC PARTNERSHIPS IT consultants consider secure, strategic partnerships to be just as important as a secure IT infrastructure (M=3.75). The weighting compared to the peace case remains almost the same (delta=0.04).

MANAGING SUPPLY CHAINS (SOFTWARE AS WELL AS HARDWARE) Supply chains will most likely shift in the defense scenario. Different hardware and software components will be used in that scenario compared to peacetime. This is reflected in the assessment of the IT consultants, who rate the management of software and hardware supply chains as important (M=3.75). Previously, in peacetime, the topic was rated as less important (delta=0.31).

PROMOTING INNOVATION The topic of innovation (M=3.63) is less important, but still mentioned as very important in the context of digital sovereignty. IT consultants continue to see the promotion of innovation as a possible driver for achieving digital sovereignty in companies. In the case of defense, IT consultants rate this topic as important (delta=0.04).

TAKING CUSTOMER NEEDS INTO ACCOUNT Taking customer needs into account is just as important as the aspect of innovation. If we assume that companies will be forced to perform in terms of defense, the IT consultants rate this topic as very important (M=3.63). Companies' customer bases will most likely change in a defense scenario. The weighting of the IT consultants changes little on average (delta=0.07).

ENSURING INTEROPERABILITY BETWEEN IT SYSTEMS In terms of digital sovereignty, ensuring interoperable systems is desirable, but is not a high-priority topic. In the defense scenario, the compatibility of systems is slightly less important (M=3.57, delta=-0.14). The IT consultants emphasize that the functionality of military systems, for example, is more important.

ENSURING INDEPENDENCE IN THE CHOICE OF IT PROVIDER As already described in the peace scenario, companies need to ensure greater digital independence from certain IT providers. 87.5% of respondents state that this cluster is more of a superordinate cluster. It is therefore classified as inapplicable for the overall overview. On average, the consultants rate the guarantee of more independence from IT providers as important (M=3.5, delta=0.21).

CONTROL OVER ACTIONS IN THE CHOICE AND DESIGN OF IT It is important to regain control of action in order to enable secure and self-determined action in the digital space for companies. In both peacetime and crisis situations, preparatory measures must be taken (delta=0.03). In both scenarios, respondents consider it very important to gain control over digital actions in order to achieve digital sovereignty (M=3.5).

(FURTHER) DEVELOPING IT SYSTEMS In the national defense scenario, there will most likely be a requirement for other IT systems or, alternatively, further development of existing IT systems. The IT consultants meet this requirement by giving a higher weighting this factor (delta=0.38). In order to be able to act with sovereignty in a national defense scenario, it is still important for IT consultants to examine the topic of developing IT systems (M=3.5).

IDENTIFY FIELDS OF ACTION The results from the interviews show that the standard deviation is particularly high because it is difficult to classify this factor (sigma=1.60). 62.5% of respondents stated that the cluster 'identification of actions' could also be allocated to the cluster of 'strategy development'.

PROMOTING AGILITY Agility, associated with flexibility and speed when procuring IT, increases slightly in importance in the scenario, as flexible action is necessary to generate speed while complying with government regulations (M=3.38, delta=0.21).

IMPROVING COMMUNICATION ON THE TOPIC OF DIGITAL SOVEREIGNTY The IT consultants believe it is important to set limits that need to be implemented in order to drive digital sovereignty. This requires communication. However, this type of communication must take place before the scenario occurs, so that freedom of decision can arise within the scenario. This cluster is estimated as less important than in the case of peace (M=3.38, delta=-0.33).

ADAPTING THE SERVICE PORTFOLIO BASED ON THE REQUIREMENTS IN THE SCENARIO It is clear that IT consultants consider it important that the service portfolio is adapted to military requirements. However, the significance of this factor for strengthening digital sovereignty in the scenario decreases in comparison to the peace case (M=3.38, delta=-0.32).

CONSIDER THE USABILITY OF THE SYSTEMS It is important in both situations, peace and defense, that IT systems are designed to be user-friendly (M=3.25). IT consultants rate the importance of user-friendliness as higher in the case of a crisis, as it can have a significant influence on the survival of soldiers (delta=-0.31).

DEFINE KEY TECHNOLOGIES When it comes to key technologies, it must be emphasized that these must be mastered in order to create meaningful IT products. According to the IT consultants, defining and mastering key technologies is useful for strengthening digital sovereignty in the scenario of peace, but in the scenario of a crisis it would be too late to deal with this until the scenario occurs (M=3.25). For this reason, they weight this factor as being lower on average (delta=-0.31).

CREATING EDUCATIONAL OPPORTUNITIES FOR DIGITAL SOVEREIGNTY The creation of educational opportunities for digital sovereignty plays a subordinate role in peacetime, but should nevertheless not be undermined. In the case of defense, the IT consultants stress that the requirements for digital sovereignty should be communicated before the scenario occurs. The topic therefore becomes less important (M=3.13, delta=-0.16).

SETTING STANDARDS FOR IT Standards for IT products, processes, and structures increase speed, which is essential in the case of defense. The IT consultants rate standards for IT systems important (M=3.13) and increase the weighting for standards compared to the peace scenario (delta=-0.13).

CREATING PROCESSES FOR COMPLIANCE WITH REGARD TO DIGITAL SOVEREIGNTY Processes remain important in the defense scenario in order to be able to function in a coordinated manner. However, their importance decrease significantly, since the processes are only seen as supporting factors in scenario by the IT consultants (M=2.88, delta=-0.50).

PROMOTING EXCHANGE WITH RESEARCH Exchange with research plays a less important role than before (M=2.88, delta=-0.73). Topics such as creating speed are now more important.

HARMONIZE IT SYSTEMS Harmonization leads to vulnerability, and therefore decreases in importance in the defense scenario (M=2.63, delta=-0.37).

SETTING UP INTERNAL COMPANY RULES FOR DIGITAL SOVEREIGNTY Regulations are slightly less important (delta=-0.51). They are needed in the scenario, but they are not the most decisive (M=2.43). Rules are also better defined in advance.

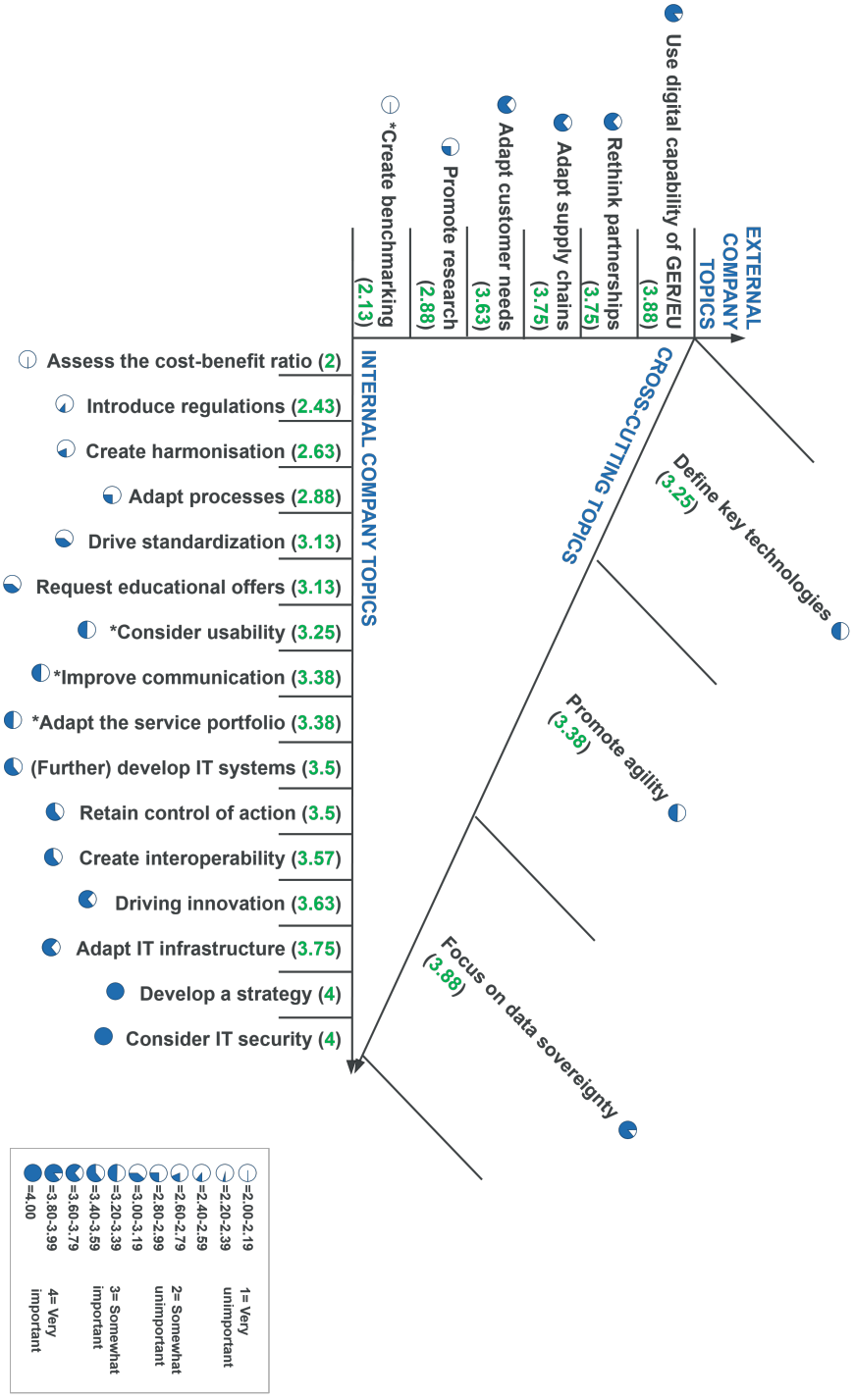
BENCHMARKING TO GATHER THE EXPERIENCE OF OTHERS The importance of benchmarking decreases in the defense scenario. The topic is only slightly important (M=2.13, delta=-0.81).

ANALYZE COST-BENEFIT RATIO The IT consultants state that costs are not important in the scenario (M=2). Ensuring security is much more important and is inevitably associated with costs. Thus, the cost-benefit ratio should not be completely ignored. The weighting here decreases sharply compared to the peace case (delta=-1.25).

TAKING SUSTAINABILITY INTO ACCOUNT Sustainability loses much of its importance in the defense scenario. All IT consultants state that it becomes highly irrelevant (M=1.25, delta=-1.53). The high standard deviation (sigma=1.28) should also be emphasized at this point. 75% of respondents see sustainability as somewhat or very unimportant in this scenario.

5.2 Model for Digital Sovereignty from the Standpoint of IT Consultancy and in a National Defense Scenario

In the following, we adapt the model as follows. First, we remove independence as a factor because the interview analysis showed that it is seen as an overarching goal. Then, we subordinate the cluster of ‘fields of action’ under ‘strategy’. A strategy can



*= factors that can not be found in models of 1) Kar and Thapa, 2020; 2) Pohle, 2020; 3) Friedewald et al., 2022; 4) Joecks-Laß and Hadwich in Hartmann, 2021; 5) Herlo et al., 2022; 6) Kagermann et al., 2021

Fig. 3: DSMIC Model in a national defense scenario

be developed by providing action fields, but it does not have to. The ‘strategy’ cluster can therefore be considered as a superordinate cluster. The factor of ‘sustainability’ is removed because its mean value is <2. Therefore it is rated as very unimportant in the scenario. And lastly, we subordinate ‘resilience’ under the clusters of ‘IT infrastructure’ and ‘IT systems’. According to our interview results, these are the main clusters in which resilience should be achieved if a company wants to strengthen digital sovereignty.

A new model was developed on the basis of the results (see Figure 3). It visualizes the weights collected from the expert interviews with the IT consultants and takes into account the adjustments described in advance.

6 Conclusion and Outlook

Strategy and IT security are top issues when it comes to digital sovereignty in a defense scenario. Adapting the IT infrastructure in terms of security and resilience, data sovereignty and development of IT systems are becoming increasingly important compared to a peace scenario. The issue of sovereignty over data can be the difference between life and death in a national defense scenario. Ensuring that IT systems function properly can give a head start. There is no room for the topics of sustainability and benchmarking in the defense scenario, although they are mentioned by authors in a peace scenario.

One limitation of this work is that only IT consultants were interviewed as experts. Another limitation is the framework in which digital sovereignty is viewed. The German perspective is assumed here, which can differ from the understanding regarding digital sovereignty of other countries.

A next research direction could be the consideration of a foreign policy crisis. In this context, researchers could examine how the results change compare to the peace and national defense scenarios. Transfer of the case study results into a reference model is planned as a next step.

Acknowledgments

This work originated in the LIONS research project. LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

References

Avila Pinto, R. (2018). Digital sovereignty or digital colonialsim. *SUR-Int'l J. on Hum Rts.*, *15*, 1–15.

Berliner Kommitee für UNESCO-Arbeit e.V. (2017). *Kultur des Friedens. Ein Beitrag zum Bildungsauftrag der UNESCO: Building Peace in the Minds of Men and Women.* Retrieved from `https://www.unesco.de/sites/default/files/2018-09/kultur_des_friedens.pdf`

Broeders, D., Cristiano, F., & Kaminska, M. (2023). In Search of Digital Sovereignty and Strategic Autonomy: Normative Power Europe to the Test of Its Geopolitical Ambitions. *JCMS: Journal of Common Market Studies*, *61*(5), 1261–1280.

Digital sovereignty. (2021). In (pp. 47–67). transcript Verlag Bielefeld.

Friedewald, M., Kreutzer, M., & Hansen, M. (2022). *Selbstbestimmung, Privatheit und Datenschutz: Gestaltungsoptionen für einen europäischen Weg*. Springer Nature.

Glasze, G., Cattaruzza, A., Douzet, F., Dammann, F., Bertran, M.-G., Bômont, C., ... Géry, A. (2023). Contested spatialities of digital sovereignty. *Geopolitics*, *28*(2), 919–958.

Grundgesetz. (2022). Retrieved from `https://www.gesetze-im-internet.de/gg/index.html#BJNR000010949BJNE011301377`

Hartmann, E. A. (2021). Digitale Souveränität in der Wirtschaft–Gegenstandsbereiche, Konzepte und Merkmale. In *Digitalisierung souverän gestalten: Innovative Impulse im Maschinenbau* (pp. 1–16).

Herlo, B., Ullrich, A., & Vladova, G. (2022). Verantwortungsvolle demokratisch nachhaltige digitale Souveränität. *Digitalisierung nachhaltig und souverän gestalten*, 1–66.

Kagermann, H., Streibich, K.-H., & Suder, K. (2021). Digitale Souveränität – Status Quo und Handlungsfelder. *acatech IMPULS*, 1–29.

Kar, R. M., & Thapa, B. E. (2020). Digitale Souveränität als strategische Autonomie. *Berlin: Kompetenzzentrum Öffentliche IT/Fraunhofer FOKUS*, 5–27.

Klare, M., & Lechner, U. (2023). Digital sovereignty from the perspective of IT consultancy in Germany: A model. *ICSOB Companion*, 1–14.

Lambach, D., & Oppermann, K. (2023). Narratives of digital sovereignty in German political discourse. *Governance*, *36*(3), 693–709.

Merkel, A. (2019). Rede von Bundeskanzlerin Angela Merkel zur Eröffnung des 14. Internet Governance Forums in Berlin. *Last modified November*, *26, 2019*. Retrieved from `https://www.bundeskanzler.de/bk-de/aktuelles/rede-von-bundeskanzlerin-angela-merkel-zur-eroeffnung-des-14-internet-governance-forums-26-november-2019-in-berlin-1698264`.

Moerel, L., & Timmers, P. (2021). Reflections on digital sovereignty. *EU Cyber Direct, Research in Focus*, 1–33.

Pohle, J., & Thiel, T. (2020). Digital sovereignty. *Internet Policy Review*, *9*(4), 1–19.

Pohle, J., Thüer, L., Dammann, F., & Winkler, J. (2022). Das Subjekt im politischen Diskurs zu „digitaler Souveränität". In *Handbuch Digitalisierung und politische Beteiligung* (pp. 1–23). Springer.

Ruohonen, J. (2021). The Treachery of Images in the Digital Sovereignty Debate. *Minds and Machines*, *31*(3), 439–456.

Weber, A., Reith, S., Kasper, M., Kuhlmann, D., Seifert, J.-P., & Krauß, C. (2018). Sovereignty in information technology. *Security, safety and fair market access by openness and control of the supply chain. Karlsruhe: KIT-ITAS. Online verfügbar unter http://www. itas. kit. edu/pub*, *2018*, 1–61.

Ethical Guidelines for DLT-based Information Systems

Isabelle Fries[1]

Abstract: There is a great desire for ethical orientation in the course of technological progress. This applies in particular to technologies that are considered new and disruptive. The demand for blockchain ethics is correspondingly high. The following research contribution not only provides guidelines, but also embeds the desire for orientation in the digital sphere within fundamental philosophical and ethical reflections, while simultaneously acknowledging legal implications as well as practical applications. It engages in an innovative design by examining existing forays into blockchain ethics in the scholarly discourse, but takes its own pioneering path on the basis of transdisciplinary exchange in the LIONS research project. At the core of the following research paper lies the specially developed "Modell der Wert-Dimensionen" (Model of Value Dimensions), encompassing the dimensions of common good, sustainability, autonomy, security, participation, transparency, and reliability. The handling of these dimensions is conceptualized as the action of ethically sovereign subjects and is addressed to IT governance as well. With regard to DLT-based information systems, concrete possibilities for conscious ethical development and design are presented. The aim is to raise awareness of ethically relevant issues and, ultimately, to promote a digital sovereignty that includes such awareness.

Keywords: Blockchain Ethics, Digital Sovereignty, Awareness, Ethical IT Governance, Value Sensitivity

PART I Challenges and Goals of "Ethical Guidelines" in the Context of Technological Innovation

In academic discourse, "ethical guidelines" for the implementation of technological innovation seem to be met with ambivalence of reservations and expectations. To outline the challenges and goals of the present "Ethical Guidelines for DLT-based Information Systems," both positions are presented in the first of two parts. In this PART I, general ethical considerations are outlined, including their concretizations regarding new technologies and a focus on information systems based on distributed ledger technology (DLT). Existing initiatives of blockchain ethics within research are systematically presented. In discursive engagement with existing approaches, the goal of contributing to ethical sovereignty in the digital sphere is justified. On this foundation, the "Wert-Dimensionen" (value dimensions) proposed for consideration from an ethical perspective are built upon. A corresponding "Modell der Wert-Dimensionen" (Model of Value Dimensions) is presented and explained in PART II.

[1] University of the Bundeswehr Munich, Neubiberg, isabelle.fries@unibw.de

I.1 Reservations and Expectations regarding "Ethical Guidelines"

I.1.1 Possible Reservations Regarding "Ethical Guidelines" from Philosophical and Theological Perspectives

Some representatives of philosophy and theology may be among those who have reservations about "ethical guidelines." The provision of a "guide" might suggest the existence of "a more or less established knowledge that is communicable and applicable" (Hubig, 1995, p. 5, translated) for a specific area of ethics; a knowledge that could presuppose consensus, which is not indisputably discernible in ethical matters. Ethicists can further substantiate their position by referring to freedom of judgment and autonomy of action. This leads to an implied pedagogical reservation, as the provision of guidelines, now understood as normatively binding and relieving of personal responsibility, replaces the independent understanding of a specific situation as an ethically relevant one, followed by autonomous ethical judgment and the taking of concrete actions. This argumentation presupposes both the emancipation of the individual in the course of the Reformation and the Enlightenment's conception of humanity. Here lie the foundations of an ethics based on freedom, rationality, and autonomy of the individual, with a connection between freedom, rationality, and autonomy subsequently present in Immanuel Kant's notion that freedom for rational action is understood in the sense of morally good conduct. The underlying conception of humanity is also followed when, since the twentieth century, the concept of "responsibility" has held a central position in many ethical designs and statements.[2]

No Step-by-Step Instruction to Moral Goodness

In light of such understanding, ethical guidelines may be suspected of being misunderstood as a reflexively followed step-by-step instruction. To illustrate a difference: in a numbered guide – for example, for assembling a cabinet – nothing should go wrong as long as one adheres accurately to the instructions. If something goes wrong despite following the instructions, one would hold the furniture manufacturer responsible and liable for any potential damage. If moral goodness were comparable to the example of the cabinet, one might occasionally compare instructions for achievement and refer to a general judgmental authority. However, the history of ethics shows that potential paths to moral goodness are more complex than cabinet assembly.

Ethics Demands and Fosters, but Does Not Replace Moral Decisions

In addition to the diversity of possible paths given in ethical approaches, an immanent reservation arises, whereby moral goodness is considered only approximately achievable (e.g., the greatest possible benefit or the least possible harm) or as a relative good

[2] In the context of technological innovations, particular attention should be given to ethical approaches that have called for responsibility in connection with technology assessment in the 20th century. In addition to Hans Jonas's well-known work "Das Prinzip Verantwortung" (The Imperative of Responsibility) from 1979, one should also consider Günter Ropohl, who has developed an engineering ethics based on the individual responsibility of developers through various contributions since 1979. Both approaches will be mentioned again below.

(e.g., as the best choice compared to alternatives). Moreover, the responsibility of an ethically capable subject cannot be completely outsourced to professional ethicists. While they can support with the knowledge and experience of their profession and offer recommendations, they cannot assume a proxy function regarding the individual conscience. Such an expectation would dangerously approach an understanding that could be encountered in Adolf Eichmann's case. The Nazi bureaucrat Eichmann not only referred to Kant's categorical imperative during his trial in Jerusalem, but also to its supposed application for ordinary people (see Wildt, 2013, p. 151f): "Act in such a way that the Fuehrer, if he knew of your actions, would approve them" (Frank, 1942, p. 15f, translated). What it means when the thoughts of an ethicist boil down to a demand to be uncritically followed was vividly illustrated by the philosopher Hannah Arendt with reference to this extreme historical example (see Arendt, 1963).

The foregoing allows conclusions to be drawn regarding ethical guidelines. It also helps to understand reservations. When technological innovations are in focus, this seems to be particularly pronounced. It is not without reason that most car users are familiar with the warning given by navigation systems that directional recommendations do not replace independent thinking, in this case not even in a moral sense.

I.1.2 Possible Expectations of "Ethical Guidelines" on the part of Applied Technological Research

The desire for guidance is also understandable, and the term "guide" encapsulates this desire. It suggests a guiding thread that one can follow, providing orientation and offering direction even in complex situations. In the context of applied technological research, "guiding" refers to the level of governance. Those intending to guide ethically can do so in conjunction with appropriate governance. Therefore, in the context of information systems, ethics should also consider IT governance.[3]

Increased Need and Expectation for Ethical Guidance

Expectations for ethical guidance are high. Gone are the days when ethical neutrality of technical artifacts or systems could be claimed in academic discourse (see e.g., Hare, 2022). The extent to which technology ethics is not just "ethics in the field of technology" and whether ethics is inherent in technology itself is now widely discussed. Referring to Artificial Intelligence (AI), the discussion not only considers how "artificial" intelligence based on algorithms programmed by humans truly is, but also how ethically relevant biases become visible in human-machine interaction.

Ethical Relevance of Programming and its Consequences

While DLT may not directly involve AI, a crucial phenomenon manifests in a similar manner. When human programmers consciously or unconsciously embed or omit ethically relevant parameters in the code, they can do so within the realm of blockchain

[3] This connection is also assumed by Lee et al., for example (Lee et al., 2022, pp. 1–3). The authors complain that although an awareness of the relevance of ethics and governance is essential for working as an IT professional, most people are ill-prepared for this at the start of their careers.

just as they can in machine learning. Furthermore, a core concern of blockchain, for understandable security reasons, lies in making future alterations as difficult as possible, if not outright excluding them. This entails consciously inflexible information systems. It also underscores the need for an immensely early awareness of the ethical relevance of one's actions, an understanding that programming does not occur in a space of moral neutrality, and an urgency concerning related assessment of consequences, which is even more pressing than in relatively flexible socio-technical systems that are easier to adjust in operation.

The relatively young history of DLT has already demonstrated on a technical-functional level that the unexpected can indeed occur. While the knowledge existed that a 51% attack was functionally possible, it was considered unlikely – yet Ethereum fell victim to a 51% attack in 2019 (see Lee et al., 2022, p. 205). It falls to others to present the various design possibilities of DLT, evaluate them technically depending on the context of application, and rightly point out alternatives to the proof of work mechanism. In an ethical perspective, the focus is to clarify: If unexpected consequences can, and indeed do, occur when considered on a functional level, this applies equally to ethical consequences, especially considering that unlike the 51% attack, there is no early awareness of the possibility of a consequence. The late French philosopher Paul Virilio once stated: "When you invent the ship, you also invent the shipwreck [...]. Every technology carries its own negativity, which is invented at the same time as technical progress" (Virilio, 1999, p. 89). This statement by Virilio can also be read very literally today. The mountains of electronic waste resulting from technological progress have recently found expression in corresponding EU legislation (see Right to Repair, Generalkommission Kommunikation, 2024).

Progress as a Moral Duty

In principle, it is plausible to expect progress from technological research. Stagnation based on a techno-pessimistic attitude, as suspected in Virilio's background, is not an adequate option from this expectation. Inaction would not be a necessarily neutral position even from an ethical standpoint. Similar considerations are reflected in criminal law, such as the concept of "acts of omission" (see §13 StGB, §323c StGB, among others). In relevant contexts, progress can also be seen as a moral duty. In the background of publicly enabled and financially supported research, there is a correspondingly positive expectation regarding technological progress.

Promotion of the Common Good as a Sociopolitical Goal

From an ethical perspective, there is an expectation that technological innovations can contribute to a societal or even global common good. Publicly funded development and operation of DLT-based information systems can be compared to the transportation infrastructure provided by Deutsche Bahn. The general infrastructural aspect of information technologies has already been emphasized by the Italian philosopher specializing in information ethics, Luciano Floridi (see Floridi, 2022, p. 31). The German Bundestag recently passed an infrastructural "orientation towards the common good" in terms of safeguarding societal interests for Deutsche Bahn (Deutscher Bundestag, Parlamentsnachrichten, 2023, translated). This expectation and political inclination

towards the railway can be implicitly transferred to or from publicly owned or operated information systems. They too can be seen as infrastructural means to enhance the common good.

By referring to the promotion of the common good, a direction is outlined that stands at least on the horizon of societal interest for every technological innovation if moral goodness is to be given shape. However, it is clear that this "shape," within an open underdetermination of the term "common good," remains a phantom in a flexible concept. Similarly, this applies to "justice," which, of course, encapsulates a clear objective of morally good actions, not only for ethicists.

Desire for a Bridge between Programming and Horizontal Values

Applied research is thus faced with the challenge of anticipating, and then implementing a path from a starting point to a particular distant goal. The starting point manifests in concrete research and requirements, sometimes even the interaction of various requirements and needs of different stakeholders, extending into the concrete code in the context of information technology. The distant goal can be described as the realization of socially desirable values and a target state desired from an ethical perspective. In the face of this challenge, "ethical guidelines" are expected to anticipate and delineate this path in its concrete steps, allowing developers and designers, for example, to use it as a guiding thread to follow. In short, the desire is for an instruction that spells out the bridge between programming and the lofty values on the horizon.

It is correct to address this desire to ethicists. But it is also correct that an ethics in dialogue with other disciplines is needed. It is equally correct that ultimately, responsible ethical decisions by individual persons matter, regardless of which discipline they belong to. It is as gratifying as it is necessary that the fine-grained differentiation into various disciplines, as observed in the scientific landscape particularly since the transition to the twentieth century, has receded into the background in the twenty-first century. To address societal questions, inter- to transdisciplinary knowledge is required again, even though this expanded knowledge, by now distributed across all areas, is distributed among various researchers. In such teams, they work together as in the LIONS research project.

I.2 Ethically Sovereign in the Digital Sphere: Components for Responsible Use

In this first part, the focus is on the challenges and possible goals of "ethical guidelines" with regard to technological innovation. General challenges in the tension of expectations have already been outlined. These considerations are also valid for information systems based on blockchain. However, what can be the goal of "Ethical Guidelines for DLT-based Information Systems" against this background? This is the subject of this section. To this end, existing research approaches to shaping the approach represented here are also taken into account.

I.2.1 Paths to Blockchain Ethics

The discrepancy between external attributions and the self-perception of ethicists has already been hinted at. While applied research in technological innovation expresses a desire for normative guidelines, Armin Grunwald, currently one of the most well-known German philosophers in the field of technology assessment, argues otherwise (Grunwald, 2013, p. 6, translated): "Society remains on its own in making decisions about the future and setting the course for scientific and technological progress." Ethics can only provide "conditional-normative advice," he continues, and ultimately makes it clear: "Technology ethics is by no means suitable as a kind of regulatory authority that can issue ethical safety certificates." This assessment is explicitly shared by the Catholic ethicist Anna Maria Riedl (Riedl, 2022, p. 288): Technology ethics does not have the task of issuing safety certificates. Riedl also rules out general norms or rules as top-down guidelines (Riedl, 2002, p. 289).

Emergent Technologies and the Search for the Known in the Unknown

But how should one productively deal with the "orientation deficits" (Grunwald, 2013, p. 2, translated) that inevitably accompany scientific and technological progress when advancing into unknown territory, where there are expanded possibilities for action but no related experiences and precedents? One possibility is to search for the known in the unknown. This not only provides researchers with a subjective sense of certainty in the face of uncertainty, but also enjoys great popularity as an objective scientific method. A chapter on literature review at the beginning of a scientific paper serves this purpose, among others. If a technology is seen as an emergent technology – and blockchain has been widely referred to as an emergent (Kirchschläger, 2021, p. 241; Sharif & Ghodoosi, 2022, p. 1009) or disruptive (Kučera & Bruckner, 2019, p. 129; Agerskov et al., 2023, p. 1; also see Feloutzis & Lekakos, 2023, for discussion) technology – pioneering efforts seem to be required. Since some years have passed since a white paper underlying DLT in general was published under the pseudonym of Satoshi Nakamoto (Nakamoto, 2008), sporadic attempts at blockchain ethics have been made, but it is still in its infancy overall (see Agerskov et al., 2022, p. 2; Sharif & Ghodoosi, 2022, p. 1011). However, it should be noted that the founding paper itself can indeed be read in the context of a moral claim. After all, the anonymous individual or group named Satoshi Nakamoto seems to be striving to address a perceived loss of integrity and trustworthiness in the financial sector (see Fries, 2022). The goal here is to achieve reliable interaction between individuals, which is nevertheless perceived as possible, verifiable, and secure only through technical mediation.

Four Discipline-Specific Starting Points of "Blockchain Ethics"

Looking at the research literature of recent years, four starting points can be schematically identified, from which intersections of blockchain and ethics are discussed:

1. **Computer Science Researchers:** Researchers in the field of computer science come across ethical issues inherent in their profession, e.g., by evaluating transparency in statements, which is often cited in the blockchain context, as a means to enhance moral integrity (e.g., Khan & Equbal, 2023).

2. **Economics-Oriented Researchers / Business Information Systems and Management Sciences:** These researchers predominantly engage with blockchain in the context of organizations, as well as the public sector. There are already several publications from this area. Ethics becomes a topic in relevant literature when the implementation of blockchain in the business sector aims to contribute to EU-required compliance with human rights along the supply chain (see OECD, 2019). Legal questions also involve ethical considerations or convictions, although they may not always become explicit research topics. From a legal perspective, for example, Tsai & Lin explore the potential of DLT to ensure human rights along global supply chains (Tsai & Lin, 2023). The Federal Office for Migration and Refugees (BAMF), on the other hand, sought a blockchain solution for asylum application management. The result was the prototype of a private permissioned blockchain, taking into account the right to correction and deletion of personal data, ensured through an off-chain central database linked via API (Rieger et al., 2019). The prototype was followed by the current successor project FLORA (Federal Office of Migration and Refugees, 2024).

3. **Application-Oriented Researchers in Philosophy, Theology, and Law:** These scholars often position blockchain under broader topics such as technology or new technologies (e.g., Kirchschläger, 2021). Frequently, only the even broader and more unspecific term of "digitization" is encountered (e.g., Held & Oorschot, 2021). In this context, themes such as "informational freedom" (Deutscher Ethikrat, 2018) are discussed. Situationally, and often with exemplary character, blockchain is then referenced, sometimes in parallel with AI. So far, a DLT-specific legal (e.g., G'sell & Martin-Bariteau, 2022) or ethical (e.g., Marković, 2022; Ishmaev, 2021) classification has only been found in exceptional cases. The societal impacts of blockchain in an ethical context were first examined by Cara LaPointe and Lara Fishbane in their work for the Beeck Center for Social Impact and Innovation at Georgetown University. While the researchers do not have a specifically academic background in philosophy and ethics, they have expertise in technologies related to social issues and societal change. As a result, they presented "The Blockchain Ethical Design Framework" in 2018 (LaPointe & Fishbane, 2018; abridged version: LaPointe & Fishbane 2019).

4. **Interdisciplinary Teams:** Researchers in interdisciplinary teams like the LIONS research project seek legal compliance and ethical value in developing and designing specific applications. The aim is to create sustainable and robust systems for, or in connection with, the public sector that comply with the required frameworks, which include ethical aspects. A similar interdisciplinarity with corresponding interests exists at the European Blockchain Center (European Blockchain Center, 2024).

The "Ethical Guidelines for Blockchain Systems" of Copenhagen

One notable attempt at a Blockchain Ethics stands out due to its precise thematic focus. Signe Agerskov, Asger Balle Pedersen, and Roman Beck, the latter being the

director of the European Blockchain Center, presented their draft of "Ethical Guidelines for Blockchain Systems" at the European Conference on Information Systems (ECIS) 2023 (Agerskov et al., 2023). If we were to categorize this authors' group, which is smaller and more specialized than the entire expert group at the European Blockchain Center, it aligns most closely with starting point (2). The researchers from the IT University of Copenhagen have backgrounds in business information systems, software development, and management. It is surprising that no professional ethical expertise was directly incorporated into these "Ethical Guidelines for Blockchain Systems," which would have shifted the focus towards starting point (4). There is no critical examination of whether normative top-down directives represent a desirable moral endpoint. Instead, the "top-down approach" with "clearly defined norms," understandable from a management and technical development perspective, is highlighted and explicitly identified as the crucial content (Agerskov et al., 2023, p. 1).

Their methodology initially followed a classic approach: they examined scholarly blockchain literature to identify and cluster the ethical fields mentioned. The research group correctly recognized that public services utilizing DLT-based software solutions should embed and promote values and norms deemed socially desirable within the information system itself (Agerskov et al., 2023.). With respect to the European Blockchain Service Infrastructure (EBSI) initiated in 2018 (European Commission, 2024b), this implies that the technological infrastructure must align with an EU value base. Consequently, the Copenhagen researchers looked to EU documents as the instance for ethical-normative goals. From these documents, they inferred an EU value base, which they proposed to operationalize in DLT systems by incorporating the ethical fields identified through literature research.

Two Pitfalls on the Path to Ethical Guidelines

What initially appears straightforward and logical in this methodology reveals pitfalls on closer inspection. Primarily, these pitfalls arise because ethical reflections in these "Ethical Guidelines" are not directly derived from ethical reflection, but are broken in two ways. The first break lies in the literature review approach. The scholars referred to blockchain literature that does not necessarily reflect an ethical perspective. Specialized (overview) literature in ethics was not included. This approach results in the identified ethical fields appearing as a somewhat arbitrary collection, captured at an equally arbitrary point in time. The fact that certain questions are already classified as ethically relevant in the relatively young and sparse research literature does not imply that this catalog of topics is complete, nor does the frequency of mentions indicate the importance of a particular ethical concern, especially when it comes to weighing up interests in practice.[4] The second break results from the fact that EU documents were used in the manner of a normative legal framework, which, while legitimate, causes ethical concerns preceding binding laws to recede into the background. Thus, ethics is viewed through the lens of the rigidity of the law. Both breaks lead to the same

[4] Sifting through data material given in research papers, with criteria including the occurrence and frequency of certain keywords, but also being able to recognize blind spots in the form of gaps, noticing exciting constellations, reflecting on them, and weighing them up: those are intellectual skills that are less easily replaced by algorithms.

effect: they shift open and still unresolved ethical reflection, particularly regarding emergent technologies, towards a time- and situation-insensitive and unequivocal clarity. This clarity is useful because it is precisely this time- and situation-insensitive, as well as irrevocable, unambiguousness that is desired from the perspective of development and management when dealing with a inflexible rule-based information system such as a blockchain, which allows subsequent system adjustments to be made only with great difficulty. However, ongoing ethical reflection that considers technological progress, potential unintended consequences, long-term user experience, and changes in societal perception and evaluation is sacrificed on the altar of utility.

Deliberative or Technically Functional?

This clearly illustrates how a specific professional background influences the choice of questions, concerns, methods, and goals. It can also introduce specific biases, which can be more readily identified and mitigated through the multiperspectivity of different disciplines. A deliberative restraint, as seen in the work of technology philosopher Armin Grunwald or theological ethicist Anna Maria Riedl, contrasts with a purely functional understanding and approach. In this context, it is essential to note:

- Ethics is not limited to binary judgments, even though programming would be easier if it were.
- Although ethics operates with rules, it cannot be exhaustively represented in a rule-based system, at least not unless one holds a purely casuistic understanding of ethics that can be concluded casuistically at a certain point in time.
- Ethics is more complex than warehouse logistics, which good management can optimize.

The Narrative of Technological Guarantee for Morality

An as yet unmentioned problem, obscured by the purely technical-functional approach, is the unconscious and misleading premise that ethical norms integrated into a technical system will automatically translate into corresponding moral efficacy. Jurica Marković highlighted in his 2018 master's thesis that there are no technical guarantees for specific moral outcomes.[5] As an officer in the Croatian military and now a philosopher with a doctorate, Marković has extensively studied the ethical foundations and moral challenges of blockchain technology. In doing so, he has already laid the foundations[6] for the realization that goals justifying the use of blockchain, such as trust and integrity, cannot be *generated* by software itself (but by the people who develop it), nor can they be *guaranteed* by software (but by the people who use it). In a similar vein, Jan Kučera and Tomáš Bruckner assert that the potential of blockchain

[5] His master's thesis from 2018 (Zagreb) was entitled "Ethical Foundations and Moral Challenges of Blockchain Technology." A revised version was later published (Marković, 2020).

[6] See Marković, 2020: "A. Tapscott and D. Tapscott will say that these values [integrity and trust] are inscribed into the blockchain code, but that does not mean that people are amnestied from them. On the contrary, if we look at technology as a product of human labour and intellectual achievement whose goal is to facilitate and simplify specific tasks, then we can acknowledge that the referential point of technology implies a human being."

depends critically on its responsible development and use (Kučera & Bruckner, 2019, pp. 129–130). Among others, they reference Kranzberg's laws of technology, which originate from the technology historian Melvin Kranzberg (Kranzberg, 1986). According to the first of these six "laws," technology itself is neither good nor bad nor neutral, and according to the sixth, technology is primarily a human activity that gives a certain qualitative character to its actions.

Autoregulatory does not automatically mean autonomous

In this context, it becomes significant whether "autonomous agents" are mentioned, as the Copenhagen research group does in reference to smart contracts (Agerskov et al., 2023, p. 1), or if "automated" is the preferred term, as used by Monica M. Sharif and Farshad Ghodoosi in the realm of business ethics, also in reference to smart contracts (Sharif & Ghodoosi, 2022, p. 1010). The term "autoregulatory," advocated by the Protestant ethicist Nicole Kunkel,[7] is also worth considering. Kunkel specifically addresses the application of lethal autoregulatory weapon systems, highlighting the importance of distinguishing this from the concept of autonomy in technical systems. The use of terms such as "automated" or "autoregulatory" maintains the understanding that ethically responsible human actors are behind every rule-based system. Referring to ethically "good" possibilities enabled by blockchain – such as democratization and increased inclusion – Marković aptly summarizes (Marković, 2020, p. 435): "Although from the technical point of view, this is ensured by the technological components, there still exists the non-technical factor." Marković also touches on the concept of the common good, viewing the critical question as being how a particular system is used (Marković, 2020, p. 442): "for general welfare or the welfare of certain individuals?" He thus considers it indispensable to enable people ethically, drawing on a tradition of virtue ethics that dates back to antiquity.

I.2.2 Guidelines as a Contribution to Ethical Sovereignty in the Context of DLT-based Information Systems

At this point, all the prerequisites and explanations leading to the approach presented here have been laid out. The goal of this "Ethical Guidelines for DLT-based Information Systems" is to contribute to ethical sovereignty in the digital sphere.

Systemic Sovereignty Model of the LIONS Research Project

The concept of sovereignty frequently appears in the context of new technologies, though not always in a consistent manner. In the LIONS research project, a model was developed that addresses the systemic complexity of digital sovereignty (Fries et al., 2023). This model considers three levels both individually and in their interconnections: (1) the level of the state or a supranational institution, (2) the level of organ-

[7] Nicole Kunkel presented her ethical considerations on autoregulatory weapon systems as a dissertation in 2023 (not yet published). In this context, she also explores the concept of autonomy in the context of technologies from an ethical perspective. See in short version: Kunkel, 2021.

izations, and (3) the level of individuals. In this context, the focus is on "ethical sovereignty," emphasizing a specific aspect. The focus is on individuals as actors capable of acting with freedom, rationality, and autonomy (see Chapter I.1.1).

Action-Theoretical Prerequisites for Speaking of "Ethical Sovereignty"

In an ethical sense, action refers to intentional and goal-directed activities (Fenner, 2020, p. 35). When individuals engage in such activities, they become subjects of action. Their ability to reflect on alternatives to their actions and to justify a particular choice of action or the means employed is fundamental to discussing individual responsibility regarding a specific action (as in deontological ethics) or its consequences (as in consequentialist ethics). Additionally, there are systemically intertwined actions whose consequences cannot be easily or solely attributed to a specific individual. This is particularly true for socio-technical systems, a term rooted in the work of technology philosopher Günter Ropohl. His habilitation thesis "Eine Systemtheorie der Technik" (Ropohl, 1979) deals with the deindividualization processes in human-machine interaction (see explanation and critique in Gräb-Schmidt, 2002, pp. 36.95–98.124). Not coincidentally, Hans Jonas's much more renowned and influential book "The Imperative of Responsibility: In Search of an Ethics for the Technological Age" (first published in German: "Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation," Jonas, 1979) was published in the same year. Both authors examined the altered conditions of human action brought about by new technologies. This was a period when the expanded possibilities of human action through technology, particularly their potentially existential threats, were being discussed on a societal level. Against this background, Jonas elevated responsibility to a principle, asserting that actions must never be taken that could destroy future life foundations (Jonas, 1979, p. 36). His own future-oriented categorical imperative clearly references Kant's (see Hubig, 2015, p. 190). Jonas called for responsible action within the scope of new technologies, even where traditional action theory seemed to lack clearly identifiable subjects of action. Regarding such seemingly responsibility-vacant areas, he urged over-individual societal players – politics and institutions, industry, and the scientific community – to engage in responsible action with a view to the future. This did not replace classical, often retrospective responsibility attribution to individuals (see "Die Kollektivitätsnatur der neuen Handlung" in Viana, 2010, pp. 78–79), but represented an attempt to address a perceived accountability vacuum in systemic actions where, in the classical sense, no single person seems responsible, yet still shape such actions with responsibility and foresight.

Against an Abandonment of Responsibility

More than half a century later, the two poles of responsibility toward the future – individual action (or inaction) and over-individual action (or inaction) – still persist in discussions. If a societal goal is to avoid any areas devoid of responsibility, because a shared future should be shaped responsibly, it is not productive to demand responsibility solely from the opposing side of the same coin, as is currently observed, especially in the context of the climate crisis. To shape the future in a desirable manner, responsibility vacuums must be avoided even on an individual level, where the convenience and attractiveness of shedding responsibility remains a constant challenge,

as Kant[8] already recognized. Therefore, the knowledge of qualified experts is needed just as much as the elected representatives of the population who engage in political action. All stakeholders are needed, as well as individual actors who are aware of their responsibility as subjects of action.

Knowledge and Knowledge Dissemination as Prerequisites

To act intentionally and purposefully, knowledge is required. Although there is a *fundamental* knowledge deficit regarding comprehensive understanding of the consequences of new technologies, there are *individual* knowledge deficits that can indeed be addressed with varying degrees of effort, as philosopher Dagmar Fenner rightly distinguishes (Fenner, 2020, pp. 47–48). In the context of information technologies, not only technical and legal knowledge is needed by various stakeholders, but also ethical knowledge. The subjects of this knowledge are all those involved in the processes of creation, usage, and the consequences of creation and usage. There is a substantive necessity for ethicists to decisively distance themselves from the understandable expectations aiming toward a functionalizable and operationalizable ethics, especially if their goal is to avoid presumed certainties, conveniences, and resulting responsibility vacuums. Only through the awareness and empowerment of all stakeholders can responsible action replace merely suggested sovereignty.

Components for Responsible Use

Initially, two possible positions regarding ethical guidelines were introduced: reservations and expectations. Both positions are constructively incorporated into the current aim of contributing to ethical sovereignty in the digital sphere. The goal cannot be a step-by-step instruction like assembling a cabinet, yet concrete components for responsible use can be offered. The following "Modell der Wert-Dimensionen" (Model of Value Dimensions) was developed for this purpose. Just as awareness of IT security risks cannot be sufficiently replaced, nor can awareness of ethical implications for DLT-based information systems be substituted. From the perspective of admins, users may be perceived as security risks due to their uncontrollable actions. Admins can restrict user rights and increase awareness of IT security issues during use, but they cannot prohibit usage entirely without rendering the system useless. Similarly, the components provided by the "Modell der Wert-Dimensionen" can help increase awareness of ethical implications. They intentionally leave room for action, allowing for situational, independent, and intentional actions during specific applications while considering the provided components.

A significant security risk lies in the illusion of security, as it fosters carelessness. This applies to the technical side of an information system as well as its ethical component. Caring for the future, however, could be described as responsible action.

[8] See Kant, 1784, p. 35 (translated): "It is so convenient to be immature. If I have a book that has reason for me, a pastor who has conscience for me, a doctor who judges my diet for me, etc., I need not trouble myself."

PART II "Wert-Dimensionen" (Value Dimensions) in the Context of DLT-based Information Systems

In PART I, the foundations and considerations supporting the argument for ethical sovereignty in the digital sphere were discussed. Building on this, PART II proposes "Wert-Dimensionen" (value dimensions) in the context of DLT-based information systems. The selection of the seven dimensions is justified, among other considerations, through engagement with already scientifically discussed ethical fields in the context of DLT. Before the relevance of the "Modell der Wert-Dimensionen" (Model of Value Dimensions) for blockchain ethics is explained, some remarks on the linguistic and philosophical understanding of "Wert-Dimensionen" are warranted. In the course of this, it will also become clear why the German term was retained.

II.1 Space for Values

When "dimensions" are mentioned below, spatial imagination is used as a helpful construct. This aligns with the original understanding of the term, which is likely more familiar to natural science disciplines than to the humanities today. The focus is on the dimensions of length, width, and height, which form the three-dimensional coordinate system of space (see Apel, 1953, p. 57). The term "dimension," which exists identically in German and French, originates from the Latin "dimensio" and literally means "measurement." The English verb "measure" – like the equivalent German "messen" or French "mesurer" – likely shares the same linguistic root as the noun "dimension." When speaking of dimensions in everyday language, meaning the "extent" of something, the original meaning and etymological origin still resonate. The term "extent" in turn comes from the Latin "extendere" and refers to spatial expansion as well. This can also be an extension "in the direction of something" (see Hau et al., 2003, pp. 322.912), which is particularly interesting later on. However, an everyday use of "dimension" and "extent" also reveals a metaphorical transfer that makes it easy to overlook the primary linguistic reference to spatial extent. The following model does not intend to rely on such everyday understanding.

Spatial Thinking for Philosophical-Ethical Insights

Both the concept of spatial extension and the idea of measurement have been repeatedly used to illustrate ethical insights. For Kant, according to common understanding, one person's freedom finds its limits at the freedom of another – a notion that continues to shape ethics and legal understanding in Europe (e.g., Schapp, 2006). French philosopher Albert Camus, often associated with existentialism against his will (see Schönherr-Mann, 2018, p. 323), built on Kant and assigned central importance to the idea of excessiveness in the sense of knowing no measure. Camus defined boundless freedom as freedom without measure and, starting from the phenomenon of excessiveness, particularly turned away from totalitarian mechanisms (see the chapter "Maß und Maßlosigkeit" in Sändig, 2004, pp. 88–102). Even in everyday language, a moral connotation is found in the discourse on measure, such as when "the right measure"

is sought or when the behavior of another person is seen as "overstepping." The inability to "maintain measure" was considered a mortal sin in antiquity, and Aristotle recommended the measure of the mean as a virtuous balance between extremes of attitude and action in his Nicomachean Ethics (on Aristotle's doctrine of the mean, see Wolf, 2010) – an ancient way of thinking that is also evident in Camus' preference for protagonists from Greek mythology.

Without mentioning Camus, philosopher Annemarie Pieper also spoke of excessive and undesirable freedom (Pieper, 2014, p. 227). What distinguishes humans as humans, she argued, is the ability to set limits for themselves, thereby becoming aware of the value of their freedom. The idea of self-limitation was often invoked in the twentieth century in the context of technologically extended human possibilities (e.g., Illich, 1973; Meadows et al., 1972) and still appears in contemporary discourse, including an awareness of the limits to growth. The discourse on measure and boundaries transfers spatial, initially territorial, conditions to notions of moral rectitude, legal permissibility, or societal desirability. Before presenting the specially conceived "Modell der Wert-Dimensionen" based on the preceding considerations on spatial-philosophical thinking, it is also necessary to address the concept of "Wert" (value).

Worthful Ways

English words "price" and "value" can be both translated into German as "Wert." Marking a difference in favor of an ethical understanding was therefore important for the German thinking and speaking philosopher Kant. He distinguished between external and relative values in economic terms (price) and inner values within the horizon of ethics (dignity).[9] In ethics, the discourse on values refers to certain goods, such as the aforementioned freedom. Values enable the moral evaluation of an action, for instance, determining whether it contributes to the protection of freedom. For the argumentation here, the traditional line of value ethics is less relevant than metaethical considerations in the sense of reflections on the linguistic nuances when talking about morally relevant issues. The German term "Wert" is associated with the suffix "…wärts" (Pfeifer, 2012, p. 1559), indicating a direction. In English – linguistically analogous in French ("valeur") – one speaks instead of "values" (Smedley et al., 1845, p. 1013). Although this term comes from the Latin "valere," which refers to a strong ability, English also knows the adjective "worth" and the directional suffix "…wards," which are cognates of the German "Wert" and "…wärts." Directional indicators derive their meaning from the concept of space, thus linking back to spatial thinking as a construct for ethics: values serve as target coordinates that specify the course of actions (and their consequences) in space (and time). Adding the dimension of time brings into focus the long-term impact of actions and technological assessments. The value of freedom, which can manifest in the technological horizon as "informational

[9] See Kant, 1785, p. 434–435 (translated): "In the realm of ends everything has either a *price* or a *dignity*. Whatever has a price, can be replaced by something else as an *equivalent*; whatever, on the other hand, is elevated above all price, and therefore allows no equivalent, has a dignity."

freedom" (as mentioned in Chapter I.2.1), becomes a spatiotemporal target point. Actions should strive towards this point if the chosen direction is to be considered ethically "good" by those who advocate and defend the value of freedom.

Governance as Navigating Space and Time

The group of individuals who adhere to certain values can be institutionalized as a state or a supranational entity, like the EU. It can also be found in corporate governance that follows specific policies, sometimes extending beyond legally mandated requirements through self-commitment. In the context of both state and organization, the term "governance" has experienced a surge in academic reflection since the 1990s (Schuppert, 2008, p. 14). A final linguistic note on "governance" is useful, especially when arguing that blockchain ethics should be anchored near IT governance.

It should no longer be surprising that "governance" etymologically points to orientation. Although the term "govern" comes through French from a Latin source, the roots go deeper. The Latin "gubernare" is a loanword from Greek, where it is the same root as that found in "cybernetics": "kybernan,"[10] which means to steer or pilot a ship. Even in ancient times, the metaphorical meaning of guiding was known (Gemoll et al., 2006, p. 484; Hau, 2003, p. 388). This can be read in Homer (ca. 800 BC), and early Christian texts also know both literal and figurative meanings (Bauer & Aland, 1988, p. 927). Governance describes the course in space and time set by those who "steer the ship," i.e., those who lead governmental or corporate affairs, thereby determining the direction and goal of actions. These directional decisions are binding for all "passengers" on the ship and affect the surrounding environment as well. Course changes are possible, but must consider the future and not occur in an empty social and ecological space. A coursebook would be useful in this context, and the extent to which ethical guidelines can serve this purpose was discussed in PART I.

II.2 "Modell der Wert-Dimensionen" (Model of Value Dimensions)

The "Modell der Wert-Dimensionen" (Model of Value Dimensions) was conceptualized by incorporating the philosophical and linguistic insights outlined previously. Thus, it is presented as a modifiable heuristic tool that utilizes spatial imagination alongside physical insights for illustration. As with any conceptual model, there are limits to analogies. Despite such weaknesses, a pictorial heuristic aids understanding.

Horizon of Justice

The starting point is the concept of a horizontal plane. It illustrates the *horizon of justice*.

[10] See Pfeifer, 2012, p. 754: The foreign word "cybernetics" only came into the German language in the twentieth century in the course of the thematization of technical system controls. This is another interesting linguistic finding when talking about IT governance.

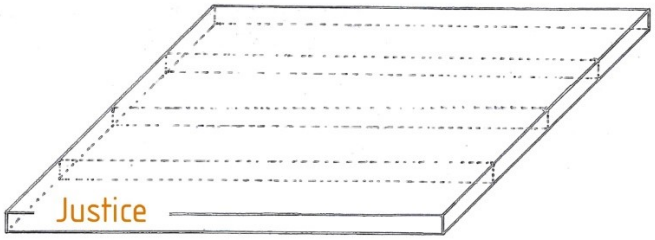


Fig. 1: Horizon of justice

One can think of the surface of a water basin. In the water basin, the density of the water below the surface differs from the lower density of the air above. If a sphere is added and thrown into the water, its density will determine whether it sinks (if denser), floats neutrally (if equally dense), or rises and floats (if less dense). The gravitational force, acting downward, opposes the buoyant force, acting upward. When these forces balance each other, hydrostatic equilibrium is achieved. For the "Modell der Wert-Dimensionen," the *equilibrium of forces* is the critical concept.[11]

"Wert-Dimensionen" within the Horizon of Justice

In the thought model, the sphere is replaced by a "Wert-Dimension," for example the dimension of the common good. There is not just one sphere placed in the thought model, but several. In principle, there could be any number of spheres in the sense of "Wert-Dimensionen."

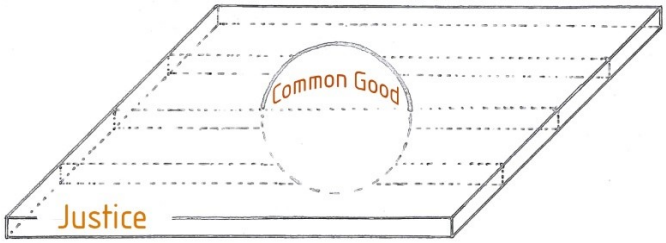


Fig. 2: The "Wert-Dimension" of the common good within the horizon of justice

Equilibrium of Forces

The crucial aspect is achieving an *equilibrium of forces* among all dimensions in their interplay. Unlike a ping-pong ball or a bowling ball, it is necessary for this conceptual model to envision the "Wert-Dimensionen" as having permeable surfaces: two dimensions can thereby form an intersection. Additionally, they can expand or contract, al-

[11] The concept of an *equilibrium of forces* within ethics evokes John Rawls. The political philosopher sought a balance in the context of justice through a different approach, known as *reflective equilibrium*. In his *Theory of Justice*, Rawls described an iterative process aimed at achieving a coherent balance between moral judgments and general principles. This methodology has a different focus than the model proposed here. Nonetheless, it could be considered when selecting the dimensions in the present model.

lowing one dimension to become larger than another. These changes can locally increase or decrease the density, affecting the *equilibrium of forces*. A temporal component, considering the consequences of actions, is also conceivable for the model. For this, the image of the water surface could be replaced by that of the Cauchy horizon in rotating black holes. However, this framework will not delve into astrophysical complexities.

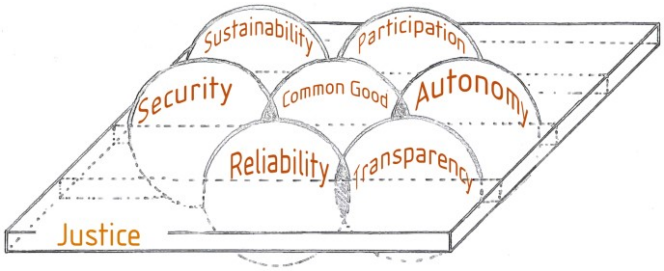


Fig. 3: The "Wert-Dimensionen" in their interplay

Weighting as Action of a Governance

Not only the selection of the "Wert-Dimensionen" is significant, but also their arrangement and respective weighting to achieve an *equilibrium of forces* in favor of justice. This also falls within the realm of governance. For example, a specific technological use case might be inherently focused on the common good, perhaps because a public contractor has specified this requirement. The ethically sovereign individual or group responsible for fulfilling the contract will also need to ensure that the common good is not weighted so heavily in the conceptual water basin that it "sinks," leaving no room for autonomy and resulting in an unjust imbalance. Similarly, determining the relationship between transparency and security requires a sensitive approach and case-specific weighting by ethically sovereign actors (weighing interests).

It is not solely due to the water metaphor that a fluid system should be considered. Every decision and weighting has systemic effects. Even if no further decisions and weightings are made, the water, in the sense of the system, does not freeze; it is subject to external influences through the passage of time, changing laws, and evolving societal notions of justice, which themselves may be subject to temporal change.

II.3 "Wert-Dimensionen" in the Context of DLT-based Information Systems

II.3.1 The Choice of the "Wert-Dimensionen"

There have already been references to advances in a blockchain ethics within the scientific discourse (see Chapter I.2.1). The ethical fields proposed for consideration in

the context of DLT do not always carry the label "values." Sometimes, the scientific discourse remains with the general mention of "ethical issues" without differentiating categories, or of "moral concerns" as seen in Ishmaev's work (Ishmaev, 2021, p. 240). The prescriptive language of "norms" was encountered in the works of Agerskov et al. (Agerskov et al., 2023; for the distinction between values and norms, see Schweppenhäuser, 2021, p. 13). Their mentioned contribution from late 2023 referred to six "ethical issues:" "Security," "Privacy," "Equality and Accessibility," "Societal and Environmental Well-being," "Accountability," and "Governance" (Agerskov et al., p. 9). LaPointe and Fishbane established what they called the "Blockchain Ethical Design Framework," avoiding speaking of ethical fields. Methodologically, they started from "Key Attributes of Blockchain" and examined each for its "social impact" (LaPointe & Fishbane, 2018). Their approach is rooted in value sensitive design. Based on the premise that values are embedded in the code, they proposed their framework to integrate values into the design process. They also named six areas for ethical reflection: "Governance," "Identity," "Verification and Authentication," "Access," "Data Ownership," and "Security" (LaPointe & Fishbane, 2018, pp. 23–24). Yong Tang et al. presented their own framework in 2019: "Ethics of Blockchain." The authors identified five relevant ethical fields: "Privacy," "Accuracy," "Property" (in the sense of data ownership), "Accessibility," and "Equality" (Tang et al., 2020, pp. 610–612). These three lists show similarities. System accessibility is always mentioned, and issues of security, privacy, and data ownership play a significant role. A different starting point compared to LaPointe and Fishbane's technical insights is found in Marković's work. Drawing on Martin Peterson's "The Ethics of Technology" (Peterson, 2017, p. 5), the author considers five moral principles as ethically significant across technologies: "cost-benefit principle, the principle of caution, the principle of sustainability, the principle of autonomy, and the principle of fairness" (Marković, 2020, p. 435). Following this, Marković questions possible "embedded values," which also leads him to value sensitive design from this perspective (Marković, 2020, p. 437). This aligns well with his advocacy for a virtue ethics-oriented blockchain ethics and also resonates with the approach represented here.

The "Wert-Dimensionen" explained below are: *common good*, *sustainability*, *autonomy*, *security*, *participation*, *transparency*, and *reliability*. As outlined above, they stand in relation to each other as well as to specific applications. This creates variables that cannot be resolved normatively. They need to be addressed and balanced by ethical actors. The guiding measure here is the concept of the *horizon of justice*. The specific selection of the seven dimensions reflects the state of research as of spring 2024, considering ethical insights and applicable laws in connection with the technological peculiarities of blockchain and drawing on scientific literature. How they relate individually to the aforementioned "ethical issues" is also demonstrated. For example, "data ownership" in the "Modell der Wert-Dimensionen" is assigned to "autonomy," while system accessibility is linked to the dimension of participation. "Governance," as listed by the Copenhagen research group, is instead understood as categorically different and as an activity of ethical actors based on values.

II.3.2 The Relevance of the Seven "Wert-Dimensionen" in Detail

The seven proposed dimensions for consideration are understood not hierarchically, but heterarchically, on an equal footing, and are further explained in the following order:

- *Common good*
- *Sustainability*
- *Autonomy*
- *Security*
- *Participation*
- *Transparency*
- *Reliability*

It is assumed that a DLT use case touches upon all dimensions, with a determination of their relationship in the specific application case, also considering intersections and partial collisions. In case of collision, weighting becomes necessary. Decision-making regarding evaluation and prioritization should be justified with respect to the *horizon of justice*. In the "Modell der Wert-Dimensionen," this is illustrated as a situation-sensitive and deliberate action by ethically sovereign actors, such as within a governance framework. Fundamentally, the "Modell der Wert-Dimensionen" is seen as open to various applications as well as the inclusion of additional dimensions that may arise. The choice of the seven dimensions outlined below was guided by two specific use cases. These are the use cases that are the subject of research in the LIONS project:

- *Implementation of DLT in the food supply chain*
- *Implementation of DLT in the area of self-sovereign identity (SSI)*

These use cases have particularities within the spectrum of DLT applications that need to be considered in ethical reflection. One initial consideration is that they are applications initiated in the *public sector*, but their implementation involves both various-sized *private enterprises* in the food supply chain and *individuals* in terms of SSI. The interaction of the different actors may lead to various interests, resulting in partial collisions within or with the proposed dimensions, as well as conflicts of interest among actors. The choice and weighting of specific dimensions are influenced by negotiation processes. The specific constellation of stakeholders and their respective heterogeneity must be taken into account, particularly in an application-specific manner. A second consideration is the *diversity of legal and cultural spaces* that the supply chain traverses. Consequently, the choice and weighting of specific dimensions are not subject to arbitrary negotiation outcomes, but are bounded by societal desirability as well as legal permissibility.

II.3.2.1 Common Good

Promotion of the common good as a sociopolitical goal has already been discussed (see Chapter I.1.2 and I.2.1). The expansion of digital infrastructure, for instance, is intended to benefit the common good. Therefore, DLT is already being used for projects that promote the common good.[12] To ensure that the common good does not become a mere empty phrase, and thus meaningless at worst, it is important to consider the ethical intention behind it. When the common good is supposed to play a role in the development and the design of an information system, it refers to the heterogeneous interests involved in the decision-making process for DLT-based information systems, their design, and subsequent use. The perspective of the common good transcends individual endeavors, taking into account the legitimate needs underlying the interests.

Human action is always motivated by interests that serve as incentives for decision-making. With regard to DLT, incentive systems are accorded correspondingly high relevance. Even the language of marketing by software providers promotes interest fulfillment. While the public sector has an aspiration for increased security, simplified administrative processes, and even digitally mediated democracy as motivations for the introduction of blockchain, its overall societal benefit[13] is also emphasized. Companies are promised, among other things, economic gains through data sharing,[14] while individuals are promised a novel relationship of trust with reduced risks in digitally mediated social interaction (see Bayerisches Staatsministerium für Digitales, 2024).

The ethically responsible use of any technology presupposes that the diffuse motivational landscape confronted by developers and designers transcends to promote the common good. The concept of the common good stands for "the well-being of all members of a community [...] as opposed to private welfare and particular interests" (Schultze, 2010, p. 299, translated; see also Gräb-Schmidt & Preul, 2014), without leveling minorities in a pluralized society. In DLT projects in the public sector, the perspective of the common good should always be present. However, it also does not stand in the way of privately oriented projects, as long as they do not exclusively serve particular interests. In practice, this may mean appropriately taxing transparent transactions and profits to benefit the general public. Questions of system accessibility (dimension of participation) and autonomy may also play a role as frequent intersections

[12] An example of the explicit focus on the common good in a DLT-based initiative: Shweta Jain and Rahul Simha have shown how donation flows can be mapped on a blockchain basis. The idea is that the ability to donate to a specific project and track the exact path of the donation increases the willingness to donate and thus also the common good. See Jain & Simha, 2018.

[13] The European Economic and Social Committee has expressed its overall optimism about blockchain technology (Europäischer Wirtschafts- und Sozialausschuss, 2019, translated): "It can bring about positive change in many areas of society, based on values such as trust and transparency, democracy and security."

[14] The IEDS project funded by the German Federal Ministry of Education and Research tested aspects including the extent to which the reference to economic successes changes the willingness to use (Fraunhofer Institut für Software und Systemtechnik ISST, 2022, p. 28, translated): "As scientific studies show, companies can increase their innovative activity, productivity, and profits by using data."

with the dimension of the common good. The conscious and differentiated reflection on whether specific *interests*, brought forth from different sides in the form of development requirements, also reflect *legitimate needs* within a common good orientation, can accompany ethical positioning and decision-making within the framework of value sensitive design. In this way, interest weighing can also be brought into an appropriate balance within the *horizon of justice*.

II.3.2.2 Sustainability

Sustainability furnishes a "Wert-Dimension," which is initially evident in its relevance from the genesis of DLT and its first use cases in the area of decentralized finance. Also, the fact that the Committee on Economic and Monetary Affairs of the European Parliament voted on the legitimacy of the proof of work mechanism (PoW) in 2022 had a significant reason in the high energy intensity of mining in the context of cryptocurrencies. The immense resource demand of mining, which Bitcoin still relies on, unlike Ethereum, prompted the call for minimum standards regarding ecological sustainability (see de Vries, 2024, for the ecological footprint of BTC). Although the proposed draft of the EU Parliament was narrowly rejected, a unified legal framework regarding cryptocurrencies in the EU was subsequently favored, intended to also consider harmful environmental impacts due to high energy consumption and resulting electronic waste, and enable regulatory measures (see Kolinska, 2022). As already mentioned (see Chapter I.1.2), the generally increasing electronic waste has meanwhile received a legal regulation in the EU.

Although alternatives to PoW, such as proof of stake within Ethereum, exist, an energy consumption that is relatively more energy-efficient compared to extremely high consumption alone does not yet justify the use of the technology from an ecological perspective. Therefore, it will be important to ensure early on in the deployment of DLT-based information systems that it is a genuine system decision that represents a justifiable sustainable choice compared to alternative solutions in terms of resource conservation. The use of DLT is to be seen in the context of comprehensive socio-ecological transformation and must justify itself before the claim of the European Digital Green Deal (see European Commission, 2024c). As a political reference framework, the Rio Conference of 1992 (see Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, 2024), the Sustainable Development Goals of 2015 within the framework of Agenda 2030 (see Bundesregierung, 2024), as well as the further development of the German Sustainability Strategy of 2021 (Bundesregierung, 2021) should also be mentioned. The triad "sufficiency, efficiency, consistency" is often cited as a criterion of sustainability as well (e.g., Reis, 2003, p. 247).

Beyond minimizing damage in the operation of IT systems, the "Digital and Ecological Transformation Agenda" of the joint project CO:DINA (IZT & Wuppertal Institut für Klima, Umwelt und Energie, 2023) goes further. Digital systems were seen here in terms of their potential to support a targeted and directed comprehensive societal sustainability strategy. A corresponding strategically acting and incentive-creating governance is rightly demanded. Sustainability thereby becomes a primary concern

and requirement for IT systems, no longer a debatable side issue through their operation. That blockchain-based information systems are now not only known for their high CO_2 footprint but are also being developed in reverse to track and subsequently reduce CO_2 emissions along a supply chain (see Seidenfad et al., 2022) is to be welcomed as a gratifying development in line with such sustainability strategy.

In ethical reflection, the call for sustainability fits into the line of responsibility ethics. Reference has already been made to a corresponding draft by the philosopher Hans Jonas (see Chapter I.1.2). A core idea lies in the recognition that today's decisions and actions determine future quality of life. Therefore, making oneself aware of the scope of one's own actions is not only a demand on experts in the field of technology assessment but remains essential for programmers, designers, operators, and users of an IT system. That sustainability is directly linked to questions of social justice (see Jenkins, 2013) will also play a significant role in balancing and weighting various "Wert-Dimensionen" within the *horizon of justice*.

II.3.2.3 Autonomy

In research discourse, autonomy in the sense of self-determination in the digital sphere is often linked to digital sovereignty. In contrast to the common good, self-determination directly aims at the well-being of individual persons. Self-determination understood as informational freedom is primarily implicated in the use of DLT in the realm of SSI. The scientific discourse on data ownership can also be framed under "autonomy." When linking data sharing to the aspect of self-determination, one represents a broad understanding of *self*. Although one might think more about a company's financial figures rather than personal data, consideration can extend beyond the individual level to include organizations or even states. This approach enables the analysis and evaluation of use cases such as supply chains from the perspective of autonomy.

Technical ethical reflections on autonomy can draw upon advanced medical ethical discourse and related analogies. Self-determination, as "Respect for Autonomy," is one of the four principles known as the "Georgetown Mantra" (Beauchamp & Childress, 1979) in scholarly literature, alongside "Nonmaleficence," "Beneficence," and "Justice." The principle of "Nonmaleficence" is also given in the "Modell der Wert-Dimensionen", along with the dimensions of security and reliability. Furthermore, "Justice" manifests beyond the *horizon of justice* in this model – for example, in the form of participatory justice (see Chapter II.3.2.5).

The dimension of autonomy is accompanied by the necessity of empowering self-determination in the digital sphere, as well as the need for corresponding technical infrastructure. Self-determined judgment and action also require the ability to acquire cognitive knowledge and practical experience. This entails an obligation on the part of system developers and operators to provide information. Regarding the goal of ensuring that individuals are "better informed about risks, costs, and charges" (Kolinska, 2022), this obligation does not solely involve making the software available as open source, which will be addressed within the transparency dimension. Additionally, legal implications need to be clarified, particularly regarding the right to rectification

under Art. 16 GDPR and the right to erasure under Art. 17 GDPR, before implementing DLT-based systems. According to the prevailing understanding, these legal claims contradict the goal of DLT, which aims to ensure complete and immutable data documentation that cannot overwrite previous entries. The attempt at a legally secure solution by BAMF has already been mentioned (see Chapter I.2.1).

At the level of IT governance, the dimension of autonomy must also be considered, especially when corporate objectives collide with the individual needs of employees. Moreover, autonomy implies that autonomous individuals should be able to freely decide against the use of DLT-based information systems, implying that DLT should not be pushed forward as the only alternative development. Lastly, legal regulations for the benefit of digital accessibility must be adhered to, which will be addressed within the explanations of the participation dimension.

II.3.2.4 Security

In the discourse surrounding DLT-based information systems, security is often cited. Several approaches in scholarly blockchain ethics address this aspect (see Chapter II.3.1). Security is seen in protection against system failure, cyber-attacks, and data manipulation. This makes DLT attractive for the public sector and relevant for interactions between government bodies and private individuals. Self-determination initially emphasizes the idea of sovereignty more strongly in relation to *individuals*. In contrast, *supra-individual* stakeholders are also more concerned with sovereignty in favor of security. This is evident in the "Cybersecurity Strategy for Germany 2021" (Bundesministerium des Innern, für Bau und Heimat, 2021, translated), which explicitly aims for digital sovereignty in the face of security threats. The hybrid warfare conducted by Russia has further intensified political traction regarding cybersecurity, especially concerning the protection of critical infrastructure (see Bundesministerium des Innern und für Heimat, 2024). At the EU level, the "Cybersecurity Strategy for the Digital Decade" of 2020 is noteworthy (European Commission, 2020). In the context of *security*, the goal of *resilience* is also articulated in the corresponding EU law of 2022 (European Commission, 2022).

A *right to security* can be derived from the right to physical integrity (Art. 2 II 1 GG). The latter is tied to personal freedom (Art. 2 II 2 GG) but is understood as "subject to balancing with other legal interests" (Kutscha, 2006, p. 41, translated) (Art. 2 II 3 GG). Explicitly, the right to security is mentioned – again linked with a right to freedom – in the EU Charter of Fundamental Rights.[15] The constitution presupposes that it is the state's task to ensure the security of the population (see Kutscha, 2006, p. 26). This touches upon a societal contractual determination between the rights of individuals and state guarantees and regulations. Ambivalences between self-determination – and the corresponding autonomy dimension in the model – and state paternalism are conceivable. The fundamental rights are intended to ward off a potential dominance of the state. Faced with a potential tension between freedom and security provision, it

[15] In addition to the "right to physical and mental integrity" (Art. 3 I Charter of Fundamental Rights of the European Union), there is also the specific "right to liberty and security" (Art. 6 CFR).

becomes the task of law to maintain an appropriate balance (see Mackenroth, 2011). This also applies to cybersecurity. It is crucial to maintain the understanding that security, provided by state bodies as well as the means employed for its realization – such as DLT-based information systems – "serve the civil freedom" (Kutscha, 2006, p. 29, translated). Therefore, mere reference to security does not constitute a legally compliant causality in case of a restriction of freedom. Instead, prioritizing security over freedom must be transparently and legally justified; measured against the *horizon of justice* with reference to the "Modell der Wert-Dimensionen."

Regarding the development and design of DLT, it is therefore necessary to address uncertainties in law, as outlined in Chapter II.3.2.3 (Autonomy), and not undermine them with mere reference to security or state digital sovereignty. At the same time, security issues are of high importance for ethical assessment. If DLT provides better protection against system failure, cyber-attacks, and data manipulation, it is ethically preferable over other system solutions. Such potential for protecting digital infrastructure is explicitly recognized by the German government in blockchain and other "future and key technologies" (Bundesministerium des Innern, für Bau und Heimat, 2021, translated). In this context, the relationship between "security" and "reliability" will be particularly important to analyze within the "Modell der Wert-Dimensionen" and to be weighted situationally in an application-oriented manner.

II.3.2.5 Participation

DLT also holds significant potential for participation. Possibilities for democratization, participation, and inclusion are envisioned (see Kossow, 2019, p. 97). In public discourse, opportunities for participation are repeatedly linked to expectations of digitization in general (see Dettling, 2019, p. 11). It is often heard that blockchain, due to its decentralized technical structure – although not unique to this technology – is particularly suitable for realizing democratization and participation in decision-making processes. However, the notion that decentralized network nodes automatically lead to balanced power relations is a misconception (see Hofman et al., 2021, p. 24). In the ethical assessment of an information technology system, it is therefore important to note that technical characteristics cannot simply be equated with social reality. The same fallacy occurred with the misunderstanding that technical systems could guarantee a certain morality upon later application (see Chapter I.1.2). Therefore, with regard to specific applications, it is necessary to assess whether increased social participation can be expected through digital participation.

In development and design, diversity of users should be considered, achieved through diverse personas – while avoiding excessive stereotyping –, corresponding user journeys, and outcome-open iterative process design with opportunities for participation to real needs (for diversity-sensitive participatory design, see Erharter, 2015, pp. 87–88). Additionally, an internal authorization concept within the system can be considered regarding the dimension of participation and discussed based on granted rights and roles, considering how the dimensions of participation, autonomy, and security relate to each other in specific cases.

Only a participation concept anchored at the level of overarching IT governance will create the conditions to facilitate participation. Public projects will particularly be committed to the *right to participation* (see Bundesministerium für Arbeit und Soziales [BMAS], 2024b). Digital accessibility should be a legally mandated matter with the Barrierefreiheitsstärkungsgesetz (see BMAS, 2024a), which also provides guidance for private sector projects (see Chapter II.3.2.3). The goal is "that all people can use digital offers, regardless of their physical and mental abilities" (see BMAS 2024a., translated). The perception of autonomy requires the possibility of participation. The German regulation is based, incidentally, on the European Accessibility Act of 2019 (see European Commission, 2024a). However, the dimension of participation concerns not only people with disabilities. It is also necessary to avoid a digital divide (see Norris, 2001, pp. 3–25), arising when system access and usage are determined by socio-economic factors and resulting unequal opportunities. This affects also companies, which may be displaced from the market due to monopolization – including or especially in the use of DLT.[16] Unequal conditions are apparent along the supply chain (see Fries & Greiner, 2023). Nevertheless, DLT has the potential to inclusively involve startups or address the trend towards the sharing economy (see Dettling, 2019, p. 15) technologically.[17] The "Wert-Dimension" of participation contributes to case differentiation from ethical perspectives.

II.3.2.6 Transparency

The dimension of transparency initially evokes thoughts of increased data integrity through DLT. Data transparency is ethically advantageous when the accurate transaction history of a blockchain complicates fraud. However, it also presents ethical and legal challenges. These arise when sensitive data is involved, and the owners desire no, only limited, recipient-specific, situational, or time-limited transparency, and such restrictions must also be legally ensured. Therefore, a legally compliant and ethically legitimate technology must be capable of situationally deciding who should have transparency over which data. Technically, this can be regulated through corresponding keys, which should also be retrofittable. Conversely, this can mean not storing any sensitive data in a blockchain (see Chapter II.3.2.3).

In practice, it is also evident that transparency requires a *culture of transparency* (see Lehner et al., 2020, p. 82). This is linked to the willingness to share non-sensitive data – especially a willingness among those who do not want to be suspected of fraudulent behavior. Those who perceive themselves as righteous in their actions may be hesitant to respond to a system that aims to prove righteousness. Therefore, it is important to

[16] If, in future, municipal contracts were to be awarded on the basis of DLT reputation tokens, this would probably exclude unreliable partners, but could also increasingly leave behind those who have never been able to demonstrate their reliability because they are too small and have never been able to earn digital merits. This could widen the gap.

[17] One possible application of DLT is in the context of electric rental cars. This would make it easier for more people to use these expensive vehicles and the energy consumption of DLT could be offset by savings associated with not using individual combustion engines.

establish the sharing mindset in a positive way, both in society as a whole and, especially, in the business sector, and to set positive incentives for sharing (see Chapter II.3.2.1). It is important to avoid damaging social resources by propagating the idea of being deceived as the normality of interpersonal, business or political relationships in a negatively delimiting argumentation, declaring a fundamental mistrust as necessary and recommending technology as a solution to the social misery described. Such an approach ultimately leads to the paradox that sown mistrust decreases willingness to share instead of increasing it. Shifting trust to technology also has insidious consequences, as it undermines awareness of technology-related security risks (see Fries, 2022). In this context, the "suggestion of security" in the context of the oracle problem becomes relevant (see Chapter II.3.2.7).

Furthermore, transparency is not limited to data transparency. Transparency regarding the opportunities and risks of a technology, as well as disclosure of motives for its use, should also be considered. Informative transparency can strengthen participation and autonomy. It is also important to note that information that is transparent is not necessarily public.[18] Therefore, providing the code as open source should be accompanied by efforts to ensure understandable interpretation – for example, based on the FAIR Principles: *F*indable, *A*ccessible, *I*nteroperable, and *R*eusable (see GO FAIR International Support and Coordination Office, 2016).

II.3.2.7 Reliability

The "Wert-Dimension" of reliability is closely related to the dimensions of security and transparency. It also encompasses the principle of nonmaleficence, as previously indicated (see Chapter II.3.2.3). In the context of DLT-based systems, reliability is intended to be achieved through decentralized structure, system-regulated processes, and resulting high data integrity, according to the general discourse. Points of reference for reliability thus include technical prerequisites and functions. This fundamental understanding – reliability through technology – aligns with the genesis of DLT. Its original claim was to replace a system whose "actions" are considered predictable and verifiable at any time with less reliable and erratic-seeming individuals or instances. Particularly, an uncontrollable human intermediary was supposed to yield to the incorruptibility of the integral, regulated, and transparent system (smart contracts). Even in scientific discourse, it is often misleadingly stated that *trust* can be replaced or generated through blockchain.[19] However, the *reliability* of a system in terms of functionality must be distinguished from trust in interpersonal interaction (see Fries, 2022). This differentiation must be maintained as a matter of urgency to prevent blind trust in the system and to prevent operators and users from becoming careless in their

[18] See Hartmann, 2020, p. 202, translated: "Transparency means the accessibility of information or the possibility of examining an issue. Publicity means that information has already been processed, i.e. that there are already people who are familiar with the information and have communicated it to other people."

[19] The aim of Satoshi Nakamoto's white paper, in which DLT originated, was to create a technical transaction system that does not require trust (see Nakamoto, 2008).

behavior due to the suggestion of reliability and trustworthiness, and thus becoming a security risk themselves (see Chapter I.2.2).

Moreover, inquiries into system reliability arise. In principle, decentralization can be seen as providing relatively higher reliability and resilience in crisis situations. In addition, the data storage in blockchain systems does not depend on whether there is a local power outage; however, the reliable usability of those affected by the power outage does. In the event of a rapid food recall, the system may *function*, but it may not be locally *usable* despite system functionality. Therefore, when deployed along global supply chains, potential region-specific diversity in terms of energy supply, as well as crisis management competence, must be considered. When various stakeholders communicate through the same system, it is also essential to keep crisis-resistant communication channels open when deployed in the critical infrastructure sector. Also worth considering is society's reliance on technologies as a whole, which creates dependencies, given that energy supply also relies on IT systems (see Voßschmidt & Karsten, 2019, p. 175).

Regarding data integrity, it is also noteworthy that the oracle problem, which concerns the quality of newly inputted data, has not yet found a solution. Data transparency is therefore not necessarily transparency regarding the realities behind the data (see Fries, 2023). There are many motives and possibilities for incorrect initial data, which can cause subsequent errors in the system. The value of DLT-based systems is therefore seen in *relative* reliability. Their introduction in a specific context should always be accompanied by efforts to minimize errors in data input – for example, by increasing the awareness of employees, providing suitable working conditions,[20] and regularly maintaining data-gathering devices or systems integrated into the process.

II.4 Outlook on Practice-Oriented Orientation in the Space of Values

PART I focused on the importance of ethical sovereignty in the digital sphere. It spoke of "Components for Responsible Use" (see Chapter I.2). As a heuristic tool for timely and situation-sensitive application, the "Modell der Wert-Dimensionen" was developed and introduced in PART II (see Chapter II.2). This model was filled with a specific set of dimensions in the context of DLT-based information systems and explained in Chapter II.3: *common good*, *sustainability*, *autonomy*, *security*, *participation*, *transparency*, and *reliability*.

In conclusion, there is an outlook. "Wert-Dimensionen," values, are not norms. The application of the "Modell der Wert-Dimensionen" lies in the hands of the readers, those who make decisions, those who develop, operate, and use systems. It lies in the

[20] This implies, for example, sufficient time for accuracy, adequate payment to avoid deliberate mistakes for financial gain, and a dual control principle. In addition, the apparent paradox of a simultaneous culture of error tolerance is conceivable in order to avoid errors based on stress or fear, as well as "meta-working," in which actual work and documented work based on requirements fall apart.

hands of persons who act responsibly and who, with the present "Ethical Guidelines for DLT-based Information Systems," find considerations and recommendations, but above all, encouragement for independent reflection and assumption of responsibility in the *horizon of justice*. They are the ones who can not only learn about the "Wert-Dimensionen," but also consider them and analyze them in the context of specific applications. They are the ones who can discuss the interaction of different dimensions and seek an *equilibrium of forces* through case-based arrangement and sensitive weighting. They are also the ones who can remain in interdisciplinary discourse so that the common reflection does not stand still and accounts for the fluid system in the *space of values* – even with regard to new developments beyond the time of the creation of this research contribution. For value-sensitive governance, all of this is as essential as it is for individual awareness, which includes the *horizon of justice* with all its possible dimensions.

Where there are no simple checklists, step-by-step instructions, and binding norms, *prudence* is required. According to philosophical understanding and with reference to the aforementioned Aristotle (see Chapter II.1), "prudence" is a virtue of the mind. With it, one can become aware of ethical goals and choose means and ways to achieve them. Prudence knows not only what is commanded but also the uniqueness of a specific situation. Moreover, it is based on practical experiences. It enables consideration, judgment, and disposition, as Thomas Aquinas pointed out (see Prechtl, 2008, p. 293). In this sense, the present "Ethical Guidelines" are rightly understood when they serve prudent use and thus increase ethical sovereignty in the context of DLT-based information systems.

Acknowledgments

This work originated in the transdisciplinary LIONS research project. I would like to express my gratitude to all LIONS project collaborators and other discussion partners from the fields of computer science, business information systems, pedagogy, psychology, philosophy, theology, ethics, and communication studies who contributed to this research outcome through engaged discourse. LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, which I gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

References

Agerskov, S., Pedersen, A. B., & Beck, R. (2023). Ethical Guidelines for Blockchain Systems. *ECIS 2023 Research Papers*. 275. https://aisel.aisnet.org/ecis2023_rp/275.

Apel, M. (1953). *Philosophisches Wörterbuch* (4th ed.). De Gruyter.

Arendt, H. (1963). *Eichmann in Jerusalem: A Report on the Banality of Evil*. Faber.

Bauer, W., Aland, K., B, & Aland, B. (1988). *Griechisch-deutsches Wörterbuch zu den Schriften des Neuen Testaments und der frühchristlichen Literatur* (6th ed.). De Gruyter.

Bayerisches Staatsministerium für Digitales (2024, May 10, accessed). *Blockchain-Strategie. Block – Chain – Trust*. https://stmd.bayern.de/themen/bavarian-center-for-blockchain/strategie.

Beauchamp, T. L., & Childress, J. F. (1979). *Principles of Biomedical Ethics*. Oxford University Press.

Bundesministerium des Innern, für Bau und Heimat (2021). *Cybersicherheitsstrategie für Deutschland 2021*. https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/2021/09/cybersicherheitsstrategie-2021.pdf.

Bundesministerium des Innern und für Heimat (2024, May 10, accessed). *Cybersicherheitspolitik des Bundes: Agenda und Strategie*. https://www.bmi.bund.de/DE/themen/it-und-digitalpolitik/it-und-cybersicherheit/cybersicherheitspolitik/cybersicherheitspolitik-node.html.

Bundesministerium für Arbeit und Soziales (2024a, May 10, accessed). *Barrierefreiheitsstärkungsgesetz*. https://bmas.de/DE/Service/Gesetze-und-Gesetzesvorhaben/barrierefreiheitsstaerkungsgesetz.html.

Bundesministerium für Arbeit und Soziales (2024b, May 10, accessed). *Bundesteilhabegesetz*. https://www.bmas.de/DE/Service/Gesetze-und-Gesetzesvorhaben/bundesteilhabegesetz.html.

Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (2024, May 10, accessed). *UN-Konferenz für Umwelt und Entwicklung (Rio-Konferenz 1992)*. https://bmz.de/de/service/lexikon/un-konferenz-fuer-umwelt-und-entwicklung-rio-konferenz-1992-22238.

Bundesregierung (2021). *Deutsche Nachhaltigkeitsstrategie: Weiterentwicklung 2021*. https://bundesregierung.de/resource/blob/975274/1873516/9d73d857a3f7f0f8df5ac1b4c349fa07/2021-03-10-dns-2021-finale-langfassung-barrierefrei-data.pdf.

Bundesregierung (2024, May 10, accessed). *Die 17 globalen Nachhaltigkeitsziele verständlich erklärt*. https://bundesregierung.de/breg-de/themen/nachhaltigkeitspolitik/nachhaltigkeitsziele-erklaert-232174.

Camus, A. (1951). *L'homme révolté*. Gallimard.

Deutscher Bundestag, Parlamentsnachrichten (2023, November 1). *DB InfraGO kommt in der Rechtsform einer Aktiengesellschaft*. https://bundestag.de/presse/hib/kurzmeldungen-974546.

Dettling, D. (2019). Zukunftswert Partizipation: Keine soziale Teilhabe ohne digitale Teilhabe. In S. Skutta & J. Steinke (Eds.), *Digitalisierung und Teilhabe. Mitmachen, mitdenken, mitgestalten!* (pp. 11–24). Nomos.

de Vries, A. (2024). Bitcoin's growing water footprint, *Cell Reports Sustainability,* vol. 1, no. 1, January 2024. https://doi.org/10.1016/j.crsus.2023.100004.

EGBE (2024). *Ethical Guidelines for Blockchain Systems.* European Blockchain Center. https://ebcc.eu/wp-content/uploads/2024/05/Ethical_Guidelines_for_Blockchain_Systems.pdf.

Erharter, D. (2015). Gender- und Diversity-Dimensionen in der Entwicklung von IKT-Projekten. In H. & J. Siegeris (Eds.), *Gender und IT-Projekte: Neue Wege zu digitaler Teilhabe* (pp. 79–94). Budrich UniPress.

Europäischer Wirtschafts- und Sozialausschuss (2019). Stellungnahme zu „Blockchain und der EU-Binnenmarkt: Wie geht es weiter?". *Amtsblatt der EU, 2020/C47/03*. https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52019IE2261.

European Commission (2024a, May 10, accessed). *European accessibility act*. https://ec.europa.eu/social/main.jsp?catId=1202.

European Commission (2022, September 15). *Cyber Resilience Act*. https://digital-strategy.ec.europa.eu/en/library/cyber-resilience-act.

European Commission (2024b, May 10, accessed*). European Blockchain Service Infrastructure*. https://digital-strategy.ec.europa.eu/en/policies/european-blockchain-sevices-infrastructure.

European Commission (2024c, May 10, accessed). *Green digital sector*. https://digital-strategy.ec.europa.eu/en/policies/green-digital.

European Commission (2020). *The EU's Cybersecurity Strategy for the Digital Decade*. https://ec.europa.eu/newsroom/dae/redirection/document/72164.

European Blockchain Center (2024, May 10, accessed). *Blockchain Ethics*. https://ebcc.eu/ethics.

Federal Office for Migration and Refugees (2024). *The federal asylum blockchain infrastructure*. https://bamf.de/EN/Themen/Digitalisierung/FLORA/flora-node.html.

Feloutzis, N., & Lekakos, Georgios (2023). Is Blockchain a Disruptive Innovation? a Systematic Literature Review. In D. H. de la Iglesia, J. F. de Paz Santana, & A. J. Lopéz Rivero (Eds.), *New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence. The DITTET 2023 Collection* (pp. 175–186). Springer.

Fenner, D. (2020). *Ethik* (2nd ed.). Attempto.

Floridi, L. (2022). The Green and the Blue: A new Political Ontology for a Mature Information Society. In L. Floridi & J. Noller (Eds.), *The Green and the Blue: Digital Politics in Philosophical Discussion* (pp. 9–51). Karl Alber.

Fries, I. (2023). Daten, Freiheit, Wirklichkeit: Deutungshoheit in Zeit und digitalem Raum. In A. Schaffer & E. Lang (Eds.), *Die Kunst vom Finden und Bauen der Wirklichkeit* (pp. 159–182). Metropolis.

Fries, I. (2022). „In Code We Trust?" – Zur Vertrauens-Verheißung der Blockchain-Technologie. *ZEE 4/2022*, 264–276.

Fries, I., & Greiner, M. (2023). Technology-enabled Fairness? Reflections on Fairness within Blockchain-based Supply Chain Consortia, 2023. In *EURAM 2023 Transforming Business for Good. Proceedings*.

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., & Wendeborn, T. (2023). Towards a Layer Model for Digital Sovereignty: A Holistic Approach. In B. Hämmerli, U. Helmbrecht, W. Hommel, L. Kunczik, & S. Pickl (Eds.), *Critical Information Infrastructures Securit. 17th International Conference, CRITIS 2022*, Munich, Germany, September 14–16, 2022, Revised Selected Papers, 119–139. Springer. https://doi.org/10.1007/978-3-031-35190-7_9.

Frank, H. (1942). *Die Technik des Staates, Mit Geleitwort von Ernst Letzgauß*. Deutscher Rechtsverlag.

Fraunhofer Institut für Software und Systemtechnik ISST (2022). *Anreizsysteme und Ökonomie des Data Sharings: Handlungsfelder des unternehmensübergreifenden Datenaustausches und Status quo der deutschen Wirtschaft*. https://ieds-projekt.de/wp-content/uploads/2022/03/IEDS-Whitepaper.pdf.

Gemoll, W., & Vretska, K. (2006). *Gemoll: Griechisch-deutsches Schul- und Handwörterbuch* (10th ed.). Oldenbourg Schulbuchverlag.

Generaldirektion Kommunikation (2024, February 2). *Kommission begrüßt politische Einigung über neue Verbraucherrechte für einfache und günstige Reparaturen*. https://ec.europa.eu/commission/presscorner/detail/de/ip_24_608.

German Ethics Council (2018). *Big Data and Health – Data Sovereignty as the Shaping of Informational Freedom.* https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/englisch/opinion-big-data-and-health-summary.pdf.

GO FAIR International Support and Coordination Office (2016). *FAIR Guiding Principles for scientific data management and stewardship*. http://go-fair.org/fair-principles.

Gräb-Schmidt, E. & Preul, R. (2014). *Gemeinwohl*. Marburger Theologische Studien (121). EVA.

Gräb-Schmidt, E. (2002). *Technikethik und ihre Fundamente: Dargestellt in Auseinandersetzung mit den technikethischen Ansätzen von Günter Ropohl und Walter Christoph Zimmerli*. Theologische Bibliothek Töpelmann (118). De Gruyter.

Grunwald, A. (2013). Einleitung. In A. Grunwald (Eds.), *Handbuch Technikethik* (pp. 1–12). J. B. Metzler.

G'sell, F., & Martin-Bariteau, F. (2022). *The impact of blockchains for Human Rights, Democracy and the Rule of Law. Report to the Council of Europa. Council of Europe.* https://edoc.coe.int/en/module/ec_addformat/download?cle=5779e947caedd3537d114bf8b0702a15&k=eace807208b9da1300b6137fdac3b312.

Hare, S. (2022). *Technology is not neutral: A short guide to technology ethics.* London Publishing Partnership.

Hartmann, M. (2020). *Vertrauen: Die unsichtbare Macht*. Fischer.

Hau, R. et al. (2003). *Pons: Wörterbuch für Schule und Studium Latein-Deutsch* (3rd ed.). Klett.

Held, B., & van Oorschot, F. (2021, Eds.). *Digitalisierung: Neue Technik – neue Ethik? Interdisziplinäre Auseinandersetzung mit den Folgen der digitalen Transformation*. FEST Forschung (1). heiBOOKS.

Hofman, D., DuPont, Q., Walch, A., & Beschastnikh, I. (2021). Blockchain Governance: De Facto (x)or Designed? In V. L. Lemieux & C. Feng (Eds.), *Building Decentralized Trust: Multidisciplinary Perspectives on the Design of Blockchains and Distributed Ledgers* (pp. 21–33). Springer.

Hubig, C. (2015). *Die Kunst des Möglichen III: Grundlinien einer dialektischen Philosophie der Technik. Macht der Technik*. Transcrcipt.

Hubig, C. (1995). *Technik- und Wissenschaftsethik. Ein Leitfaden* (2nd ed.). Springer.

Illich, I. (1973). *Tools for Conviviality*. Harper & Row (German translation, 1975: *Selbstbegrenzung: Eine politische Kritik der Techni*k, translated by Ylva Erikkson-Kuchenbuch. Rowohlt).

Ishmaev, G. (2021). Sovereignty, privacy, and ethics in blockchain-based identity management systems. *Ethics and Information Technology 23,* 239–252. https://doi.org/10.1007/s10676-020-09563-x.

IZT – Institut für Zukunftsstudien und Technologiebewertung gemeinnützige GmbH / Wuppertal Institut für Klima, Umwelt, Energie GmbH (2023). *Systemwandel by sustainable design oder digital disruption: Impulse für eine digital-ökologische Transformationsagenda.* https://codina-transformation.de/wp-content/uploads/Ramesohl-et-al-2023_CODINA-Transfomationsreport.pdf.

Jain, S., Simha, R. (2018). Blockchain for the Common Good: A Digital Currency for Citizen Philanthropy and Social Entrepreneurship. *IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Los Alamitos / CA 2018*, 1387–1394. https://doi.org/10.1109/Cybermatics_2018.2018.00238.

Jenkins, W. (2013). *The Future of Ethics: Sustainability, Social Justice, and Religious Creativity*. Georgetown University Press.

Jonas, H. (1979). *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation*. Insel.

Kant, I. (1784). *Beantwortung der Frage: Was ist Aufklärung?* In AA VIII (pp. 33–42).

Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten*. In AA VI.

Kant, I. (1793). *Über den Gemeinspruch. Das mag in der Theorie richtig sein, taugt aber nicht in der Praxis* (1793). In AA VIII (pp. 273–313).

Kant, I. (1786). *Was heißt: sich im Denken orientieren?* In AA VIII (pp. 131–147).

Khan, A., & Equbal, M. T. (2023). Exploring the Potential of Blockchain in Data Science. *International Journal of Scientific Research in Engineering and Management*, 07(09). https://doi.org/10.55041/IJSREM25669.

Kirchschläger, P. G. (2021). *Digital Transformation and Ethics: Ethical Considerations on the Robotization and Automation of Society and the Economy and the Use of Artificial Intelligence*. Nomos.

Kolinska, D. (2022, March 14). *Cryptocurrencies in the EU: New rules to boost benefits and curb threats*. European Parlament. https://europarl.europa.eu/news/en/press-room/20220309IPR25162/cryptocurrencies-in-the-eu-new-rules-to-boost-benefits-and-curb-threats.

Körtner, U. H. J. (2017). *Diakonie und Öffentliche Theologie*. Vandenhoeck & Ruprecht.

Kossow, N. (2019). Blockchain: viel Potential, begrenzte Umsetzbarkeit. In S. Skutt & J. Steinke, *Digitalisierung und Teilhabe: Mitmachen, mitdenken, mitgestalten!* (pp. 97–112). Nomos.

Kranzberg, M. (1986). Technology and History: "Kranzberg's Laws". *Technology and Culture. 27 (3),* 544–560.

Kučera, J., & Bruckner, T. (2019). Blockchain and Ethics: A Brief Overview of the Emerging Initiatives. *CEUR Workshop Proceedings, vol. 2443*, 129–139. https://ceur-ws.org/Vol-2443/paper12.pdf.

Kunkel, N. (2021). Autoregulative Waffensysteme. Automatisierung als Friedensethische Herausforderung – ein Werkstattbericht. *ethikundgesellschaft 2/2021*. https://ethik-und-gesellschaft.de/ojs/index.php/eug/article/view/950.

Kutscha, M. (2006). Innere Sicherheit und Verfassung. In M. Kutscha & F. Roggan (Eds.), *Handbuch zum Recht der Inneren Sicherheit* (2nd ed.) (pp. 24–104). Berliner Wissenschafts-Verlag.

LaPointe, C., & Fishbane, L. (2018). *The Blockchain Ethical Design Framework*. Beeck Center Georgetown. https://beeckcenter.georgetown.edu/wp-content/uploads/2018/06/The-Blockchain-Ethical-Design-Framework.pdf.

LaPointe, C., & Fishbane, L. (2019). The Blockchain Ethical Design Framework. *Technology, Governance, Globalization 12 (3–4)*, 50–71. https://doi.org/10.1162/inov_a_00275

Lee, D., Lim, J., Phoon, K. F., & Wang, Y. (2022), *Applications and Trends in Fintech I. Governance, AI, and Blockchain Design Thinking*. World Scientific Publishing Co.

Lehner, J., Schützeneder, P., & Sametinger, J. (2020). Custom Tokens und Smart Contracts zur Projektsteuerung. In H.-G. Fill & A. Meier (Eds.), *Blockchain: Grundlagen, Anwendungsszenarien und Nutzungspotenziale* (pp. 65–85). Springer.

Mackenroth, G. (2011). *Der Rechtsstaat in der Zwickmühle? Zur Balance von Freiheit und Sicherheit.* Nomos.

Marković, J. (2020). Ethical Foundation of the Blockchain Technology – an Introduction Inquiry. *Synthesis philosophica 70 (2/2020)*, 425–452. https://doi.org/10.21464/sp35209.

Meadows, D. H., Meadows, D. L., Randers, J., Behrens, W. W. (1972). *The Limits to Growth: A report for the CLUB OF ROME'S Project on the Predicament of Mankind*. Universe Books.

Nakamoto, S. (2008, October 31). *Bitcoin: A Peer-to-Peer-Electronic Cash System*. Satoshi Nakamoto Institute. https://nakamotoinstitute.org/bitcoin.

Norris, P. (2001). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge University Press.

OECD (2019). *Is there a role for blockchain in responsible supply chains?* OECD. https://mneguidelines.oecd.org/Is-there-a-role-for-blockchain-in-responsible-supply-chains.pdf.

Peterson, M. (2017). *The Ethics of Technology: A Geometric Analysis of Five Moral Principles.* Oxford University Press.

Pieper, A. (2021). *Denkanstösse zu unseren Sinnfragen*. Schwabe.

Pieper, A. (2014). *Nachgedacht: Philosophische Streifzüge durch unseren Alltag*. Schwabe.

Pilkington, M. (2016). Blockchain Technology: Principles and Applications. In X. Olleros & M. Zhegu (Eds.), *Research Handbook on Digital Transformations* (pp. 225–253). Edward Elgar. https://doi.org/10.4337/9781784717766.00019.

Pfeifer, W. (2012). *Etymologisches Wörterbuch des Deutschen*. Edition Kramer.

Prechtl, P. (2008). Klugheit. In P- Prechtl & F.-P. Burkhard (Eds.), *Metzler Lexikon Philosophie: Begriffe und Definitionen* (3rd ed.) (pp. 293–294). Springer.

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.

Reis, O. (2003). *Nachhaltigkeit – Ethik – Theologie: Eine theologische Beobachtung der Nachhaltigkeitsdebatte*. LIT.

Riedl, A. M. (2022). Technik. In M. Heimbach-Steins, M. Becka, J. J. Frühbauer, & G. Kruip (Eds.), *Christliche Sozialethik: Grundlagen – Kontexte – Themen. Ein Lehr- und Studienbuch* (pp. 280–299). Pustet.

Rieger, A., Guggenmos, F., Lockl, J., Fridgen, G., & Urbach, N. (2019): Building a Blockchain Application that Complies with the EU General Data Protection Regulation. *MIS Quarterly Executive, 18(4)*, 263–279. https://doi.org/10.17705/2msqe.00020.

Ropohl, G. (1979). *Eine Systemtheorie der Technik: Zur Grundlegung der Allgemeinen Technologie*. Hanser.

Sändig, B. (2004). *Albert Camus: Autonomie und Solidarität*. Königshausen & Neumann.

Sarmah, S. S. (2018). Understanding Blockchain Technology. *Computer Science and Engineering 2018, 8(2)*, 23–29. https://researchgate.net/publication/336130918.

Schapp, J. (2006). Die Grenzen der Freiheit. *JuristenZeitung Nr. 61 (12/2006)*, 581–586.

Schönherr-Mann, H.-M. (2018). Albert Camus, Der Mensch in der Revolte (1951). In M. Brocker (Eds.), *Geschichte des politischen Denkens. Das 20. Jahrhundert* (pp. 323–337). Suhrkamp Taschenbuch.

Schultze, R.-O. (2010). Gemeinwohl. In R.-O. Schulze & D. Nohlen (Eds.), *Lexikon der Politikwissenschaften 1* (4th ed.) (pp. 299–302). C. H. Beck.

Schuppert, G. F. (2008). Governance – auf der Suche nach Konturen eines „anerkannt uneindeutigen Begriffs". In G. F. Schuppert & M. Zürn (Eds.), *Governance in einer sich wandelnden Welt* (pp. 14–40). VS Verlag für Sozialwissenschaften.

Schweppenhäuser, G. (2021). *Grundbegriffe der Ethik* (2nd ed.). Reclam.

Seidenfad, K., Wagner, T., Hrestic, R., & Lechner, U. (2022). Demonstrating Feasibility of Blockchain-Driven Carbon Accounting – A Design Study and Demonstrator. In Phillipson, F., Eichler, G., Erfurth, C., & Fahrnberger, G. (Eds.), *Innovations for Community Services. I4CS 2022. Communications in Computer and Information Science, vol 1585* (pp. 28–46). Springer. https://doi.org/10.1007/978-3-031-06668-9_5.

Sharif, M. M., & Ghodoosi, F. (2022). The Ethics of Blockchain in Organizations. *Journal of Business Ethics, vol. 178*, 1009–1025. https://doi.org/10.1007/s10551-022-05058-5.

Smedley, E., Rose, H. J., & Rose, H. J. (1845). *Encyclopaedia Metropolitana; or, Universal Dictionary of Knowledge, On an Original Plan: comprising the twofold advantage of a philosophical and an alphabetical arrangement, with appropriate engravings* XXV. Griffin.

Stegmeier, W. (2008). *Philosophie der Orientierung*. De Gruyter.

Tang, Y., Xiong, J., Becerill-Arreola, R., & Iyer, L. (2020). Ethics of blockchain: A framework of technology, applications, impacts and research directions. *Information Technology & People, vol. 33 no. 2*, 602–632. https://doi.org/10.1108/ITP-10-2018-0491.

Tsai, C., & Lin, C. (2023). Shedding New Light on Multinational Corporations and Human Rights: Promises and Limits of 'Blockchainizing' the Global Supply Chain. *Michigan Journal of International Law, vol. 44, no. 1*, 117–155. https://doi.org/10.36642/mjil.44.1.shedding.

Viana, W. C. (2010). *Das Prinzip Verantwortung von Hans Jonas aus der Perspektive des objektiven Idealismus der Intersubjektivität von Vittorio Hösle*. Königshausen & Neumann.

Virilio, P. (1999). *Politics of the Very Worst: An Interview with Philippe Petit*. S. Lotringer (Eds.), translated by Michael Cavaliere. Semiotext(e).

Voßschmidt, S., & Karsten, A. (2019). *Resilienz und Kritische Infrastrukturen: Aufrechterhaltung von Versorgungsstrukturen im Krisenfall*. Kohlhammer.

Wildt, M. (2013). Eichmann und der Kategorische Imperativ, oder: Gibt es eine nationalsozialistische Moral? In N. Kampe & P. Klein (Eds.), *Die Wannsee-Konferenz am 20. Januar 1942. Dokumente. Forschungsstand. Kontroversen* (pp. 151–166). Böhlau.

Wolf, U. (2010). Über den Sinn der aristotelischen Mesotes-Lehre (II). In O. Höffe (Eds.), *Aristoteles: Nikomachische Ethik* (pp. 83–108). Akademie Verlag.

Designing Sovereign Information Systems

Technology-Enabled Fairness?

Reflections on Fairness within Blockchain-Based Supply Chain Consortia

Isabelle Fries[1] and Maximilian Greiner[2]

Abstract: Can technologies improve social interaction morally? Blockchain is said to lead to greater fairness through transparency and accountability. We explore the extent to which this is true. Using the coffee supply chain as an example, we examine the extent to which the use of blockchain enables fairness in global supply chains. We argue against equating transparency and fairness and cite other factors in favor of fairness, such as laws that legally enforce compliance with human rights. The transdisciplinary approach combines ethics and computer science. Ethical research is based on linguistic findings and John Rawls's concept of *Justice as Fairness*. The governance addressed is also viewed from the technological side. We argue for fairness to be a guiding principle within the governance of blockchain consortia. This relates to trust, legal compliance, decision-making, decision rights, and responsibilities. Empirically, we conducted a literature review as well as interviews with various actors in a specialty coffee supply chain. We conclude that blockchain is a suitable way to give technological form to a societal will for more fairness in global supply chains, but that it can never replace this human will.

Keywords: Blockchain Ethics, Fairness, Governance, Blockchain Consortium, Supply Chain

1 Introduction

How is technology changing society? – This is one of the emerging social questions that is quickly followed by an ethical question: how can technology change society for the better? The field that this opens up is vast. In the following, we focus on a particular technology and consider it in a particular context. We ask to what extent blockchain can contribute to greater fairness in the global coffee supply chain. This implies two assumptions. First, we assume that there is a need for optimization in the existing coffee supply chain in terms of fairness. Second, we assume that the use of blockchain technology is a viable way to contribute to such optimization. The first point is largely based on interviews we conducted with stakeholders along the supply chain. Considering the second one, it is by no means arbitrary for the authors – an ethicist and a computer scientist – to deal with blockchain and fairness in supply chains. This connection can be found in societal expectations (Geißler, 2022) and the marketing of providers of corresponding software solutions (Sopek, 2022). One finds

[1] University of the Bundeswehr Munich, Neubiberg, isabelle.fries@unibw.de
[2] University of the Bundeswehr Munich, Neubiberg, maximilian.greiner@unibw.de
This research paper was first presented at the EURAM 2023 Conference in Dublin (June 14–16, 2023). It is part of the conference proceedings (ISSN 2466-7498). The version now published has been abridged and revised.

it in the research literature as well (Miatton & Amado, 2020), sometimes in an axiomatic way. This is the starting point of investigating the extent to which the use of blockchain in a specific use case can contribute to fairness and the extent to which there are nevertheless permanent limits to this noble desire. Our research question includes three sub-questions: Why should the coffee supply chain be fair at all? What about fairness in conventional processes? Which requirements need to be considered for blockchain-based governance structures to increase fairness?

The above already indicates that the coffee chain as a global supply chain has a particularly high discrepancy in the living conditions of those involved. On the one hand, there are coffee farmers in the country of origin. They are predominantly characterized by high dependence on exporters and precarious living conditions. On the other hand, there are the consumers in the destination country. They enjoy a luxury good that is not essential to life, but has become an everyday commodity. In between, there are stakeholders lined up along the chain with a division of profits that is rarely made in favor of the coffee farmers. Only about 10% of the profits are received by farmers in the "bean belt" due primarily to dependencies and unequal power relations (Samper et al., 2017, p. 3). 40% of the sub-Saharan population lives below the international threshold of one dollar a day (Naudé, 2010, p. 101). Efforts are being made to remedy this state of affairs, which is perceived as a malaise. The motives for this vary. Some have long been committed to fair trade for ethical reasons. Others are confronted with the need to change their existing processes, at the latest with the European Supply Chain Act. An underlying "Proposal for a Directive on corporate sustainability due diligence" was adopted by the European Commission in February 2022. The aim is "to foster sustainable and responsible corporate behavior and to anchor human rights and environmental considerations in companies' operations and corporate governance" (European Commission, 2022). The imposed duty of diligence, which will already apply under national law in Germany, for example, from 2023 (BMWK, 2021), also comes with a documentation requirement, which is what makes blockchain interesting as a technology of transparency from a corporate perspective.

In an interdisciplinary research approach, we take two starting points. In our ethical approach, the topic of governance is understood as one of political philosophy, and fairness is explained in its meaning according to John Rawls. In our technological approach, the characteristics of blockchain consortia will be discussed in more detail. We will identify problems and argue for fairness as a topic of IT governance. We then present the selected use case of the coffee supply chain with reference to interviews conducted. Finally, we form a synthesis of technological approach and ethical approach to make well-founded statements about the potential of blockchain technology in the presented use case and put them up for further discussion.

2 Research Design

Our research goal was to use a specific use case to investigate the extent to which blockchain can contribute to greater fairness compared to the previous process without this technology. We chose the case of a specialty coffee supply chain. The situation

in the food supply chain is particularly well suited for the use of blockchain, as traceability and transparency about processes are required. In addition, the aforementioned supply chain laws ensure that corresponding technologies are also in demand from a corporate perspective. This can be seen in the fact that IBM Food Trust offers corresponding software solutions. Furthermore, coffee is an excellent product for thinking about fairness along the supply chain, because the existing processes are perceived as unfair by the various stakeholders, as will be shown. That this example is predestined to reflect fairness is shown not least by the fact that it has already been considered in the research literature: Miatton and Amado examined fairness in the context of transparency and traceability in the coffee value chain using blockchain (Miatton & Amado, 2020) and introduced the idea of a Commodity Fairness Index used to measure inequality or economic imbalance. Finally, we chose specialty coffee because it is a clearly definable food supply chain that is also small enough to be adequately captured in our study. Investigating fairness in this small industry setting also offers a high potential for generalizing our findings to other agricultural supply chains.

To determine the realistic needs along the coffee supply chain and get a more practice-based idea of fairness, five interviews were conducted between April and July 2022. The interviewees consisted of a coffee farm owner from Vietnam as well as managing directors of roasteries and distributors within the specialty coffee supply chain. Each interview lasted approximately 90 minutes, conducted online using common communication platforms. This allowed us to include the perspective of coffee farmers in a country of origin alongside roasteries or distributors in a destination country. The interview guide is based on the IT governance dimensions of decisions and decision rights, accountabilities, and incentives by Weill (Weill, 2004). One focus was on possible links governance and the use of blockchain (section 4).

The combination of technology and ethics led us to choose an interdisciplinary approach, which is shown below in starting points A and B. From the two complementary professional starting points, supplemented by insights from practice, we finally formed a synthesis to answer the question of the extent to which the use of blockchain in the specific use case is suitable for achieving greater fairness.

3 Starting Point A: Ethical Approach to Fairness

In this section, we first explore the meaning of “fairness” before examining fairness in the context of blockchain. We connect both with an approach of political philosophy according to Rawls and relate it to governance in blockchain consortia.

3.1 The Meaning of Fairness

When asking about “fairness,” we first do so literally. Something that is “fair” has a positive connotation from the start. This is philosophically remarkable, because “fairness” is on the side of the good and desirable from the outset. Something pleasant or

beautiful was once called fair (Harper, 2014). This good was later understood as morally good. The word of Gothic origin "fair" then meant the same as the word of Latin origin "just." This is how it has remained until today. The moral understanding found its way into the field of competition, in sports (with the early counter term "foul"), then also in trade (with the early counter term "contraband") (Weekley, 1967, p. 544). "Fair" refers to a particular type of interaction within a multi-actor system that is considered good and desirable. A person cannot be called "fair" in isolation. What is called fair is a person's action or disposition within a space of action shared with other actors. This space of action can be characterized by fair conditions. If one expects fair conditions, one expects that things should be done honestly and justly. However, this presupposes the conception of something honest and just, over which it is necessary to come to an agreement. The idea of competition in sports and trade also refers to an interaction in which the advantages of one are directly related to the disadvantages of the other. Fairness then means that competition and cooperation are in balance (Copray, 2012, p. 504). Therefore, the notion of fair trade or fair play is associated first with the expectation that benefits should be obtained only in a fair manner, and second that all participants have fair opportunities to obtain those benefits in the long run. In this respect, fairness can also be understood as a "mediating idea," as Fischer puts it (Fischer, 2012, p. 16). Fairness refers to a process or a solution that is supported by the majority of those involved (Fischer, 2012, p. 15).

3.2 Fairness in Blockchain-Based Supply Chains

As supply chains are a system that connects different players, it is not surprising that fairness is also mentioned here. The research literature on fairness concerning supply chains is mainly concerned with fairness in processes and distribution, and how stakeholders interact (Chen et al., 2022, p. 67). To come closer to the ideal of fair cooperation in real interactions, there are rules in trade just as there are in games. If everyone abides by the same rules, at least a formal fairness is fulfilled. Informal fairness would deal with the attitudes of the players or traders (Loland & Court, 2003, pp. xiii f.).

The concept of formal compliance with rules makes it possible to speak of fairness in blockchain. The technology is associated with the promise of a system in which all participants have no choice but to abide by the rules – at least not without being convicted of breaking rules and having to expect punishment. In this respect, it is obvious to even think of a particularly fair socio-technical system. An inevitable conformity to the rules promises more fairness than a card game, in which there are rules but cheating is not necessarily immediately noticeable; or a soccer game, in which the question of whether the rules were still observed is to a certain extent left to the discretion of the arbitrator. But who actually made the rules? And to what extent do these rules also meet the needs and expectations of those who join later, or the needs and expectations of those who do not participate directly in the interaction but are affected by its consequences? So who guarantees that the rules applied are those in favor of fairness? – These questions will lead us to the level of governance in the following.

First, it is noteworthy that blockchain is associated with fairness with the argument that the provided transparency enables a more equitable distribution of profits along a supply chain. However, there is often a gap in this argument when it comes to how exactly transparency enables fairness. The FairChain Foundation simplistically suggests that transparency contributes to a more equitable distribution of wealth (FairChain, 2019). However, the fact that it is visible that farmers receive the smallest share of profits is not a compelling reason for the current beneficiaries to change this distribution. It is therefore important to focus on motives and incentives for fair behavior. Thus, informal fairness must also be thought about beyond technical possibilities. This is also the case with FairChain. The foundation wants to replace development aid with sustainable consumption. This motive differs from a profit-oriented one, which may also be interested in fairness, but rather according to the motto "fairness pays off" (Amesberger, 2015, p. 8). Fairness could pay off. At least, that is what consumers said when asked about the criteria in their purchasing decisions (Butera, 2011).

There is a strong case for adopting fairness as a guiding principle within the supply chain, particularly the global coffee supply chain. We also assume that this guiding principle should be at the core of the IT governance of blockchain supply chain consortia. Before we elaborate on these ideas, we will first take a look at Rawls's concept of fairness. We do so not only because within ethics, the topic of fairness is linked to his *Theory of Justice* (Rawls, 1971), but precisely because his anchoring of the idea of fairness on governance level is also instructive for the design of IT governance.

3.3 Justice as Fairness in a Blockchain Consortium

The following considerations are based on Rawls's concept of *Justice as Fairness* (Rawls, 2001). However, it is not necessary to be already familiar with this concept or to know it in detail. The most important points for our purpose will be mentioned. We consider his theory under the question of the extent to which it can be applied to blockchain consortia. To do so, it is helpful to consider Rawls's presuppositions.

First, Rawls notes a plural starting point. Some doctrines make statements about what is just, but these doctrines do not always agree. This results in different conceptions of justice. The second assumption lies in the democratic constitution of the society in which Rawls finds himself. Consequently, the answer to the question of what is just lies neither with a tyrant nor with oligarchs, but with free and equal citizens. Both together raise Rawls's initial question: What conception of justice can free and equal citizens agree upon despite differing views? Rawls's answer was that it is crucial that the process of agreeing on certain principles itself takes place under fair conditions. The concept he was looking for was the concept of *Justice as Fairness*.

Anyone who is asked to design IT governance is faced with the same basic problem as Rawls. There are a lot of requirements. They do not seem compatible, but need to be mapped with one concept. If a requirement is fairness, the designer must know what is meant by fairness. There are stakeholders with different definitions. Should the designer now simply go along with an idea that is personally meaningful? Or is there an authority that makes a hierarchical decision? A designer seeking to follow

Rawls will look for a democratic consensus (cf. §11 *The Idea of an Overlapping Consensus* | Rawls, 2011). Then it will probably occur to the designer that blockchain could be a suitable technology. After all, it has been written again and again, especially in the early days of the technology, that blockchain would bring democratic conditions and could reduce dependencies on monopolists. Reference was often made to technical decentralization, which was equated with decentralization of power. We advise against equating a technical possibility with a social reality. Magnuson has noticed how quickly social hopes faded in the face of technical realities (Magnuson, 2020, p. 90). Hermstrüwer also criticized any hasty equation of decentralization and democratization and preferred to look for concrete democratic design options (Hermstrüwer, 2019). We agree with this line of thinking and assume that the blockchain consortium is a good technical basis for negotiating interests; not because data storage is decentralized, but because it is technically possible to involve different stakeholders equally in a decision-making process. We also assume that it will be possible in this way to approach fairness understood as procedural justice enabled by technology.

One could say our designer is where Rawls's fictional protagonists are in his thought experiment of the *Original Position* (cf. §6 *The Idea of the Original Position* | Rawls, 2001). A designer seeking to think about requirements in favor of fairness within a coffee supply chain, would have to think of representatives of all stakeholders. They should not know whether they will be coffee farmers or consumers in the end. Now they should agree on which principles of fairness should apply within the blockchain consortium. Assuming the same initial conditions as in Rawls's experiment, the exemplary protagonists of the blockchain consortia should come to the same conclusion: justice means equal access to the same system of freedoms. Inequalities are only allowed if everyone has a fair chance to reach a better position and if even the least advantaged are still better off under the inequalities than under conceivable alternatives (cf. §13 *Two Principles of Justice*). Just as Rawls sees society as a fair system of cooperation (cf. §2 *Society as a Fair System of Cooperation*), the protagonists of the supply chain are likely to see the blockchain consortium as just such a system. They, too, will seek a model of cooperation that is not only fair, but also beneficial to all.

The question of how realistic the thought experiment is, is futile but was often asked, not least during Rawls's lifetime. From our point of view, it raises awareness of the problem of defining criteria within IT governance, especially if it is not to be hierarchically prescribed, but democratic and, above all, fair. At the same time, the undeniable weaknesses of a thought experiment show how important it is to know the real interests of the various participants. All designers, like all ethicists, must come up against human limitations when they try to abstract from themselves and imagine, for example, what fairness means to a coffee farmer in Vietnam. Not least for this reason, we talked to stakeholders and experts to elicit a realistic understanding of fairness and unfairness in the specialty coffee supply chain we chose as an example. We will present the results later. In the evaluation, we will also come back to Rawls and ask to what extent a blockchain consortium is comparable to Rawls's concept of the well-ordered society based on cooperation (cf. §2 *Society as a Fair System of Cooperation*). First, however, we return to starting point B, asking about fairness from the computer scientist's point of view.

4 Starting Point B: Blockchain's Promise of Fairness

Within this section, we begin by outlining blockchain technology including its technical background as well as its potential. Furthermore, the current state of research on blockchain governance about fairness is discussed.

4.1 Blockchain as a Disruptive Technology

Blockchain has a number of distinctive features that set it apart from other technologies. Key characteristics include decentralization, immutability, security, and transparency (Sultan et al., 2018). A blockchain is a digital accounting system used to track and secure transactions. It consists of blocks of transactions that are chained together to form a record of transactions. Each block contains a record of transactions, a timestamp, and a reference to the previous block. The blocks are linked together using a cryptographic function that makes it impossible to subsequently change a block without also changing all subsequent blocks. That is why blockchain is considered to be very secure. Lu also points to trustworthiness (Lu, 2018). However, this is misleading, insofar as a technology can neither be compared to nor replace a trusted person (Fries, 2022). What blockchain does is provide security through the visibility of data and the aforementioned cryptography, which can take the previous place of trust.

There are several types of blockchain technologies that differ in their architecture, purpose, and scope of use. Public blockchains are accessible to anyone and allow any user to perform transactions and become part of the network. Examples of public blockchains are Bitcoin and Ethereum (Sheth & Dattani, 2019). Private blockchains are reserved for specific individuals or organizations with permitted access. They are often used in companies and organizations to optimize and accelerate internal processes. Consortium blockchains are operated by a group of companies or organizations and are not accessible to the public. They are often used to improve processes in industries where multiple companies collaborate, such as the financial or supply chain sector (Dib et al., 2018). Lastly, hybrid blockchains combine elements of public and private blockchains. They offer the flexibility and adaptability of private blockchains, and the security and transparency of public blockchains (Alkhateeb et al., 2022).

The functionality in combination with the various properties and types described opens great potential for different fields of application. Especially in the supply chain, transparency can be improved. By using blockchain technology, all parties involved can see at any time where a product is in the delivery process and who handled it beforehand. Furthermore, faster processing of transactions is possible as no manual processing steps are required. The immutable records on the blockchain can help prevent fraud because every transaction is recorded immutably. Last, it enables all participants to collaborate and exchange information on a common platform (Chang & Chen, 2020). In terms of ethics, blockchain is considered as an enabler for improved compliance. On the one hand, the technology offers the technical potential to document, track, and prevent the use of child labor and other violations of labor standards. The blockchain can also support compliance with other human rights. Accurate data

documentation and its traceability is promised to ensure that no practices that violate human rights are taking place along the supply chain. In addition, companies can prove and track compliance with environmental standards, e.g., by tracking CO2 emissions along the supply chain. On the other hand, the technology can also help companies track their compliance with safety standards in the supply chain and ensure that no dangerous or harmful products are produced or distributed (Hyrynsalmi et al., 2019).

To fully realize the potential, a solid governance structure is necessary that regulates the use of blockchain and ensures that the technology is used ethically.

4.2 Blockchain Governance towards Fairness

Governance generally refers to the rules, procedures, and mechanisms that determine how an organization or system is directed, controlled, and managed. It encompasses the full range of decision-making processes and structures required to manage and control an organization or system. The governance goal is to achieve the goals of the organization or system and to protect the interests of the stakeholders (Benz, 2004).

Governance has an important role in promoting fairness in an organization or system. Fairness refers to ensuring that all stakeholders are treated equally and that decisions are made based on justice and equity. An important element of governance is adherence to rules and procedures that ensure decisions are made in a transparent and accountable manner. This strengthens the reliability of governance structures. To ensure fairness in governance, it is also important that all stakeholders have access to information and have the opportunity to represent their interests. This can be achieved through mechanisms such as open decision-making processes, stakeholder participation, and embracing diversity and inclusion (Baker et al., 2016).

The starting point for successful governance is the different application areas and business models as well as the specific challenges that blockchain technology implies. A first approach to address this issue is proposed by Beck et al. The authors discuss the dimensions (decision rights, responsibilities, incentives) of IT governance defined by Weill along the blockchain economy and propose key questions of the dimensions in the form of a research agenda (Weill 2004; Beck et al., 2018). So far, only a few scientists have made reference to ethics. Hofman et al. take up Beck's agenda and create a governance analysis framework that attempts to capture the embeddedness of blockchain solutions in the broader world. They relate existing power structures (legal, political, environmental, social) to the 5W1H method and describe this as follows: Who? *(actors and stakeholders)*, What? *(data, records, and logs)*, Why? *(use cases and added values)*, When? *(temporality and change over time)*, Where? *(geography of instantiation)*, How? *(instantiation)* (Hofman et al., 2021). This is the first use of the term "social" in connection with blockchain governance. Yue et al. propose a framework that consists of six governance attributes and 13 sub-attributes (Yue et al., 2021). The governance attributes include decision-making processes, accountability and verifiability, privacy and security, trust, incentives, and effectiveness. From their perspective, successful implementation of governance measures can be achieved from

a mix of two ways: organizational mechanisms and human interaction. Thereby, "trust" is associated with fairness, as effective governance is based on participants trusting that decisions are fair, properly executed, privacy-protected, and highly tamper-resistant, with decision policies that are transparent and accessible (Yue et al., 2021). The authors also consider ethical issues including six principles. Fairness is mentioned within two principles: First, blockchain governance should enable transparent decision processes to have insights into reasonableness and traceability. Second, successful governance should manage legal compliance and ethical responsibility. This should ensure that all governance-related decisions and processes conform to legal regulations and ethical responsibilities (Liu et al., 2021). Anthony Jnr. refers to the structures and concepts of governance in companies, describing them in relation to the blockchain. In addition to economic, technological, and political factors, the social factor is also mentioned. Fairness here is mainly characterized by the distribution of decision rights and a clear structure in decision-making (Anthony Jnr., 2022).

Overall, the key aspects that are addressed within blockchain governance towards fairness are trust, legal compliance, decision-making, decision rights, and accountability. Here, the central question now arises: how can fairness be established through trust, decisions, and decision rights, as well as responsibilities within a consortium?

5 Case Study: Towards Fairness in a Coffee Supply Chain

As described above, interviews were conducted based on Weill's IT governance dimensions: incentives, decisions, and decision rights, as well as accountabilities (Weill, 2004). Ethical issues were addressed in each case. To describe how the interviewees envision fairness in a consortium, the current situation and then the opportunities and challenges in implementing fairness in a consortium are outlined.

5.1 Current State of Fairness

Considering fairness, the interview partners clearly state that there are major deficits even in the specialty coffee supply chain.

Starting from the perspective of the farmers, unfairness can be viewed from different angles. In complicated years, such as those of the COVID-19 pandemic, coffee farmers are on their own. To a large extent, there is no support from the other stakeholders within the supply chain, which is also evidenced by several reviews (Aday & Aday 2020; de la Peña García et al., 2020). While other companies in consumer countries receive government support, farmers must rely on revenues from previous years, which is not always a given. Although coffee is considered one of the best-selling products and "trails only oil in global trade volume," as Smith notes (Smith, 2013, p. 163), the farmers, who grow the coffee and are largely responsible for the final product, belong to the social underclass. This is also accompanied by a lack of understanding and use of technology in producer countries. Technologies enable data and information visibility that can support and improve a company's strategic goals. Based on

the missing data insights, the farmers lack the know-how how to improve quality, as it is difficult for them to get feedback about their product. Furthermore, the farmers often have few educational opportunities, which in turn leads to a lack of knowledge in how to allocate funds and how to invest in innovative machines or support tools. Last, the interview partners recognized a problem considering the pricing of the coffee along the value chain. Farmers have to sell a certain amount of coffee to keep a minimum standard of living. This results in limited co-determination rights in pricing, as they are dependent on the quantity purchased by buyers.

From the perspective of roasteries and distributors, the interviewees agree, that the difficult communications between farmers and other stakeholders support unfairness. Thus, sellers in consumer countries rely on information from intermediaries. Littler or no information is available about the working conditions and social circumstances. This often leads to greenwashing and an exchange of information that trivializes the social circumstances in the country of origin. Although this is not as common in specialty coffee as in industrial coffee, it is a major problem.

The results show that distributors or roasters who are in direct contact with customers feel a great responsibility towards farmers. However, implementing the improvement in living conditions is not simple, as the customer must be enlightened about why the coffee is more expensive, even though no improved quality is presented in the short term. These aspects lead to injustice and unfairness towards farmers. The question arises: Which issues could a blockchain-based consortium address and counteract to increase fairness towards farmers as well as other supply chain stakeholders?

5.2 Blockchain Potentials and Challenges towards Fairness

The potential of blockchain has already been addressed in section 4. If we add the interview findings, we can focus on the question of the extent to which the use of technology can improve the exchange and transparency of information within the coffee supply chain. The fairness to be improved, then, consists first of fair access to information. Improved fairness in terms of access to information benefits not only coffee farmers in the growing countries, but also traders, roasters, and customers. Farmers recognize the value of their product in the consumer countries and can thus improve their negotiating position. Traders can increase efficiency and save time and storage costs. Roasters have better arguments for their price calculations vis-à-vis customers. Customers can check whether their investments are investments in fair trade.

The interviews show that the long-term relationships enabled by a blockchain-based consortium in particular could improve both quality and fairness. The secondary fairness goes beyond information fairness and can extend into improving living conditions in the country of origin. The blockchain consortium could counter opportunistic behavior by participants, leading to a fairer environment for all stakeholders by strengthening reliability among participants. As all stakeholders, including the customer, are involved in a possible consortium, communication can be improved. A fair and sustainably sourced product leads to strengthened marketing for distributors, creating a unique selling point. Based on this transparency, the coffee farmers could be

involved in the coffee pricing and long-term contracts could be concluded to enable prepayment processes for the farmers. This leads to improved quality of the initial product. Farmers can invest in machines and further technological innovations to automate their work, and thus enhance fairness. The last potential is about consulting strategies regarding technological and organizational perspectives. Companies from consumer countries could support the farmers and simultaneously improve their own product and marketing strategies. Consulting could take place considering the allocation of funds and investments as well as in the form of social projects.

As outlined above, blockchain provides many potentials, but there are also technological as well as organizational challenges for improving fairness within a consortium. As the interviewees state, on the one hand the implementation of blockchain at the beginning of the chain would be a major challenge. This relates primarily to infrastructure, education, regulation, acceptance, and financing. In many developing countries, the technical prerequisites for using blockchain may not exist or may be insufficiently developed. These include, for example, stable power supply, Internet access, and hardware. To this end, there may be a lack of qualified professionals capable of implementing and managing blockchain technology. There is also a lack of regulatory mechanisms or legal frameworks that support or enable the use of the technology in these countries. Due to the lack of understanding, acceptance of the technologies also suffers. Lastly, integrating blockchain technology in developing countries may require significant investments that may not be available or may be difficult to obtain.

6 Conclusion: Blockchain's Contribution to Greater Fairness in a Coffee Supply Chain

To what extent can blockchain technology enable fairness in the coffee supply chain? It is noteworthy that the term chosen is "enable." No technology is a panacea that, once implemented, will eliminate social ills. We should not expect that from blockchain either, although the media sometimes gives a different impression. We are dealing with hopes. However, it is important to distinguish between technical possibilities and exaggerated expectations if we want to sensitize people to their responsibility.

The starting point in philosophy shows that fairness does not have to be understood as the utopia of philanthropic moralists. According to Rawls, fairness is a rational behavior of people in social interaction. If you behave reasonably, you behave fairly, we can take from this. This philosophical insight, which is known to be found in various game theories (Laden, 1991), coincides with statements we have gained from the practice within the coffee supply chain. The roasters also benefit if the farmers have better living and working conditions. Perhaps intermediaries would choose differently because they may not directly benefit to the same extent: this is where applicable laws come into play. With the EU Supply Chain Act, it is worthwhile to prove that all stakeholders are acting in a fair and sustainable manner so as not to risk penalties.

The question of why one should behave fairly at all, which is a variation on the question "Why be moral?" has found three possible answers. The first already confronted

us linguistically: fairness as the good is understood, at least by many people, also as the desirable, which they can intrinsically want to achieve. The other possibility can be described as rational or game-theoretical, insofar as one behaves fairly because it also increases one's advantages. A third possibility comes into play with laws and fear of punishment. In the literature, fairness is also referred to as human rights (Amesberger, 2015, p. 7). This implies that there should be structures to enforce fairness; also in the economic sector and thus also in the supply chain. One technical-structural basis can be seen in blockchain technology. This is true irrespective of the challenges mentioned in section 5. To overcome these challenges, it can be helpful to work closely with local communities and affected stakeholders to address their needs. It may also be helpful to seek support from international organizations and development banks to promote the integration of the technology in developing countries.

One difference between a blockchain consortium and Rawls's well-ordered society is the respective scope. Rawls thought of a democratically constituted state with a binding constitution in a congruent legal space. The coffee supply chain as a global supply chain moves across national and legal borders. The same is true for the blockchain consortium. The consortium will not adopt a state constitution and the corresponding sanction options will also remain absent. However, if blockchain is used as a common technical basis, this can be compensated for to a certain extent. Even if the blockchain consortium is more similar to an association than to a state, it is still possible to identify similar mechanisms in regulation, which are reflected in IT governance. Governance ethics is addressed in the context of business and virtue ethics (Wieland, 2006). Schramm speaks of a "management of moral interests" (Schramm, 2017). We have seen that concepts of IT governance increasingly also refer to ethical foundations. Although "fairness" mostly remains fuzzy in the still-new field of blockchain governance, it is mentioned in connection with business ethics and responsibility. It should be recalled that governance is in line with the corporate goals. Whether a company is advocating fairness out of conviction, improving efficiency, or complying with the law, it will come back to fairness. Blockchain in the coffee supply chain combines aspects of fair trade, fair governance, and fair participation. The benefits are not only for the farmers, but also for the participants within a blockchain consortium in consumer countries, who can derive considerable added value from increased fairness. The fact that fairness is also seen as contributing to long-term economic success (Fröhlich & Glaner, 2007, p. 40) makes it so appealing for company use, even for non-ethically argued reasons. Whether out of intrinsic interest, due to legal requirements, or because this is currently demanded and rewarded by customers (Auger et al., 2003), companies themselves are committed to a transformation of business for good.

The importance of transparency regarding processes and data is repeatedly emphasized. This is evident both in the research literature and in the conversations with the interview partners presented above. More transparency is often expected to lead to more fairness as if it were a matter of course. Via the link of the transparency idea, one then very quickly comes to the assumption that blockchain as a technology of transparency automatically leads to more fairness. We have shown why we disagree with this equation. Still, there are good reasons to associate blockchain with fairness. It is important to distinguish two dimensions or scopes of addressed fairness here.

The first refers to the technical possibility of fair participation. This concerns the equal representation of different stakeholders as equally empowered members of the blockchain consortium. This can be associated with equally strong decision-making rights and with the technical possibility to achieve a fair balance of interests, which also includes fair access to information. Whether these technical possibilities find technical realizations, however, stands and falls with the leading governance that supports this way of fairness or not. The same applies here to the promise of democratization that once accompanied the emergence of blockchain. Just like democracy, fairness is not a self-propelling process that technology would simply set in motion. Technology can support social change processes, but it is still people who determine the course of these changes. What we can also observe is that there is currently very much a societal will towards more fairness and sustainability, especially in the supply chain. Under democratic conditions, corresponding laws that are just emerging in the EU must be seen as a reaction to this societal will, and not vice versa. Blockchain does not replace this will, but it can be seen as a very good way to give this will a technical basis for implementation. If this sequence is adhered to, Thomason et al. rightly speak of blockchain as "a game changer for the poor" (Thomason et al., 2018, p. 138).

The second dimension of addressed fairness points beyond technical reality. Here, it is about the extent to which the use of technology can also positively influence social circumstances, for example about better wages and better working conditions. Respondents indicated that greater fairness also translates into better product quality, which in turn leads to expectations of long-term economic relationships and better prices. It may indeed be the case that transparency over pricing for all parties involved can also indirectly influence the distribution of profits. This assumption is suggested by the interviews conducted. The customers, who can also be participants in an affiliated public blockchain, can exert particular influence here. If they see, through the accurate data documentation, that the additional price actually arrives in the country of production, and if they even see which farmer exactly receives this additional price, they may also be more willing to pay it. A study has shown that even then the majority of them will not pay any price, continuing to act according to their own interests, but they are more willing to pay more than before (Degli Antoni & Faillo, 2022).

It is important to point out a persistent difficulty that limits the ability of technology to change the extra-technological world. In the blockchain literature, reference is made to the "oracle problem," which concerns interfaces between the digital and analog worlds. Applied to the coffee supply chain, this means that the customer in the country of consumption cannot be 100% sure that no child labor took place, even if the data fed into the blockchain is supposed to prove exactly that. It is still humans who make these inputs and are in turn verified by other humans in their workflows. So the reality of data, even with the best technology, is a human-driven reality (Fries, 2023). However, it is still the case that blockchain is a technology that makes it much harder for false data not to be seen. It also makes corruption much more difficult. The technical structure also makes it particularly well suited to making processes fairer, as we have seen. And yet, lastly, it must be emphasized once again that blockchain will never by itself lead to specific ethical goals. It is always based on the human will to change the social world for the better, which can be enabled by technology.

Acknowledgments

This research paper is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU. We also thank the LIONS consortium and our research partners for their contributions to this research activity.

Authors have been listed in alphabetical order and have contributed as follows: Isabelle Fries took the ethical perspective. She wrote the introduction, the research description, the first ethical approach, and the conclusion, and ensured overall coherence. Maximilian Greiner wrote the second approach as a computer scientist. He conducted the interviews, evaluated them, and wrote the case study section.

References

Aday, S., & Aday, M. S. (2020). Impact of COVID-19 on the food supply chain. *Food Quality and Safety, 4(4)*, 167–180.

Alkhateeb, A., Catal, C., Kar, G., & Mishra, A. (2022). Hybrid Blockchain Platforms for the Internet of Things (IoT): A Systematic Literature Review. *Sensors, 22(4)*, 1304.

Amesberger, G. (2015). Vorwort I. In M. Dimitriou & G. Schweiger (Eds.), *Fairness und Fairplay. Interdisziplinäre Perspektiven* (pp. 7–9). Springer.

Anthony Jnr., B. (2022). *Toward a collaborative governance model for distributed ledger technology adoption in organizations. Environment Systems and Decisions*. https://doi.org/10.1007/s10669-022-09852-4.

Auger, P., Burke, P., Devinney, T. M., & Louviere, J. J. (2003). What Will Consumers Pay for Social Product Features? *Journal of Business Ethics, 42(3)*, 281–304.

Baker, D. B., Kaye, J., & Terry, S. F. (2016). Governance Through Privacy, Fairness, and Respect for Individuals. *EGEMs, 4(2)*, 1207.

Beck, R., Müller-Bloch, C., & King, J. L. (2018). Governance in the Blockchain Economy: A Framework and Research Agenda. *Journal of the Association for Information Systems*, 1020–1034.

Benz, A. (2004). Governance – Modebegriff oder nützliches sozialwissenschaftliches Konzept? In A. Benz (Eds.), *Governance – Regieren in komplexen Regelsystemen: Eine Einführung* (pp. 11–28). VS Verlag für Sozialwissenschaften.

BMWK (2021). *Lieferkettensorgfaltspflichtengesetz*. https:// bmwk.de/Redaktion/DE/Gesetze/Wirtschaft/lieferkettensorgfaltspflichtengesetz.html.

Butera, J. (2011). *Fairness-Barometer*, Powered by Fairness-Stiftung. https://www.fairness-barometer.de.

Chang, S. E., & Chen, Y. (2020). When Blockchain Meets Supply Chain: A Systematic Literature Review on Current Development and Potential Applications. *IEEE Access, 8: 62478–62494*. Presented at the IEEE Access.

Chen, L., Lee, H., & Tang, C. S. (2022). Fair Price, Fair Trade, and Fair Pay in Supply Chains. In H. Lee, R. Ernst, A. Huchzermeier, S. Cui (Eds.), *Creating Values with Operations*

and Analytics: A Tribute to the Contributions of Professor Morris Cohen (pp. 65–81). Springer.

Copray, N. (2012). Fairness als spiritueller Impuls im Unternehmen. In H. Schoenauer (Eds.), *Spiritualität und innovative Unternehmensführung* (pp. 494–507). Kohlhammer.

de la Peña García, A., Zimmermann, S. A., & Eleuterio, A. A. (2020). Food Supply Chains, Family Farming, and Food Policies under the COVID-19 Pandemic in a Brazilian City. *Human Organization, 79(4)*, 323–332.

Degli Antoni, G., & Faillo, M. (2022). Ethical consumerism and wage levels: Evidence from an experimental market. *Business Ethics, the Environment & Responsibility, 31(3)*, 875–887. https://doi.org/10.1111/beer.12447.

Dib, O., Brousmiche, K.-L., Durand, A., Thea, E., & Hamida, E. B. (2018). *Consortium Blockchains: Overview, Applications and Challenges.*

European Commission, E. (2022). *Corporate sustainability due diligence*. https://commission.europa.eu/business-economy-euro/doing-business-eu/corporate-sustainability-due-diligence_en.

FairChain (2019). *FairChain Foundation – Returning production and profit to the countries of origin*. https://fairchain.org.

Fischer, D. H. (2012). *Fairness and Freedom: A History of Two Open Societies: New Zealand and the United States*. Oxford University Press.

Fries, I. (2023). Daten, Freiheit, Wirklichkeit. Deutungshoheit in Zeit und digitalem Raum. In A. Schaffer, E. Lang, & S. Hartard (Eds.), *Wirklichkeitskonstruktionen* (pp. 159–183). Metropolis.

Fries, I. (2022). "In Code We Trust?" – Zur Vertrauens-Verheißung der Blockchain-Technologie. *ZEE 2022(4)*, 264–276. https://doi.org/10.14315/zee-2022-660405.

Fröhlich, M., & Glasner, K. (Eds.) (2007). *IT Governance: Leitfaden für eine praxisgerechte Implementierung*. Gabler.

Geißler, O. (2022). *Mehr Fairness für die Medien dank Blockchain*. https://www.dev-insider.de/mehr-fairness-fuer-die-medien-dank-blockchain-a-093cba540d77ad3ea39fadec1dee8e66.

Harper, D. (2014). *Fairness | Etymology, origin and meaning of fairness by etymonline*. https://www.etymonline.com/word/fairness.

Hermstrüwer, Y. (2019). Democratic blockchain design. J*ournal of Institutional and Theoretical Economics, 175(1)*, 163–177.

Hofman, D., DuPont, Q., Walch, A., & Beschastnikh, I. (2021). Blockchain Governance: De Facto (x)or Designed? In V. L. Lemieux & C. Feng (Eds.), *Building Decentralized Trust* (pp. 21–33). Springer.

Hyrynsalmi, S., Hyrynsalmi, S. M., & Kimppa, K. K. (2020). Blockchain Ethics: A Systematic Literature Review of Blockchain Research. In M. Cacace, R. Halonen, H. Li, T. P. Orrensalo, C. Li, et al. (Eds.), *Well-Being in the Information Society. Fruits of Respect* (pp. 145–155). Springer.

Laden, A. (1991). Games, Fairness, and Rawls's A Theory of Justice. P*hilosophy & Public Affairs, 20(3)*, 189–222.

Liu, Y., Lu, Q., Paik, H.-Y., & Zhu, L. (2021). *Defining Blockchain Governance Principles: A Comprehensive Framework*. ArXiv:2110.13374 [Cs]. http://arxiv.org/abs/2110.13374.

Liu, Y., Lu, Q., Zhu, L., Paik, H.-Y., & Staples, M. (2021). *A Systematic Literature Review on Blockchain Governance*. ArXiv:2105.05460 [Cs]. http://arxiv.org/abs/2105.05460.

Loland, S., & Court, J. (2003). Fair Play in Sport: A Moral Norm System. *German Journal of Exercise and Sport Research, 33(1)*, 90–92.

Lu, Y. (2018). Blockchain: A Survey on Functions, Applications and Open Issues. *Journal of Industrial Integration and Management, 03(04)*. https://doi.org/10.1142/S2424862218500015X.

Magnuson, W. (2020). *Blockchain Democracy: Technology, Law and the Rule of the Crowd*. Cambridge University Press.

Miatton, F., & Amado, L. (2020). Fairness, Transparency and Traceability in the Coffee Value Chain through Blockchain Innovation. *2020 International Conference on Technology and Entrepreneurship – Virtual (ICTE-V)*. https://doi.org./10.1109/ICTE-V50708.2020.9113785.

Naudé, P. J. (2010). Fair Global Trade: A Perspective from Africa. In G. Moore (Eds.), *Fairness in International Trade* (pp. 99–115). Springer Netherlands.

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.

Rawls, J. (2001). *Justice as Fairness: A Restatement*. Harvard University Press.

Samper, L. F., Giovannucci, D., & Vieira, L. M. (2017). *The powerful role of intangibles in the coffee value chain*. https://doi.org/10.34667/tind.29021.

Schramm, M. (2017). Das Management moralischer Interessen: Zur Praxisrelevanz von Tugenden in der Wirtschafts-und Unternehmensethik. In J. Wieland (Eds.), *Die Tugend der Governance* (pp. 5–82). Metropolis.

Sheth, H., & Dattani, J. (2019). Overview of Blockchain Technology. *Asian Journal For Convergence In Technology (AJCT) ISSN-2350-1146*. https://asianssr.org/index.php/ajct/article/view/728.

Smith, D. (2013). Ethical Product Sourcing in the Starbucks Coffee Supply Chain. In C. Munson (Eds.), *The Supply Chain Management Casebook: Comprehensive Coverage and Best Practices in SCM* (pp. 163–170). FT Press.

Sopek, M. (2022). *Products, Services and Research. Quantum Blockchains.* https://www.quantumblockchains.io/products-services.

Sultan, K., Ruhi, U., & Lakhani, R. (2018, June 10). *Conceptualizing Blockchains: Characteristics & Applications*. arXiv. https://doi.org/10.48550/arXiv.1806.03693.

Tang, Y., Xiong, J., Becerril-Arreola, R., & Iyer, L. (2019). Ethics of blockchain: A framework of technology, applications, impacts, and research directions. *Information Technology & People, 33(2)*, 602–632.

Thomason, J., Ahmad, M., Bronder, P., Hoyt, E., et al. (2018). Blockchain – Powering and empowering the poor in developing countries. In A. Marke (Eds.), *Transforming climate finance and green investment with blockchains* (pp. 137–152). Elsevier.

Weekley, E. (1967). A*n etymological dictionary of modern English*. Vol. I. Dover Publications.

Weill, P. (2004). Don't Just Lead, Govern: How Top-Performing Firms Govern IT. *MIS Quarterly Executive, 3*, 1–17.

Wieland, J. (2006). *Die Tugend der Governance*. Metropolis.

Yue, K.-B., Kallempudi, P., Sha, K., Wei, W., & Liu, X. (2021). *Governance Attributes of Consortium Blockchain Applications*.

Towards a Governance Model Design for Traceability Systems

Maximilian Greiner[1], Christian Zeiß[2], Ulrike Lechner[3], and Axel Winkelmann[4]

Abstract: Lack of social fairness and increasing legal and regulatory obligations for traceability along the supply chain cause companies to face complex challenges. As a promising technology for supply chains, blockchain has the potential to address these challenges. This research paper focuses on governance for resilience in information systems for supply chain consortia. As instantiation, we aim to develop a governance model for blockchain-based traceability systems in supply chain consortia within an agricultural environment. To set a foundation and narrow down the research interests, in this article, we utilized a design science research approach to elicit seven tentative design principles (DP) for an agricultural supply chain consortia governance model using blockchain-based traceability systems. Drawing on existing literature and expert interviews requirements, we identified the design principles of data, legislation and regulation, roles and responsibilities, decisions, decisions rights, decision management, system as a service, social - communication and fairness, as well as an incentive system. The elaborated DP can be used as a foundation for researchers and practitioners to design a governance model, including roles, rules, incentives, structures, and processes with associated possible alternatives. This research paper was first presented at the 18th International Conference on Design Science Research in Information Systems and Technology in Pretoria, South Africa (May 31st–June 2nd, 2023). It is not part of the conference proceedings.

Keywords: Governance, Blockchain, Supply Chain, Traceability, Resilience, Design Principles

1 Motivation and Problem Definition

1.1 Motivation

In today's globalized world, supply chains face various challenges impacting their efficiency, effectiveness, and sustainability that must be overcome (Sunyaev et al., 2021). Considered one of the most critical supply chains is the agricultural supply chain (Kharas, McArthur, & Ohno, 2022). In particular, agricultural supply chains face challenges including the lack of transparency driven by difficult communication among stakeholders because of language barriers. In addition, especially in agricultural supply chains with large social imbalances between actors, greenwashing is widespread to create an environmentally friendly and responsible image for customers. Furthermore, social, cultural, and technological gaps between growing and consuming countries lead to exploitation and fraud (Abbasi, 2017). To counteract this, increasing regulatory changes are demanding accountability from companies to improve environmental,

1 University of the Bundeswehr Munich, Neubiberg, maximilian.greiner@unibw.de
2 University of Würzburg, Würzburg, christian.zeiss@uni-wuerzburg.de
3 University of the Bundeswehr Munich, Neubiberg, ulrike.lechner@unibw.de
4 University of Würzburg, Würzburg, axel.winkelmann@uni-wuerzburg.de

human, and child rights protections along global supply chains. In addition, companies lack guidelines on how to prepare for adopting recent legislation.

An information system in the form of a traceability system can provide assistance in uniquely identifying physical goods, documenting transactions, and storing states and environments in attributes while ensuring compliance (Tian, 2017). However, existing systems using common databases neglect aspects such as fairness, trust, and intercultural boundaries (Bosona & Gebresenbet, 2013). To address these issues, traditional traceability systems are extended by new technologies. For example, recent approaches integrate blockchain within these traceability systems from a technical view (Tian, 2017), which still provides a lack of organizational perspectives. Blockchain, a growing technology, has gained significant attention from companies and researchers. In addition to sovereignty, this technology provides the potential to increase efficiency, resilience, security, fairness, and transparency (Gurtu & Johny, 2019). Considering the challenges above, one possibility would be in implementing blockchain-based traceability systems (Chang & Chen, 2020). Research in business models and technical implementations of this technology within supply chains is matured to the point that the organizational perspective for operating a blockchain should be included. Based on the blockchain operation categories, it now requires governance structures and processes to enable transparent and resilient control and regulation for traceability systems within the development, operation, and evolvement of blockchain-based consortia (Hoiss, Seidenfad, & Lechner, 2021).

1.2 Research Problem and Objective

However, research and practice indicate a gap considering organizational governance issues within supply chain consortia in information systems.

As an instantiation, we focus on a specific environment, the agricultural supply chain. There is a lack of understanding of the requirements for designing a governance model for blockchain-based traceability systems in agricultural supply chain consortia. Furthermore, there is a lack of governance guidelines on developing, operating, and evolving blockchain-based consortia regarding incentives, decisions, decision rights, and accountabilities.

We approach the problem from two sides. On the one hand, we consider the research of trust and collaboration through blockchain-based traceability systems in the supply chain domain. On the other hand, we investigate the inter-organizational governance perspective of consortia using the blockchain life cycle and strengthening resilience in supply chains.

To contribute to this research gap, our common interest lies in the governance of blockchain platforms within industrial application fields in supply chains to provide a foundation for further individual research. We identified missing design knowledge about governance providing resilience in a blockchain-based traceability system within supply

chain consortia. Therefore, we want to identify design requirements, design principles, and design features for a governance model addressing resilience in blockchain-based traceability systems in inter-organizational supply chain consortia. Our primary focus is the business value proposition for all participants, including customers and stakeholders.

This approach leads to the following overall research question: *"How to design a governance model for agricultural supply chain consortia using a blockchain-based traceability system?"*

Guided by this research question, we want to investigate how design requirements, principles, and features can be mapped to a governance model for agricultural supply chain consortia using a blockchain-based traceability system.

1.3 Structure

The structure of this paper is as follows: First, the foundations (Section 2) representing existing input knowledge and concepts used for our research are presented, followed by the applied research design (Section 3). Next, section 4 points out the expected results, including meta-design requirements, design principles, and a first suggestion of the governance model. Finally, this research-in-progress paper is concluded by discussing the contributions and future work (Section 5).

2 Foundations and Theoretical Background

2.1 Service-Dominant Logic

As our focus for developing a governance model is on the business value proposition for all participating actors, service science, especially service-dominant (S-D) logic, is considered. This research aims to link S-D logic and governance of blockchain-based traceability systems within supply chain consortia through co-creation and emphasizing relationships and trust. In our context, blockchain technology facilitates collaboration through a traceability system by providing all parties with access to a shared, immutable record of transactions. In S-D logic and supply chains, trust is a key driver of value creation and building strong relationships with suppliers, producers, or customers. By using blockchain technology to improve transparency and security of transactions, companies can help build trust and enhance value creation throughout the supply chain (Azzi, Chamoun, & Sokhn, 2019). Robert Lusch (Lusch, 2011) describes that shifting dominant thinking of supply chain management toward the concepts of service, value co-creation, value propositions, operant resources, networks, and service ecosystems opens up many research opportunities and strategies for improved organizational performance.

2.2 Blockchain Governance

As further input knowledge and concepts, we investigate governance, especially the governance of blockchain networks, as the technology provides the basis for the consortium. Blockchain governance can be captured as the integration of norms and culture, the laws and the code, the people and the institutions that facilitate coordination and determine a given organization (Fischer & Valiente, 2021). According to Weill (Weill & Ross, 2004), IT governance consists of three fundamental dimensions: incentives, decisions, and decision rights, as well as accountabilities. Considering blockchain governance, these dimensions have been retrieved, adapted, and extended by several researchers to capture the challenges, as well as research directions of this field (Beck, Müller-Bloch, & King, 2018; Yue, Kallempudi, Sha, Wei, & Liu, 2021; Pelt, Jansen, Baars, & Overbeek, 2021; Tan, Mahula, & Crompvoets, 2022).

2.3 Traceability Systems

Furthermore, transparent supply chain processes are becoming crucial for companies and organizations in fulfilling the current legislation and regulations (Hofmann, Pytel, & Winkelmann, 2020). Therefore, we need a blockchain-based traceability system – an information system – within a consortium capturing objects and events as information on the blockchain and analyzing data with forward and backward tracing (Sunyaev et al., 2021; Tian, 2017). Forcing traceability in organizations, the lack of governance is an emerging topic (Asprion, Hübner, & Moriggl, 2019).

3 Research Design

3.1 Overall Research

Our overall research goal is to develop a governance model for agriculture supply chain consortia, ensuring resilience in traceability systems.

This combines our previous outlined research interests of blockchain-based traceability systems and the inter-organizational governance perspective of consortia that are illustrated within two Design Science Research (DSR) grids according to vom Brocke and Maedche (Vom Brocke & Maedche, 2019) (see Fig. 1, Fig. 2).

For this purpose, we use the DSR approach according to Kuechler and Vaishnavi (Kuechler & Vaishnavi, 2008) to design our governance model. We include different stakeholder perspectives and combine different knowledge bases by using S-D logic as a kernel theory. S-D logic should thereby ensure a focus on aiming for customer and stakeholder needs as well as wants. We adopt our research approach with a method (Möller, Guggenberger, & Otto, 2020) for systematically generating design principles in an iterative supportive, or reflective way. This approach focuses on the design of

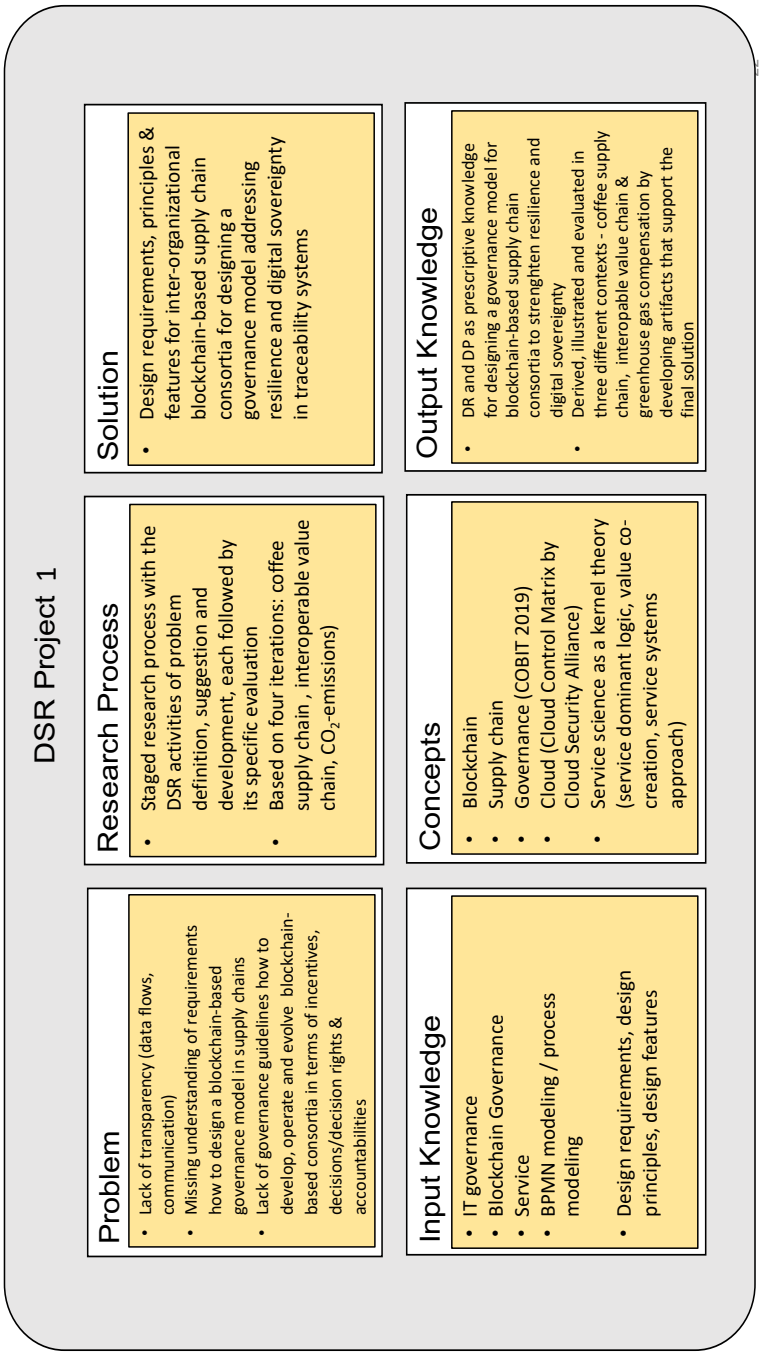


Fig. 1: DSR grid author 1 according to vom Brocke and Maedche (2019)

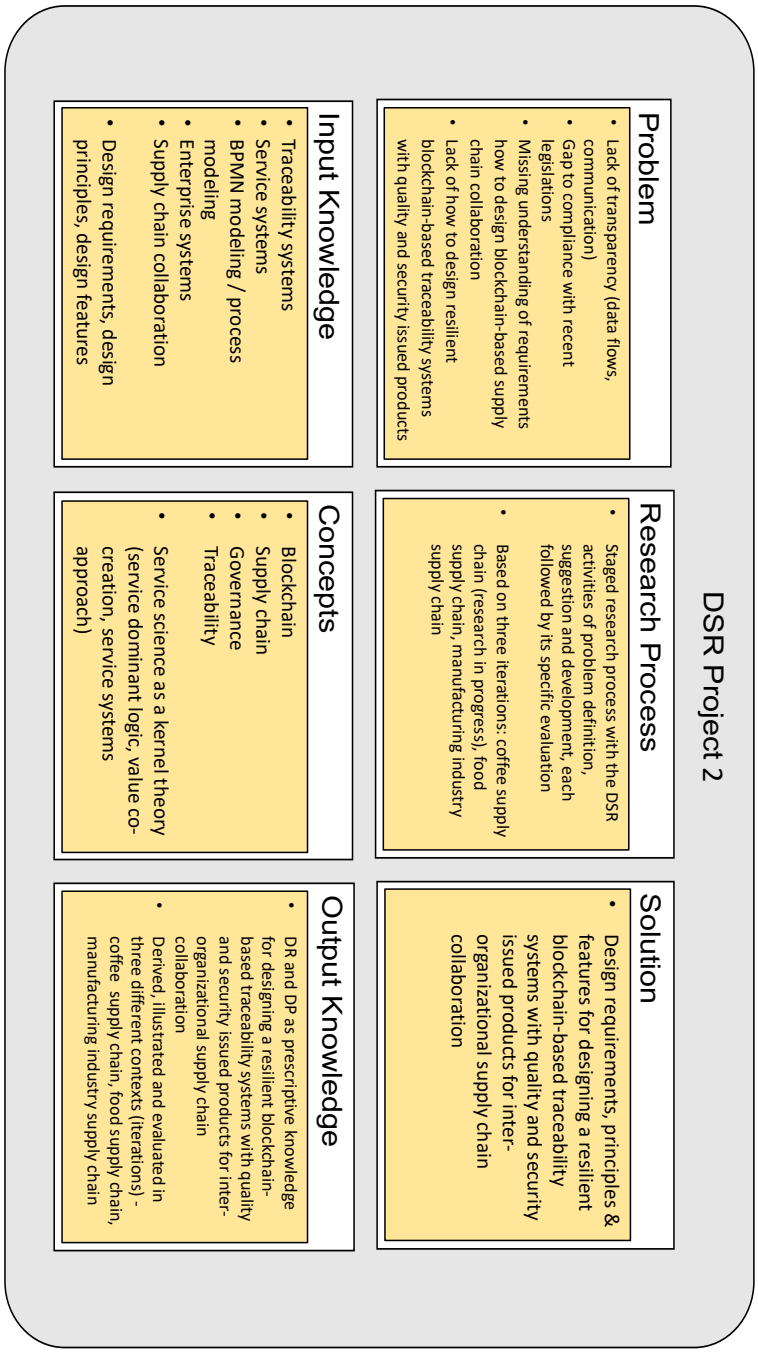


Fig. 2: DSR grid author 2 according to vom Brocke and Maedche (2019)

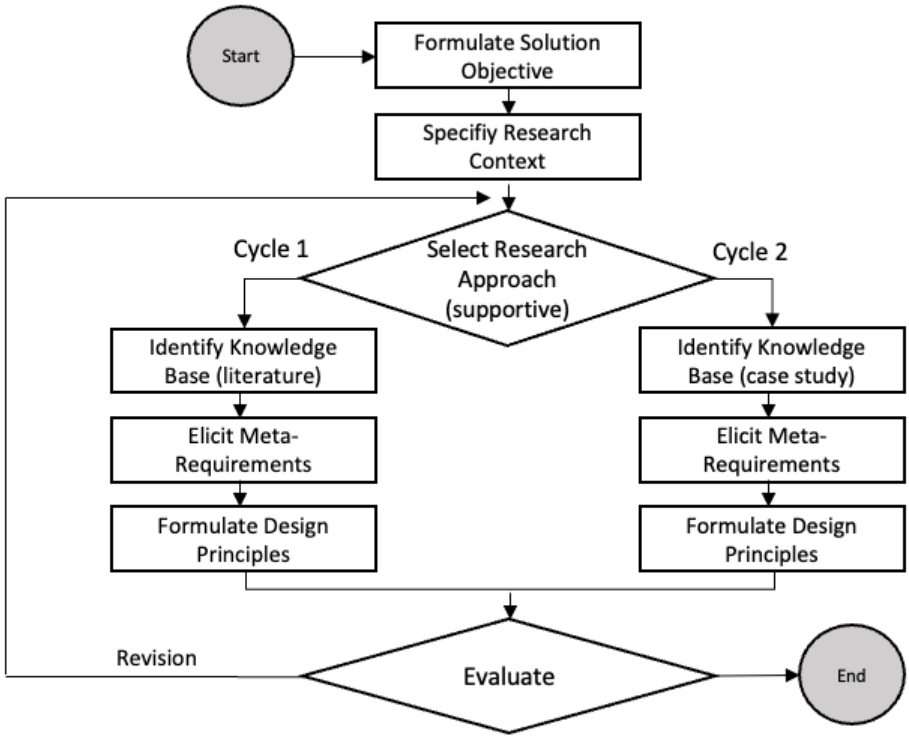


Fig. 3: Design science research model according to Möller et al. (2020)

an artifact and the theory associated with highlighting the importance of the iterative learning cycle of design science research considering the DSR design decision taxonomy (Smuts, Winter, Gerber, & van der Merwe, 2022).

In our overall research agenda, we systematically elaborate design requirements (DR), design principles (DP), and, additionally, design features (DF) for our governance model. This governance model includes roles, rules, and incentives, as well as structures and processes with associated possible alternatives to provide guidance for companies in participating or developing a supply chain consortium using blockchain-based traceability systems.

3.2 Research-in-progress

In this research-in-progress paper, we present the exploration of our DR and DP in terms of the first two steps, including problem definition and suggestion of the DSR process model (Kuechler & Vaishnavi, 2008). For generating DP, we use an iterative method (see Fig. 3) (Möller et al., 2020). We choose secondary sources (scientific literature) in the first iteration as a knowledge base. Next, we conduct a structured literature review

according to vom Brocke et al. (vom Brocke et al., 2009) on blockchain governance including databases *Web of Science, IEEE, EBSCO and ACM*. From this, we extract the first set of DR. Furthermore, we derive DP as general guidance. Afterwards, we evaluate our DP based on internal reviews by blockchain and supply chain experts. We use expert interviews as a new knowledge base in a second iteration. Therefore, we use semi-structured interviews (Glaser & Strauss, 1967) with five managing directors within the coffee supply chain. To analyse the results we used case study research following Yin (Yin, 2018). We refine our DR and DP with this method. In the final evaluation, we compare our DR and DP. Additionally, we demonstrate a prototypical suggestion for the governance model. Subsequently, we intend to extend the evaluation of iteration two in future research by conducting a survey with blockchain and supply chain experts.

4 Designing a Governance Model for Blockchain-based Traceability Systems in Agricultural Supply Chains

4.1 Identification of Meta-design Requirements (MDR) in Iteration 1

We start our first iteration with a literature review on blockchain governance as a knowledge base. After analyzing the results, 29 design requirements can be identified and aggregated into MDRs.

MDR1 Stakeholder trust addresses the need to understand how the blockchain economy can achieve trust between agricultural stakeholders (Beck et al., 2018; Yue et al., 2021) or how transparent decision processes can establish trust in an global environment.

MDR2 Compliance and laws include ensuring the integration of current and future cross-national compliance and legislative requirements towards the traceability system (Yue et al., 2021; Jnr., 2022).

MDR3 Transparent information flows represent the requirement of stakeholders to receive the data for decisions in the blockchain-based traceability system according to previously agreed upon policies and rights (Beck et al., 2018; Yue et al., 2021; Pelt et al., 2021).

In the next step, DPs were derived from the MDRs. At the end of the iteration, an internal evaluation based on blockchain and supply chain experts showed that especially the user perspective and industry-specific insights were not captured by the previously extracted DPs.

4.2 Identification of MDRs in Iteration 2

In the second iteration, a new knowledge base will be implemented as a foundation through interviews with agriculture supply chain experts. The interview analysis results in 18 design requirements, which are generalized into four meta-design requirements.

MDR4 Coordination & control covers the need for external deployment and maintenance of the blockchain-based traceability system considering the technological imbalances of the growing countries.

MDR5 Social & ethical awareness describes the responsibility of the individual participants of the value chain towards the social situation and environment of the farmers or producers.

MDR6 Inter-organizational collaboration gathers the requirements of experts in terms of data exchange, collaboration, and communication, as well as support and consulting activities for growing countries.

MDR7 Resilience refers to the adoption of changing circumstances, such as fluctuations in demand, disruptions in the availability of materials or components, or transportation delays caused by the global agricultural supply chain.

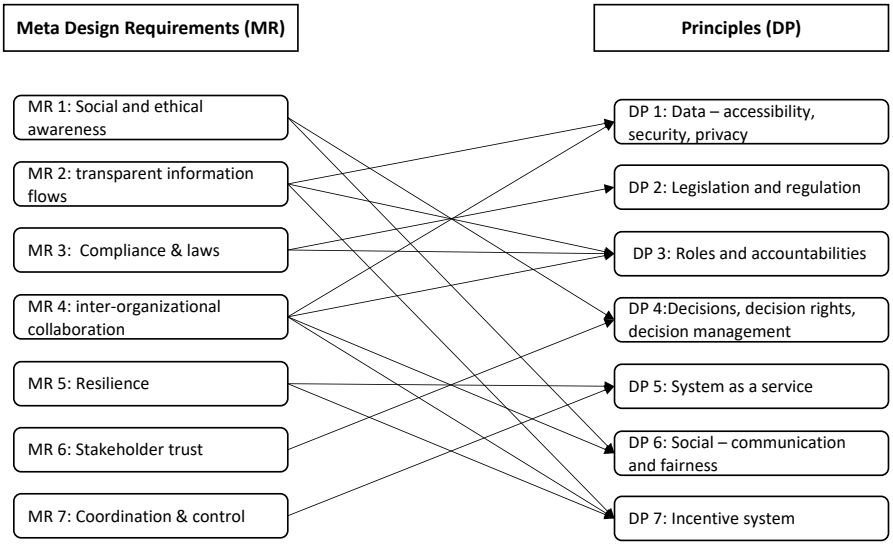


Fig. 4: Meta-design requirements and design principle - mapping diagram (own illustration)

We evaluate our DPs at the end of iteration 2 with a final comparison of all MDRs and DPs. In the next step we conclude our evaluation with a survey of external supply chain and blockchain experts (not part of this article).

4.3 Design principle deviation

We present our DP in a mapping diagram (see Fig. 4). This highlights the links between the DP and MDR.

DP1 Data covers data management of accessibility, security, and privacy in a cross-organizational consortium.

DP2 Legislation and regulation describe a mechanism to implement current and future regulations and laws in the structures and processes of the traceability system.

DP3 Roles and accountabilities address a transparent rights system that can be derived from existing organizational structures.

DP4 Decisions and decision rights cover trust and social factors in decision processes involving actors of producer countries.

DP5 System as a Service describes the provision of the traceability system as a service ensuring the inclusion of S-D logic by value creation for all participants and customers.

DP6 Social - Environment & fairness addresses the disbalances between consumer and producer countries and should enable improvements in working and environmental conditions.

DP7 Incentive system should provide incentives for stakeholders and customers to capture tensions for continuous participation within the consortium.

We understand our artifact, the elaborated DPs, as a starting point to provide resilience. The DPs can be used by traceability system providers or agricultural supply chain consortia to drive the building of their own governance model and implement it later into their blockchain-based traceability system. Our research also expands the knowledge base of the underlying domains. In the future, other researchers may incorporate our results, the DPs, into their own research.

4.4 Governance Model Approach

As already pointed out, the development of DRs and DPs is intended to lead to the identification of design features. These features will be the foundations for the design of concrete roles, rules, and incentives, as well as structures and processes, including associated alternatives, which in turn should be the basis for our overall artifact – the governance model (see Fig. 5).

This includes the description, explanation, and design of our information system, the blockchain-based traceability system, the associated processes, and process models.

The governance model for blockchain-based traceability systems in agricultural supply chain consortia is divided into three stages. The first stage represents the members and relationships, including the supply chain consortium: for example, stakeholders of the supply chain, the service providers, the regulator, and the customer. The second stage is given by key considerations developed from the design features identified by our proposed DPs, representing the processes and structures included in our governance domains (third stage). Finally, the model links the stakeholder to the respective processes and structures.

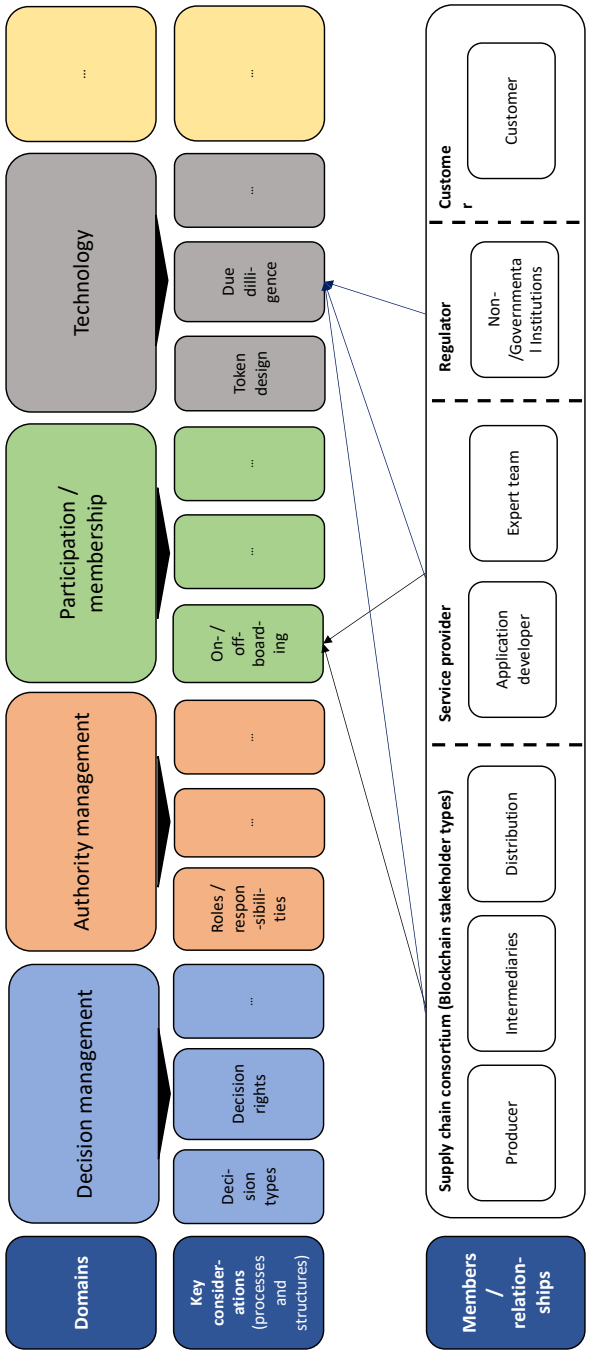


Fig. 5: Governance model for blockchain-based supply chain consortia (own illustration)

5 Discussion, Research-in-Progress, and Outlook

5.1 Summary

This research-in-progress paper addresses the lack of research on how to design a governance model for blockchain-based traceability systems in agricultural supply chain consortia considering the business value proposition for all participants, including customers and stakeholders. To this end, we derive seven tentative design principles from seven meta-design requirements based on 47 DRs (29 theoretical and 18 practical DRs). Due to our instantiation, we provide cumulative prescriptive knowledge and thus contribute to the knowledge base of blockchain governance and blockchain-based traceability systems. Furthermore, our proposed artifact (DPs) can be generalized to expand the understanding of governance for resilience in information systems wihtin supply chain consortia.

5.2 Further Research

In subsequent research, a survey addressing external block-chain and supply chain experts to validate the identified design principles is planned which are the basis for our governance domains (see Fig. 2). Afterwards, we want to implement the design principles into a governance model by developing design features. The latter are intended to support the governance model with concrete incentives, rules, processes, structures and alternatives that are demonstrated within the key consideration stage. For the evaluation of our governance model, we plan to use focus groups and surveys as well as evaluation of the fulfillment of the requirements to conclude the first iteration.

After this, the research interests will split. On the one hand, the research into blockchain governance for resilience and digital sovereignty plans to carry out three iterations.

This contributional research focused on setting the foundation, the coffee supply chain (Iteration 1). In the future, a second and third iteration will include greenhouse gas emissions and interoperable product and information flow of agricultural supply chains to derive further implications for governance and subsequent processes, structures, and alternatives. Therefore, further input knowledge in supply chain resilience, coordination, and digital sovereignty will be included. Within the final governance model, these processes and structures are intended to enable guidance for specific situations in developing, operating, or evolving a blockchain-based consortium to strengthen resilience and digital sovereignty. On the other hand, future research on traceability systems will focus on mapping organization structures on rights and accountabilities of the governance model, securing backward and forward traceability with governance guidelines, performance measures, and inter-organizational collaboration of the supply chain on the enterprise level.

5.3 Limitations

The preliminary nature and high abstraction of our DPs also represent a limitation emphasizing the lack of completeness at this stage. Furthermore, our DPs are based on literature reviews as well as five expert interviews validated by an internal evaluation. Before developing our design features, a comprehensive second evaluation, which is already in the development phase, needs to take place. Since our paper exclusively addresses a blockchain-based traceability system, a detailed investigation of alternatives, such as distributed or federated databases, should be included in the future.

5.4 Conclusion

Nevertheless, our research provides an initial foundation for further research on governance within blockchain-based consortia focusing on resilience as well as digital sovereignty.

Acknowledgments

This work originated in the LIONS project. Project LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU. The interviews that serve as empirical data were conducted within a cross-case study of this research project.

Contributions can be attributed to the coauthors as described below:

- Maximilian Greiner conducted the interview study and executed the data analysis (Sect. 4). Furthermore, he contributed the input knowledge in service science and blockchain governance within Section 2.
- Christian Zeiß designed the research approach (Sect.3) and executed the data analysis (Sect.4). In addition, he contributed the input knowledge for traceability systems and supply chain (Sect.2).
- The authors jointly wrote the introduction (Sect. 1), methodology (Sect. 2), and conclusion (Sect. 5).
- Ulrike Lechner contributed to the interpretation and revisions of the article.
- Axel Winkelmann contributed to the interpretation and revisions of the article.

References

Abbasi, M. (2017). Towards socially sustainable supply chains–themes and challenges. *European Business Review*.

Asprion, P., Hübner, P., & Moriggl, P. (2019, 01). Towards a distributed ledger system for supply chains.. doi: 10.24251/HICSS.2019.561

Azzi, R., Chamoun, R. K., & Sokhn, M. (2019). The power of a blockchain-based supply chain. *Computers & Industrial Engineering*, *135*, 582–592.

Beck, R., Müller-Bloch, C., & King, J. L. (2018). Governance in the blockchain economy: A framework and research agenda. *Journal of the Association for Information Systems*, *19*(10), 1.

Bosona, T., & Gebresenbet, G. (2013). Food traceability as an integral part of logistics management in food and agricultural supply chain. *Food Control*, *33*(1), 32-48. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0956713513000790` doi: /10.1016/j.foodcont.2013.02.004

Chang, S. E., & Chen, Y. (2020). When blockchain meets supply chain: A systematic literature review on current development and potential applications. *IEEE Access*, *8*, 62478–62494.

Fischer, A., & Valiente, M.-C. (2021). Blockchain governance. *Internet Policy Review*, *10*(2), 1–10.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine Transaction. Retrieved from `https://books.google.de/books?id=oUxEAQAAIAAJ`

Gurtu, A., & Johny, J. (2019). Potential of blockchain technology in supply chain management: A literature review. *International Journal of Physical Distribution & Logistics Management*.

Hofmann, A., Pytel, N., & Winkelmann, A. (2020, 11). Tracing back the value stream with colored coins..

Hoiss, T., Seidenfad, K., & Lechner, U. (2021). Blockchain service operations-a structured approach to operate a blockchain solution. In *2021 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)* (pp. 11–19).

Jnr., B. A. (2022, June). Toward a collaborative governance model for distributed ledger technology adoption in organizations. *Environment Systems and Decisions*, *42*(2), 276-294. doi: 10.1007/s10669-022-09852-

Kharas, H., McArthur, J. W., & Ohno, I. (2022). *Breakthrough: The promise of frontier technologies for sustainable development*. Brookings Institution Press.

Kuechler, W., & Vaishnavi, V. (2008, 01). On theory development in design science research: Anatomy of a research project. *EJIS*, *17*, 489-504.

Lusch, R. F. (2011). Reframing supply chain management: A service-dominant logic perspective. *Journal of Supply Chain Management*, *47*(1), 14–18.

Möller, F., Guggenberger, T., & Otto, B. (2020, 11). Towards a method for design principle development in information systems. In (p. 208-220). doi: 10.1007/978-3-030-64823

-7_20

Pelt, R. v., Jansen, S., Baars, D., & Overbeek, S. (2021). Defining blockchain governance: a framework for analysis and comparison. *Information Systems Management*, *38*(1), 21–41.

Smuts, H., Winter, R., Gerber, A., & van der Merwe, A. (2022). "designing" design science research–a taxonomy for supporting study design decisions. In *International Conference on Design Science Research in Information Systems and Technology* (pp. 483–495).

Sunyaev, A., Kannengießer, N., Beck, R., Treiblmaier, H., Lacity, M., Kranz, J., & Luckow, A. (2021). Token economy. *Business Information Systems Engineering*, *63*(4), 457-478. doi: 10.1007/s12599-021-00684-1

Tan, E., Mahula, S., & Crompvoets, J. (2022). Blockchain governance in the public sector: A conceptual framework for public management. *Government Information Quarterly*, *39*(1), 101625.

Tian, F. (2017). A supply chain traceability system for food safety based on haccp, blockchain internet of things. In *2017 International Conference on Service Systems and Service Management* (p. 1-6). doi: 10.1109/ICSSSM.2017.7996119

Vom Brocke, J., & Maedche, A. (2019). The dsr grid: Six core dimensions for effectively planning and communicating design science research projects. *Electronic Markets*, *29*(3), 379–385.

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. In *European Conference on Information Systems.*

Weill, P., & Ross, J. W. (2004). *It governance: How top performers manage it decision rights for superior results*. Harvard Business Press.

Yin, R. K. (2018). *Case study research and applications*. Sage.

Yue, K.-B., Kallempudi, P., Sha, K., Wei, W., & Liu, X. (2021). Governance attributes of consortium blockchain applications. In *Twenty-seventh americas conference on information systems, montreal, 2021*.

Sovereign Skill-Constrained Project Scheduling under Uncertainty and Semi-Autonomous Workforces

Andreas Fink[1]

Abstract: Effective project management is of high relevance for business and administration. Current challenges include the allocation of human resources for knowledge tasks under consideration of scarce skills, uncertainties within a dynamic environment, and cross-organizational project work with external contractors. While these aspects are increasingly addressed in the scientific literature, there is a need to elaborate the corresponding problem characteristics and solution approaches to support the sovereign planning and control of projects. This paper builds on recent work on the multi-skilled, resource-constrained project scheduling problem and extends it to include labor costs, uncertain activity durations, and partially autonomous personnel from outside the project owner's organization. The objective of this study is to examine various approaches to modeling and solving such problems and to discuss selected results from computational simulation experiments.

Keywords: Project scheduling, Multi-skilling, Stochastic optimization, Robust optimization, Collaborative planning

1 Introduction

Projects, and the task of planning and scheduling them, are pervasive in practice, including the fields of research and development, digital engineering, manufacturing, and supply chain management. Effective project management contributes to organizations' value creation and profitability. Managing projects involves breaking down the overall undertaking into multiple activities (jobs, work items) and then purposefully scheduling these intertwined activities under consideration of scarce resources. The decision-making process involves determining which activities should be executed at what time and with what allocated resources, under consideration of the workforce with diverse skill profiles. Project scheduling is often made more difficult by the need to deal with uncertainty, such as the time required to complete activities. In addition, there can be the challenge of coordination in projects where semi-autonomous workforces collaborate across organizational boundaries. In this context, we understand sovereignty as the ability to effectively plan and control projects in the digital sphere in a self-determined manner from an organizational perspective. This involves optimizing a system's performance metrics, encompassing economic efficiency and resilience in the face of a dynamic environment. It is therefore valuable to hedge against adverse conditions that may be associated with uncertain activity durations, depending on one's own decisions and on the environment. Furthermore, with respect

[1] Helmut Schmidt University, Hamburg, andreas.fink@hsu-hh.de

to projects that extend beyond the boundaries of the project owner's organization and where other companies, contractors, or freelancers may be working together on projects with in-house personnel, principal-agent challenges may arise due to information asymmetries and conflicting goals among the project participants. This calls for appropriate decentralized coordination mechanisms, which in turn can contribute to the sovereign execution of projects in enterprise networks.

Given the complexity inherent in many real-world decision scenarios, the incorporation of uncertainty considerations is critical, and means of dealing with sequential decision making are of interest (Powell, 2022). However, a large proportion of optimization studies in the literature are based on deterministic data and related deterministic optimization techniques. Such data may come from a previous prediction step. In essence, one might start the endeavor with a world model that incorporates uncertainties, but then a point forecast is made and the resulting "best guess" (e.g., the most likely activity durations) is fed into a deterministic optimization procedure. While this should at least be followed by a sensitivity study, analyzing the potential impact of the original uncertainty and the variability in the prediction (considering distributional forecasts) can be complex and is often neglected. As an extension of this "predict, then optimize/act" paradigm, recent work has examined how to improve the coupling between these steps, for example by devising an adapted loss function for training prediction models depending on the decision error induced by a prediction (Elmachtoub & Grigas, 2022). Nevertheless, the fundamental challenge of developing effective decision policies remains, given the question of how to define the overall objective and how to evaluate the outcome under different types of uncertainty.

The goals and contributions of this study are i) to identify certain characteristics of project scheduling problems with semi-autonomous multi-skilled workforces and uncertainty, ii) to discuss different approaches to modeling and solving such problems, and iii) to computationally explore approaches that facilitate the sovereign planning and control of projects in terms of economic efficiency and robustness. In the following sections, we will first describe basic considerations and challenges, and then discuss different solution approaches. These will be analyzed using simulation experiments, as computational simulations are increasingly used to examine not only physical systems, but also socio-technical systems and digital artifacts (Beese et al., 2019).

2 Problem Description

2.1 Basic Model: Deterministic Data and Centralized Planning

We draw on the multi-skilled resource-constrained project scheduling problem (MSRCPSP, Snauwaert & Vanhoucke, 2022, 2023), but with the extension of costs for the human resources (workforce). In this model, a project basically consists of a

set N of n non-preemptable activities i with durations d_i (processing times) and finish-to-start precedence relations $(i, i') \in N \times N$, which can be modeled with a directed acyclic graph (with the activities on the nodes and the arcs depicting finish-to-start precedence constraints). Ignoring resource constraints at first, a project scheduling problem can be processed using the critical path method, which efficiently computes a project schedule with minimum makespan (project duration, as given by the completion time of the last activity of the project). The makespan is the most widely used objective in the project scheduling literature, but in practice additional cost factors may be of interest, too. The introduction of resource constraints means that we need to allocate limited resources to complete the activities. Under the common assumption of a set of renewable resources R – i.e., the workforce is not exhausted but is available again from period to period – and that each activity requires certain amounts of specific resources in each period to be performed, this leads to the *NP*-hard resource-constrained project scheduling problem (RCPSP), which is also a generalization of the job shop scheduling problem from production operations management (Blazewicz et al., 1983; Hartmann & Briskorn, 2022).

More recently, as an extension of the RCPSP, the multi-skilled resource-constrained project scheduling problem (MSRCPSP) deals not only with scarce resources, but also with skill scarcity associated with a set of categorical skills (Snauwaert & Vanhoucke, 2023): For each resource $k \in R$ (workforce), we know whether the resource has certain skills $j \in S$ (competence profile), and for each activity i, we know what skills j are required in what number r_{ij} to perform the activity. For example, in a digital engineering project the desired functional and non-functional features may be associated with activities that require a specific set of coding skills or proof of the corresponding professional certifications from the assigned workforce. In the MSRCPSP model, a resource can perform no more than one skill per activity and cannot be assigned to more than one activity at the same point in time. We extend the MSRCPSP model by additionally considering costs for the workers (resources), similar to the software project scheduling model (Alba & Chicano, 2007; Vega-Velázquez et al., 2018). For each assignment of a worker k to a scheduled activity i and thus the corresponding time span, a worker incurs costs according to a worker's price rate per time unit p_k. The MSRCPSP with workforce costs is illustrated in Fig. 1; in this example the critical path length, without observing resource restrictions and without considering workforce costs, would be 15 (activities 2, 3, 5, 7, 10).

Feasible project schedules can be represented by a vector of start time decisions s_i for all activities i and associated binary assignment decisions x_{ijk} for worker k to skill j of activity i, such that all precedence and skill resource constraints are observed. Within the set of feasible solutions, one aims at some performance measure (objective function). Here, we consider the minimization of the total cost in terms of a weighted sum of the project makespan and the resource costs (payments to the workers depending on the assignments and price rates), which is a common approach to balancing trade-offs between competing objectives.

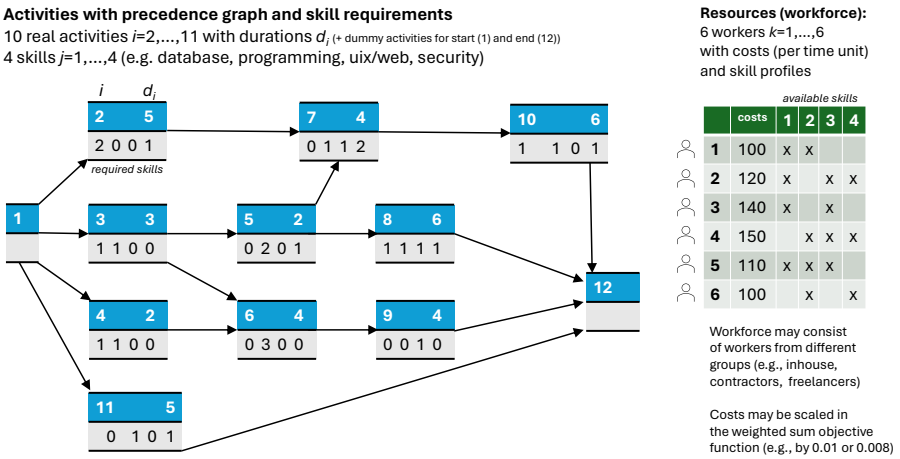


Fig. 1: Example of the considered project scheduling problem

Addressing the example shown in Fig. 1, and now paying attention to the resource constraints with skill requirements and taking into account workforce costs in the objective function (sum of the makespan and the workforce costs with the latter initially scaled by 0.01), we obtain a best project schedule, as shown in Fig. 2 at the top, with a makespan of 32 and a scaled workforce cost of 122 (using a deterministic optimization method as described in Section 3.1). Changing the weighting of the workforce costs, for example, by applying a reduced scaling factor of 0.008, may result in a different solution. In this case, we obtain a more compact schedule with a makespan of 27 and an adjusted workforce assignment with a scaled workforce cost of 101.6. While in this example the higher-cost workers are deployed less frequently, this may vary in other examples depending on the skill profile, scarcity, and cost.

For formal models of the problem and further literature references, we refer to the Appendix and the literature sources cited above. Many studies in the literature on the MSRCPSP address makespan minimization and do not consider resource costs. In situations where there is a deadline for the overall project completion yet minimizing the makespan is not a primary objective, an adapted approach could set an upper bound on the makespan and attempt to minimize the resource costs within the given time frame. Furthermore, the literature on multi-skilled resource-constrained project scheduling typically assumes that all data is known with certainty in advance and that a centralized planning approach is appropriate. Recent research suggests that it is necessary to go beyond these assumptions to address conditions that prevail in the real world. Snauwaert & Vanhoucke (2023) point out that there is "a clear future research perspective […] to study the impact of stochasticity on the MSRCPSP". In addition, recent work highlights the potential of decentralized project scheduling approaches (e.g., Wang et al., 2024; You et al., 2024). Both aspects will be addressed in the following sections, focusing on key aspects in connection with selected computational simulation experiments.

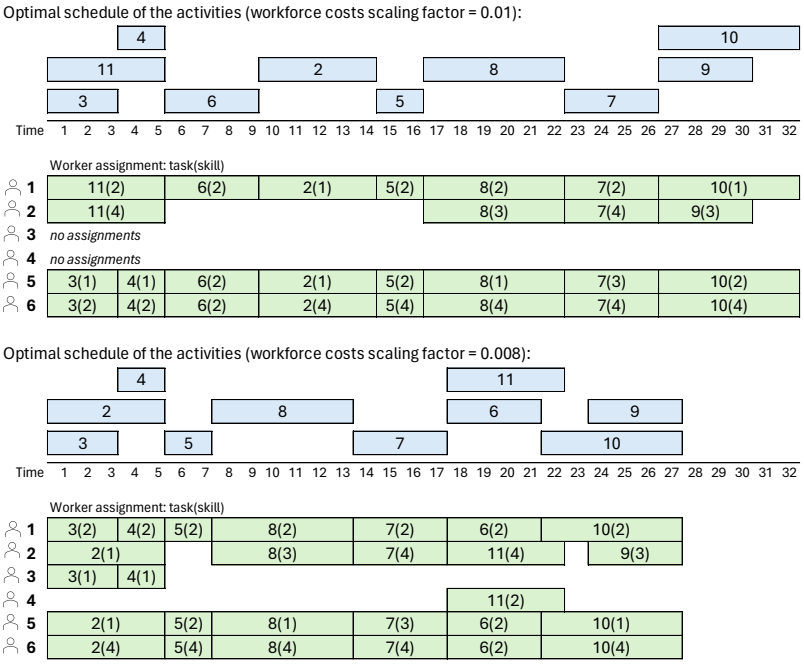


Fig. 2: Solution schedules for the exemplary project scheduling problem instance

2.2 Extension: Uncertainty About Activity Durations

In practice, project scheduling can face various types of uncertainty. For a summary of uncertainties in projects see Hazır & Ulusoy (2020) and for a general discussion of sources of uncertainties in the context of decision making see Powell (2022, Chap. 10.1). We also refer to Vanhoucke (2023), who emphasizes the importance of data availability for project management and discusses the different approaches that prevail in theory (academic world) and practice (professional world). In this paper, we focus on uncertainty in terms of activity durations that may be subject to random influences. When modeling uncertain activity durations or considering corresponding scenarios, one can distinguish between possible changes that involve both reductions and delays of the originally assumed activity durations, and the more common setting that encompasses only potential delays. In an advanced model, for example, random activity durations may be contingent upon effects associated with the assigned workers, which would require assessing the quality of workers and using this information in the selection and assignment of workers. Depending on the circumstances of the practical situation, one must be careful when predicting random distributions based on prior observations and when deriving associated scenarios by random sampling. In our study, we consider situations in which skill-related disruptions may occur or certain skill-related issues may become more challenging than expected beforehand, and

thus the completion of affected activities of the project, depending on the skill requirements, may be delayed. For example, it may turn out that working on high-performance data processing or addressing IT security issues becomes more complex and requires more work, thereby extending the duration of the affected activities, than anticipated in the original planning and scheduling of the project. Any approaches that can deal with such disruptions contribute to sovereign decision-making and resilient project scheduling.

For example, for the case of four skills $j \in \{1, \dots, 4\}$, each of which may or may not be affected by a complication, this leads to up to $2^4 = 16$ basic constellations by assuming that random delays depend on whether required skills are affected or not. If we restrict to scenarios where at most two out of four skills can be affected, this results in 11 constellations. Considering that an activity i requires r_{ij} resources (workers) for a skill j to be executed, we calculate an upper bound for the duration delay depending on whether skill j is in the set E of affected skills and an additional extent parameter e (e.g., $e = 1.0$), by $d_i' := \lceil d_i e \sum_{j \in E} r_{ij} \rceil$. Then, a realization of a potential random delay for activity i is obtained by drawing the duration uniformly from the range of integers $[d_i, \dots, d_i + d_i']$. Those scenarios with a larger number of affected skills will, on average, be associated with longer delays and thus higher makespan values. Depending on the circumstances, a finite set of scenarios may fully represent the set of possible realizations of the uncertainty, or the set of scenarios may only represent a random sample from a larger, and potentially infinite, set of random future paths.

Once one has a good understanding of the uncertainty involved in a decision situation, there are various approaches to dealing with that uncertainty depending on the principal goal of a decision-maker. One may seek to optimize the average expected outcome given some probability distributions or over a set of scenarios. With a focus on risk management, a risk-adjusted objective function may be appropriate that models risk-averse decision making with the goal of reducing variability in the outcome. A cautious decision-maker may seek a plan that focuses on value-at-risk or that performs as well as possible under all scenarios (worst-case consideration). The latter is related to approaches known as robust optimization or min-max optimization and does not necessarily require a probability distribution on the random variables (i.e., uncertain data within bounds, but not stochastic data) (Ben-Tal et al., 2009). For the classical RCPSP with the extension of activity duration uncertainty, several specific approaches have been discussed in the literature. For example, Ballestin (2007) describes and compares different heuristic methods, Bruni et al. (2015) study a constraint-based concept for determining schedules that should be able to absorb dynamic changes in activity durations, and Chakrabortty et al. (2017) follow a robust optimization concept and propose a specific branch-and-cut algorithm. However, none of these works consider skill requirements or the issue of labor costs.

2.3 Extension: Semi-Autonomous Workforce

A project's workforce may consist of individuals from different groups, including those employed by the organization that owns the project and those from outside the organization, such as personnel from contractor firms or freelancers. The latter groups may be autonomous in terms of whether they participate in the project at all and, if so, at what compensation (pay rates); but when they enter into agreements to participate in the project, this is subject to the constraints associated with feasible project planning and control (hence the use of the term "semi-autonomous"). In a broader context, addressing the sovereign sourcing of contractors for knowledge tasks involves the challenges of behavioral uncertainty, quality assessment, and reputation measurement (Benson et al., 2020). For example, Clausen et al. (2018) examine performance evaluation in project-based collaborations in the context of digital labor markets. The effective matching of supply and demand for highly skilled labor can be supported by digital labor platforms (Gussek & Wiesche, 2023; Wagner et al., 2021) and the promotion of technology standards to foster the interoperability between organizations (Fries et al., 2023). This may include the use of blockchains and digital ledger technology as a trusted repository of data about agreements, monitoring events, etc. in project planning and control (Seidenfad et al., 2023; Sonmez et al., 2023). Another sovereignty issue is how a project owner can avoid being exploited by inflated pay rates. In order to be able to choose from an external workforce in a cost-efficient manner, multi-sourcing is generally sought and an attempt is made to avoid heavy reliance on scarce personnel. This can be facilitated by a sensible planning of the project activities with the skill requirements, as discussed above for the MSRCPSP, and by establishing interoperable digital connections to the external labor market for participation in the project. In this way, companies can reduce dependencies and promote sovereignty.

In this study, we are interested in the impact of the strategic behavior of semi-autonomous workforces with respect to their requested pay rates on the allocation of workers to project activities. Specifically, prior to the actual project scheduling, workers can participate in a bidding process by announcing their desired pay rates. Workers may honestly announce their usual base rates, but they may also pursue an aggressive strategy by asking for inflated rates. The strategy adopted is unobservable, since the associated resource costs depend on the private information of workers about their initial costs, current workloads, and usual profit margins. Consequently, we are faced with the problem of decentralized decision making and the issue of coordination that takes into account self-interested worker agents and their strategic behavior.

2.4 Data

To analyze solution approaches, we draw on selected problem instances from the comprehensive benchmark dataset of Snauwaert & Vanhoucke (2023). To analyze the scalability of the computational methods, we consider project instances from the MSLIB with $n = 30$ (2 ×), 60, and 90 real activities, with five instances for each group

and medium parameter settings.[2] There are four skill types for all instances. We extend these instance files with worker pay rates taken from the corresponding SPLIB instance files. These salary values are added with a scaling by 10^{-4} to roughly balance the weight of project duration values (and thus the makespan) with labor costs.

3 Methods and Results

3.1 Solving the Deterministic Optimization Problem

Considering at first the basic model of the MSRCPSP of Section 2.1 (with deterministic data and centralized planning), both exact and heuristic solution approaches have been described in the literature (Snauwaert & Vanhoucke, 2021, 2022). In this chapter, we do not want to develop new specialized algorithms for this particular setting, but rather to use optimization models and general solver technology aimed at building a decision support system in an economical way to address specific features according to respective practical needs (without requiring highly skilled personnel, which would be necessary to design and implement problem-specific methods from scratch, albeit admittedly potentially leading to even better solution quality and improved runtime efficiency). In particular, we use the Gurobi mixed-integer programming (MIP) solver, which is widely recognized as a leading-edge solution.

The optimization models were built, and the solver software was accessed through Python programming interfaces using Pyomo (an open-source optimization modeling package). This involves modeling the problem in terms of the decision variables, the constraint expressions to be satisfied for a feasible solution, and the objective function depending on the data and the decision variables. The MIP formulation is based on that described by Snauwaert & Vanhoucke (2023). It models the problem using start time variables s_i for all activities i, binary decision variables x_{ijk} representing the assignment of workers k to skills j of activities i, and binary sequencing variables $z_{ii'}$ representing the sequencing requirements between activities i and i' in view of the precedence constraints as well as the endogenous disjunctive assignment of workers to activities. The objective function is given by $f(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{z}) = w_1 s_n + w_2 \sum_{ijk} d_i p_k x_{ijk}$ (with the makespan being represented by s_n that denotes the start time of the final dummy activity with zero duration). The complete MIP formulation is provided in the Appendix for reference.

Table 1 shows the computational results of using the MIP solver for the selected problem instances. The first two groups of project instances have 30 real project activities (plus two dummy activities for the start and end, i.e., $n = 32$); for the third and fourth

[2] This consists of the following instances with serial-parallel indicator *SP*=0.5, project skill strength *SS*=0.5, and resource availability *RA*=0.4: MSLIB_Set2_142# (*n*=30), MSLIB_Set2_172# (*n*=30), MSLIB_Set2_442# (*n*=60), and MSLIB_Set2_742# (*n*=90), each with # ∈ {1, ... ,5}.
Data source: https://www.projectmanagement.ugent.be/research/project_scheduling/MSRCPSP.

groups we have 60 and 90 real activities respectively. For reference and validation purposes, in columns 2–5 of the table we show the results for the MSRCPSP with the makespan-only objective function ($w_2 = 0$). For all considered instances of this type, a project schedule with a minimum makespan could be computed with runtimes of less than one minute, which is consistent with the results of the computational study by Snauwaert & Vanhoucke (2022).

Instance	**Results for the makespan-only objective function**				**Results for the total cost objective function (10 repetitions, each with max. runtime 600s)**					
MSLIB_ Set2_	Opt. [SV23]	Result	Gap [%]	Run-time [s]	Total costs [avg.]	Total costs [min]	Total costs [max]	Make-span [avg]	Lower bound [avg]	Gap [avg.]
1421	93	93	0%	0.8	322.9	322.9	323.2	93.4	322.1	0.3%
1422	91	91	0%	0.9	358.7	358.7	358.7	91.0	358.7	0.0%
1423	89	89	0%	0.7	272.1	272.1	272.1	93.0	272.0	0.0%
1424	83	83	0%	0.8	326.2	326.2	326.2	94.0	326.2	0.0%
1425	101	101	0%	0.8	340.3	340.3	340.3	101.0	340.3	0.0%
1721	72	72	0%	3.6	478.0	475.6	481.3	102.2	423.7	11.4%
1722	80	80	0%	4.1	445.8	443.7	447.8	85.6	428.4	3.9%
1723	91	91	0%	3.4	542.2	541.0	543.4	115.4	498.2	8.1%
1724	96	96	0%	3.6	459.2	459.1	459.5	109.4	447.7	2.5%
1725	81	81	0%	3.2	582.5	580.3	585.5	114.1	522.4	10.3%
4421	161	161	0%	4.6	557.7	557.5	558.2	169.6	553.8	0.7%
4422	175	175	0%	5.6	492.1	492.1	492.1	181.0	492.1	0.0%
4423	187	187	0%	4.5	546.0	546.0	546.0	187.0	546.0	0.0%
4424	167	167	0%	6.7	523.7	522.7	524.8	170.0	517.3	1.2%
4425	153	153	0%	6.4	512.5	511.4	513.7	153.8	508.3	0.8%
7421	244	244	0%	25.3	742.5	742.1	743.5	245.1	738.1	0.6%
7422	231	231	0%	20.2	767.6	759.9	777.9	247.1	742.2	3.3%
7423	281	281	0%	21.3	904.0	904.0	904.1	281.0	903.6	0.0%
7424	285	285	0%	17.2	866.4	865.7	867.0	286.0	861.2	0.6%
7425	250	250	0%	16.0	805.9	805.7	806.2	250.0	804.0	0.2%

Tab. 1: Results obtained with the MIP solver Gurobi 11.0 (with a maximum runtime of 600s on Xeon 8360Y processors utilizing max. four physical cores)

The main part of the table (columns 6–11) shows the results for the extended total cost objective function of our study. Due to the phenomenon of performance variability of MIP solvers (Lodi & Tramontani, 2013), we repeated each run ten times with different seed values to get a more robust picture. As expected, solving the problem becomes more difficult for the considered extended total cost objective function (sum of makespan and workforce costs weighted by $w_2 = 10^{-4}$ as described above). As shown in the table, in several cases the MIP solver was not able to identify provably optimal solutions within a ten-minute runtime limit (see the remaining gap from the lower bound in the respective column). However, the overall quality of the results is good for the most part. It is noteworthy that the gap values are larger for the second group of instances. Closer inspection of the data has shown that these cases involve a larger number of workers (82–96) than those in the first group (50–55), and in addition the skill requirements of the activities for the second group are broader than those in the first group, which is also associated with higher makespan values. This confirms the need for caution when performing computational studies and drawing conclusions.

However, even if the solver is able to determine high-quality solutions for the formulated model most of the time, the usefulness of such solutions in practice still depends on whether the "right" problem has been solved. If, after validation, it is found that the quality of the solution deteriorates when the real data deviates from the assumed expected case scenario, it seems appropriate to incorporate uncertainties into the solution approach instead of working only on improving the quality of the solution for the assumed deterministic case.

3.2 Uncertainty and Sequential Decision Making

To deal with nondeterministic optimization problems in general, we refer to approaches discussed under terms such as stochastic optimization, stochastic programming, sequential decision making, simulation optimization, and robust optimization (Birge and Louveaux, 2011; Powell, 2022; Fu, 2015). A typical framework for modeling and solving related optimization problems is to divide decision-making into stages, with a solution policy defining a decision procedure and the resulting action at each stage. In the common case of two stages, the goal is to first determine a promising initial solution, which in our study is an appropriate assignment of workers to project activities given skill requirements and capabilities. This is done in the absence of complete information, such as uncertain activity durations. The initial plan, which should comprise preventive schedule characteristics to account for the anticipated uncertainty of the project, can then be dynamically adjusted in one or more subsequent stages as the previously uncertain data becomes known. This is implemented through recourse actions that take into account the constraints and the remaining degrees of freedom for decision-making and action.

As described in Section 2.2, we assume that the uncertainty in activity durations depends on concerns related to the skills required. For example, IT security issues may become more complex, and the corresponding activities may take longer than expected. When this kind of randomness is represented by scenarios, the scenario data differs from the baseline data in activity durations while the overall project structure with skill requirements and the workforce remains the same. We generate scenarios using Monte-Carlo-based sampling with the distribution of activity duration delays as described above. Each scenario is considered equally likely to occur. Given a set of scenarios Q, we consider two alternative principal goals and corresponding decision strategies. First, we are interested in solving an extended MIP model that aims at the average expected outcome for a given set of scenarios; second, we are interested in robust optimization that focuses on worst-case considerations.

We model a two-stage stochastic program for the considered application in an extensive form by an extended MIP model that includes a set of scenarios (see the Appendix). The model involves scenario-specific activity durations d_{iq} and includes, as before, binary decision variables x_{ijk} representing a cross-scenario assignment of workers k to skills j of activities i to model the main allocation decisions (which should be

kept unchanged after the initial project planning stage). The second-stage decisions are scenario-specific and aim at executing the project plan with the changes necessary to cope with the changed data (i.e., adjusting the timeline as needed while maintaining the workforce allocation). Therefore, there are adjusted start time variables s_{iq} for all activities i and scenarios $q \in Q$ and the related binary sequencing variables $z_{ii'q}$ that address the sequencing requirements between activities i and i' in scenario q. The first-stage expected value objective function calculates the average outcome from the considered set of scenarios by $f_Q(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{z}) = \sum_{q \in Q}(w_1 s_{nq} + w_2 \sum_{ijk} d_{iq} p_k x_{ijk})/|Q|$, which is to be minimized. The main challenge in applying such kinds of stochastic programming approaches, which integrate the consideration of scenarios into deterministic non-convex MIP models, is that these models can become very large and may only be approximately solvable for a limited number of scenarios. For the robust optimization approach, we use a slightly adapted MIP model, with a worst-case objective function with a decision variable u to be minimized that is bounded by the scenario-specific total cost realizations: $u \geq w_1 s_{nq} + w_2 \sum_{ijk} d_{iq} p_k x_{ijk} \;\; \forall q \in Q$.

For the average outcome objective, we can on the one hand consider the set of scenarios as a representation of all possible realizations of the uncertainty. Then we could analyze the results for individual scenarios by iterating over the given set of scenarios. On the other hand, if the set of scenarios is to be understood as a random sample from a larger distribution of random future paths, one could analyze the outcome for a new test set of scenarios, and the original set of scenarios can be seen as a kind of training set. That is, the results obtained after solving the corresponding MIP formulation for the original training set of scenarios Q can be validated by an eventual simulation study for newly generated scenarios. As a benchmark, we consider a conventional "predict, then optimize/act" approach. For this, we calculate the average activity durations over the scenarios and use them within a corresponding MIP model. Below is a description of the overall experimental procedure:

1. Generate scenario set Q (with scenario-specific activity durations d_{iq})
2. For the scenario set Q, solve the first-stage extended MIP and store the determined cross-scenario worker assignments X_{scen}
3. Calculate average activity durations $\bar{d}_i = (\sum_{q \in Q} d_{iq})/|Q|$, solve the original MIP with this data, and store the worker assignments X_{avg} (benchmark)
4. Either set $Q' = Q$ or generate a new test set of scenarios Q', then repeat for each $q' \in Q'$ (second stage):
    a. Solve the original MIP with the worker assignments fixed to X_{scen}
    b. Solve the original MIP with the worker assignments fixed to X_{avg}
5. Compare the outcome of 4.a with the outcome of 4.b.

Figure 3 shows selected results for the first set of the problem instances with 30 real activities, considering the setting where the scenarios fully cover the uncertainty. We

use 16 randomly generated scenarios corresponding to the 16 basic skill-related constellations described in Sect. 2.2. The box plots depict the differences between the results of Step 4.a and Step 4.b (extended MIP with integrated scenarios and conventional MIP with average durations respectively). As can be seen, the results of the benchmark approach are mostly outperformed for the considered setup. However, comprehensive further computational experiments are needed to thoroughly analyze the performance of the considered approaches and to draw any general conclusions.

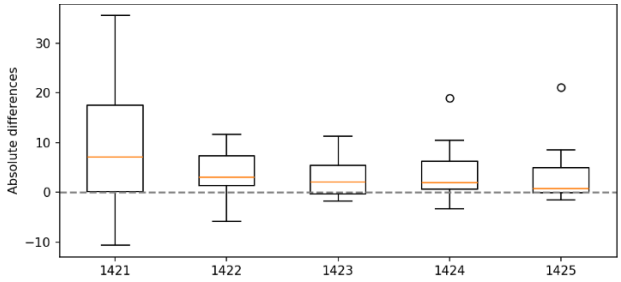


Fig. 3: Differences in solutions obtained by approximately solving the extended MIP with integrated scenarios vs. the conventional benchmark MIP with average durations, for the instances MSLIB_Set2_142x; positive values are those where the extended MIP with integrated scenarios performs better than the conventional benchmark MIP

The considered robust optimization approach can be useful when it is not easy to identify the worst-case scenario a priori. For example, in our setting the worst results (in terms of total cost) could result from data where each activity duration is individually set to its maximum value with respect to the individual upper bounds for the activity durations. However, if the set of scenarios contains scattered activity durations, a mechanism to identify the worst case is needed. To investigate such a situation, we take an example scenario set comprising the 11 constellations where at most two of the four skills are affected (see Sect. 2.2). We compare the result of the robust optimization procedure with the result of solving the original MIP with average activity durations, and follow an experimental procedure like the one described above (now for the robust optimization procedure, with the goal of identifying the worst-case scenario instead of the extended MIP with averaging the scenarios). As shown in Fig. 4, the robust optimization approach produces the expected results: For each worst-case scenario, the total cost is reduced, but this type of cautious decision-making comes at the cost of worse average results over the whole set of scenarios.

Problem instance MSLIB_Set2_...	1421	1422	1423	1424	1425
Absolute difference for the worst scenario	15.1	18.3	11.6	6.4	8.6
Absolute avg. difference over all scenarios	-3.6	-9.0	-20.5	-2.8	-3.0

Fig. 4: Effect of the robust optimization procedure over the scenario set for the instance set MSLIB_Set2_142x; negative values are those where the result of the robust optimization procedure is worse than the conventional benchmark MIP with average activity durations

3.3 Strategic Pricing Behavior of the Workforce

We consider a setting where the workforce consists of two groups (e.g., from different contracting firms) who act in a self-interested manner. For the computational experiments, we generate worker groups of equal size by randomly partitioning the total workforce (if the partitioning is not equal, the group sizes may differ by one). These groups aim to maximize their income or profit, which depends on the worker-specific pay rates and the endogenous allocation of workers to activities. As described in Sect. 2.3, the focus of this study is on the strategic behavior of workers with respect to their desired pay rates within a bidding process prior to the actual project scheduling. Two main strategies are considered: either workers announce their true pay rates (as provided in the respective dataset, see Sect. 2.4), or they covertly ask for inflated rates (e.g., 25% surplus). While in a competitive environment, higher pay rates may lead to a reduction in the number of assignments to activities, the relative income for the remaining assignments may increase. Given this trade-off, it is of interest to evaluate different pricing strategies.

Figure 5 (left) illustrates the effect of the considered strategic pricing behavior within the framework of non-cooperative game theory by means of a payoff table (for a single typical exemplary case). For each considered pricing strategy combination, the rounded payoff numbers represent the income for the two players (groups 1 and 2). Assuming common knowledge, or at least expectations, about the ordinal relations in the payoff table and examining for each group the best "answer" to each choice of the other group, the dominance criterion leads to an equilibrium where both groups keep up with the base rate as the dominant individual strategy (since inflated rates lead to inferior income values). The outcome may be different, for example, in less competitive cases where certain skills are particularly scarce or mainly available in a single group, thus increasing the bargaining power of such a group. Such cases should of course be avoided in view of the desired sovereignty of the project owner. It should also be noted that non-cooperative game theory usually assumes that binding agreements between the players are not possible, but in our application, there is a danger that autonomous workforce groups could collude to charge inflated prices; this would lead to an overall improvement for the workers at the expense of the project owner.

Payoff:		Group 2			
Income		**1.00**		**1.25**	
Group 1	**1.00**	130	97	179	67
	1.25	104	153	158	127

Payoff:		Group 2			
Income - Costs		**1.00**		**1.25**	
Group 1	**1.00**	30	27	46	25
	1.25	43	36	61	53

Fig. 5: Exemplary effects of pricing strategies for the instance MSLIB_Set2_1421 and a random partition of the workforce in two groups that either select their default price rate (1.00) or one with 25% surplus (1.25)

While it should be kept in mind that the income of a group of workers who demand excessive prices may decrease due to fewer assignments, this raises the question of whether some of the workers can be partly deployed elsewhere (outside the project). This can be taken into account by opportunity costs, which can be dynamic depending

on the actual utilization. To illustrate respective effects, we assume that a worker's opportunity cost is 75% of the respective base rate, and then calculate the payoff of a group as the sum of the income values minus the respective opportunity costs. As shown in Fig. 5 on the right, this can lead to different behaviors. Here, group 1 will use inflated prices to maximize profits (regardless of the group 2's strategy), while for group 2 there is no clear strategy to prefer (since the best "answer" varies depending on the group 1's strategy). In general, such results depend on the specific setting and data. This necessitates the use of prediction and data-driven decision making in conjunction with the actual practical circumstances.

4 Conclusion

In this paper, we have addressed the question of how to sovereignly plan and control resource-constrained projects characterized by skill requirements, in view of the challenge of scheduling knowledge workers towards resilient solutions. In addition to the literature, we have considered an extended objective regarding the cost of human resources, a dynamic setting with uncertain data, and the issue of cross-organizational projects with external contractors. We have described different approaches to modeling and solving such problems, and we have provided selected computational results from related simulation experiments that demonstrate the usefulness of the approaches for the settings explored. While we thus contribute to both theory and practice, our explorations are limited in terms of the scenarios considered, the limited elaboration of the methods, and the selective computational experiments. Therefore, more comprehensive studies are needed to provide further generalizable findings.

Regarding future research, project scheduling and control in an uncertain environment can be pursued more intensively in the form of multi-stage sequential decision making, which allows one to better respond to random effects in the course of the project (instead of constructing and partially following an initial plan with adjustments for the entire planning horizon). In principle, project planning lends itself to sequential decision making, where one decides "online" which activities to start next from the current set of feasible candidates subject to precedence and resource constraints. This allows better use of the information available up to that point in time. This leads to agile planning under a rolling planning horizon, where, if required, a project may even be started without knowing about all the activities to be performed later. In any case, sequential decision-making policies will require certain knowledge of the uncertainties to be considered in order to plan effectively based on respective predictions and lookahead considerations. Furthermore, coordination between the self-interested groups of workers involved might be improved by extending the project scheduling procedure to include elements of negotiation (collaborative planning in a multi-agent setting), rather than relying solely on a price request at the beginning (see, e.g., Fink & Gerhards, 2021; Homberger & Fink, 2017; Wang et al., 2024). This can also be extended to the domain of multi-project management (see, e.g., Fink & Homberger, 2015; You et al., 2024).

References

Alba, E., & Chicano, J. F. (2007). Software project management with GAs. *Information Sciences, 177*(11*)*, 2380–2401.

Ballestín, F. (2007). When it is worthwhile to work with the stochastic RCPSP*? Journal of Scheduling, 10*(3), 153–166.

Beese, J., Haki, M. K., Aier, S., & Winter, R. (2019). Simulation-based research in information systems: Epistemic implications and a review of the status quo. *Business & Information Systems Engineering, 61*(4), 503–521.

Benson, A., Sojourner, A., & Umyarov, A. (2020): Can reputation discipline the gig economy? experimental evidence from an online labor market. *Management Science 66*(5), 1802–1825.

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). Robust Optimization. Princeton University Press.

Birge, J. R., & Louveaux, F. (2011). Introduction to Stochastic Programming. Springer.

Blazewicz, J., Lenstra, J. K., & Kan, A. R. (1983). Scheduling subject to resource constraints: classification and complexity. *Discrete Applied Mathematics, 5*(1), 11–24.

Bruni, M. E., Beraldi, P., & Guerriero, F. (2015). The stochastic resource-constrained project scheduling problem. In C. Schwindt & J. Zimmermann (Eds.), *Handbook on Project Management and Scheduling* Vol. 2 (pp. 811–835), Springer.

Chakrabortty, R. K., Sarker, R. A., & Essam, D. L. (2017). Resource constrained project scheduling with uncertain activity durations. *Computers & Industrial Engineering, 112*, 537–550.

Claussen, J., Khashabi, P., Kretschmer, T., & Seifried, M. (2018). Knowledge work in the sharing economy: What drives project success in online labor markets? SSRN 3102865.

Elmachtoub, A. N., & Grigas, P. (2022). Smart "predict, then optimize". *Management Science, 68*(1), 9–26.

Fink, A., & Gerhards, P. (2021): Negotiation mechanisms for the multi-agent multi-mode resource investment problem. *European Journal of Operational Research, 295*(1), 261–274.

Fink, A., & Homberger, J. (2015). Decentralized multi-project scheduling. In C. Schwindt & J. Zimmermann (Eds.), *Handbook on Project Management and Scheduling* Vol. 2 (pp. 685–706), Springer.

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., & Wendeborn, T. (2023). Towards a layer model for digital sovereignty: a holistic approach. In Hämmerli, B., et al., *Critical Information Infrastructures Security, 17th International Conference, CRITIS 2022, LNCS 13723* (pp. 119–139), Springer.

Fu, M. C. (Ed.) (2015). Handbook of Simulation Optimization (Vol. 216). Springer.

Gussek, L., & Wiesche, M. (2023). IT professionals in the gig economy: The success of IT freelancers on digital labor platforms. *Business & Information Systems Engineering 65*(5), 555–575.

Hartmann, S., & Briskorn, D. (2022). An updated survey of variants and extensions of the resource-constrained project scheduling problem. *European Journal of Operational Research, 297*(1), 1–14.

Hazır, Ö., & Ulusoy, G. (2020). A classification and review of approaches and methods for modeling uncertainty in projects. *International Journal of Production Economics, 223*, 107522.

Homberger, J., & Fink, A. (2017). Generic negotiation mechanisms with side payments – Design, analysis and application for decentralized resource-constrained multi-project scheduling problems. *European Journal of Oper. Research 261*(3), 1001–1012.

Lodi, A., & Tramontani, A. (2013). Performance variability in mixed-integer programming. In *Theory Driven by Influential Applications* (pp. 1–12). INFORMS.

Powell, W. B. (2022). Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions. Wiley.

Seidenfad, K., Greiner, M., Biermann, J., & Lechner, U. (2023). CarbonEdge: Collaborative blockchain-based monitoring, reporting, and verification of greenhouse gas emissions on the edge. In Krieger, U. R., et al. (eds.) *Innovations for Community Services – I4CS 2023* (pp. 123–147), Springer.

Snauwaert, J., & Vanhoucke, M. (2021). A new algorithm for resource-constrained project scheduling with breadth and depth of skills. *European Journal of Operational Research, 292*(1), 43–59.

Snauwaert, J., & Vanhoucke, M. (2022). Mathematical formulations for project scheduling problems with categorical and hierarchical skills. *Computers & Industrial Engineering, 169*, 108147.

Snauwaert, J., & Vanhoucke, M. (2023). A classification and new benchmark instances for the multi-skilled resource-constrained project scheduling problem. *European Journal of Operational Research, 307*(1), 1–19.

Sonmez, R., Sönmez, F.Ö. & Ahmadisheykhsarmast, S. (2023). Blockchain in project management: a systematic review of use cases and a design decision framework. *Journal of Ambient Intelligence and Humanized Computing, 14*, 8433–8447.

Vanhoucke, M. (2023). The Illusion of Control: Project Data, Computer Algorithms and Human Intuition for Project Management and Control. Springer Nature.

Vega-Velázquez, M. Á., García-Nájera, A., & Cervantes, H. (2018). A survey on the software project scheduling problem. *Int. Journal of Production Economics, 202*, 145–161.

Wagner, G., Prester, J., & Paré, G. (2021). Exploring the boundaries and processes of digital platforms for knowledge work: A review of information systems research. *The Journal of Strategic Information Systems 30*(4), 101694.

Wang, X., Lu, S., Qian, X., Hu, C., & Liu, X. (2024). Dynamic scheduling of decentralized high-end equipment R&D projects via deep reinforcement learning. *Computers & Industrial Engineering*, 110018.

You, W., Xu, Z., Yu, Y., & Zhao, S. (2024). A two-layer approach for solving robust decentralized multiproject scheduling problem with multi-skilled staff. *International Transactions in Operational Research, 31*(3), 1631–1670.

Appendix: Mathematical Optimization Models

Deterministic MIP Model

Data:

N	set of n non-preemptable activities $i, i \in \{1, \dots, n\}$, with 1 and n representing the first and final dummy activities, respectively
N'	$= N \setminus \{1, n\}$, set of the real project activities
P	set of finish-to-start precedence relations $(i, i') \in N \times N$
d_i	duration of activity i (processing time)
l_i	earliest starting time of activity i (0 or a better lower bound)
M	upper bound for the makespan, e.g. $= \sum_i d_i$
R	set of renewable resources k (workers)
S	set of skills j
S_k	skill set of resource $k \in R$, $S_k \subseteq S$
r_{ij}	number of resources with skill j that are required for activity i
N_k	set of activities that require a skill that resource k masters, $N_k \subseteq N'$
p_k	price rate (per time unit) of resource k
w_1	weight factor of makespan in objective function
w_2	weight factor of resource costs in objective function

Decision variables:

s_i	start time of activity i
x_{ijk}	binary assignment with a value of 1 if and only if resource k is assigned to skill j of activity i
$z_{ii'}$	binary sequencing with a value of 1 if and only if activity i is finished before activity i' starts

Objective:

Minimize $f(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{z}) = w_1 s_n + w_2 \sum_{i \in N, j \in S, k \in R} d_i p_k x_{ijk}$

Constraints:

$\sum_{k \in R} x_{ijk} = r_{ij}$	$\forall i \in N', j \in S$
$s_i + d_i \leq s_{i'}$	$\forall (i, i') \in P$
$s_i + d_i - M(1 - z_{ii'}) \leq s_{i'}$	$\forall i \in N, i' \in N$: $i \neq i' \wedge (i, i') \notin P$
$\sum_{j \in S_k : r_{ij} \geq 1} x_{ijk} \leq 1$	$\forall k \in R, i \in N_k$
$z_{ii'} + z_{i'i} \leq 1$	$\forall i \in N, i' \in N$: $i < i' \wedge (i, i') \notin P$

$\sum_{j \in S_k : r_{ij} \geq 1} x_{ijk} + \sum_{j \in S_k : r_{i'j} \geq 1} x_{i'jk} \leq 1 + z_{ii'} + z_{i'i}$	$\forall k \in R, i \in N_k, i' \in N_k$: $i \neq i' \wedge (i,i') \notin P$
$x_{ijk} = 0$	$\forall i \in N, k \in R, j \in S \setminus S_k$
$s_i \geq l_i$	$\forall i \in N$
$x_{ijk} \in \{0,1\}$	$\forall i \in N, j \in S, k \in R$
$z_{ii'} \in \{0,1\}$	$\forall i \in N, i' \in N$

Extended MIP Model with Scenarios

Additional scenario data:

Q	set of scenarios q
d_{iq}	duration of activity i in scenario q
l_{iq}	earliest starting time of activity i in scenario q (0 or a better lower bound)
M_q	upper bound for the makespan in scenario q, e.g. $= \sum_i d_{iq}$

Decision variables:

s_{iq}	start time of activity i in scenario q
x_{ijk}	binary assignment with a value of 1 if and only if resource k is assigned to skill j of activity i (cross-scenario)
$z_{ii'q}$	binary sequencing with a value of 1 if and only if activity i is finished before activity i' starts in scenario q

Objective (average over the scenarios):

Minimize $f_Q(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{z}) = \sum_{q \in Q} (w_1 s_{nq} + w_2 \sum_{i \in N, j \in S, k \in R} d_{iq} p_k x_{ijk}) / |Q|$

Constraints:

$\sum_{k \in R} x_{ijk} = r_{ij}$	$\forall i \in N', j \in S$
$s_{iq} + d_{iq} \leq s_{i'q}$	$\forall (i,i') \in P, q \in Q$
$s_{iq} + d_{iq} - M_q(1 - z_{ii'q}) \leq s_{i'q}$	$\forall i \in N, i' \in N, q \in Q$: $i \neq i' \wedge (i,i') \notin P$
$\sum_{j \in S_k : r_{ij} \geq 1} x_{ijk} \leq 1$	$\forall k \in R, i \in N_k$
$z_{ii'q} + z_{i'iq} \leq 1$	$\forall i \in N, i' \in N, q \in Q$: $i < i' \wedge (i,i') \notin P$
$\sum_{j \in S_k : r_{ij} \geq 1} x_{ijk} + \sum_{j \in S_k : r_{i'j} \geq 1} x_{i'jk} \leq 1 + z_{ii'q} + z_{i'iq}$	$\forall k \in R, i \in N_k, i' \in N_k$, $q \in Q$: $i \neq i' \wedge (i,i') \notin P$
$x_{ijk} = 0$	$\forall i \in N, k \in R, j \in S \setminus S_k$
$s_{iq} \geq l_{iq}$	$\forall i \in N, q \in Q$
$x_{ijk} \in \{0,1\}$	$\forall i \in N, j \in S, k \in R$
$z_{ii'q} \in \{0,1\}$	$\forall i \in N, i' \in N, q \in Q$

System Design for Electronic Signatures within Supply Chains using Blockchain Technology and Self-Sovereign Identities

Michael Hofmeier[1], Michael Grabatin[2], and Wolfgang Hommel[3]

Abstract: This paper presents the development of a system for electronic signatures and identity management using a self-sovereign approach and distributed ledger technology, enabling deployment and operation within supply chains for greater sovereignty of the organizations and the individual users. For this purpose, the current regulations of Europe and the United States are compiled, existing technologies are examined, and then an own design is derived, showing the technical and regulatory hurdles and resulting design decisions such as off-chain databases.

Keywords: Electronic Signatures, Self-sovereign Identities, Distributed Ledger Technology

1 Introduction

Within supply chains, where participants might be in a competitive relationship, **blockchain technology** can help to increase trust in the truth about data related to business processes. At the same time, the **sovereignty** of organizations is increased through participation in the network. By combining this with methods such as SSI, the sovereignty of individual employees can also be increased. Consequently, the approach described in this work has an impact on the digital sovereignty on the **individual** and **organizational** layers defined by Fries et al. (Fries et al., 2022).

In the fields of self-sovereign identities (SSI) and electronic signatures, there are various solutions and approaches with their respective advantages and disadvantages. The concept presented in this paper designs a system that brings both fields together, while pursuing an approach that attempts to maximize self-sovereignty, data protection, and –as the top priority, end-user friendliness with a major focus on electronic signatures.

Signatures play a decisive role in the digitization of business and organizational processes. If documents, information, instructions, or inquiries are transmitted electronically, their authenticity must be validated by the recipient, meaning that the sender's identity must be verifiable. This can be ensured by the use of electronic signatures. The difficulty is, on the one hand, to make the system self-explanatory for the end user and, on the other hand, not to become dependent on single third-party providers.

In this paper, we set up a design for a signature system based on SSI and distributed ledger technology (DLT). In this project, the user-friendly combination of electronic signatures, DLT and SSI is the main challenge.

1 University of the Bundeswehr Munich, Neubiberg, michael.hofmeier@unibw.de
2 University of the Bundeswehr Munich, Neubiberg, michael.grabatin@unibw.de
3 University of the Bundeswehr Munich, Neubiberg, wolfgang.hommel@unibw.de

First, the legal and technical framework conditions must be defined and assembled. The goal chosen in this case is the legal validity in Europe. In order to allow comparison and compatibility of the requirements with other representatives of the western economic area, the regulations of the USA as the major partner are also included.

Furthermore, a high level of security, but also above all a high level of usability, is aimed for. A positive user experience and easy comprehensibility are prerequisites for acceptance by individuals without prior experience or relevant technical background (Srivastava, 2012).

A further goal is an approach that is as self-sovereign as possible. Since electronic signatures cannot function without electronic identities, a choice must be made. SSI technology does not require a central provider and is a feasible choice, especially for use cases where there is a high demand for keeping the data trustworthy, secure, and private because it has the potential to revolutionize data exchange (Laatikainen, Kolehmainen, & Abrahamsson, 2021).

The identity part of the system is newly designed according to the signing requirements, and possible compatibility with existing techniques or protocols is to be checked afterwards. It has similarities to existing systems like Sovrin, such as the use of blockchains and asymmetric cryptography, which results from the same constraints and similar objectives, but it differs in the prioritization of objectives, especially in terms of usability. Also, this system aims to avoid tombstoning (non-delivery of privacy-problematic transactions in the blockchain). Boundary conditions and the resulting design decisions become visible in this development process.

The last decisive goal is to store the signatures as decentralized as possible and not to subject them to the control of a single provider or to presuppose trust in this provider. This is where DLT comes into play, e. g., as a private permissioned blockchain with a few but trusted instances in case of application within specific supply chains or as a public blockchain for universal use. This paper focuses on use within a consortium of organizations that also operates the system (see Section 2). In the case of blockchains, care must be taken to ensure that no personal data can enter the ledger. This is where a further major challenge lies in this project.

The rest of this paper is organized as follows: First, Section 2 defines the abstract application scenario. Section 3 describes the legal requirements for electronic signatures within the EU and the United States. The current state of the art for electronic and digital signatures is presented in Section 4. As the core contribution of this paper, Section 5 documents a novel approach for electronic signatures using DLT and SSI technology. Future considerations and ongoing research are outlined in Section 9.

2 Application scenario

We first assume an abstract scenario in which different companies of a supply chain, which may partly be in competition with each other, agree on the use of the system and also operate it, but without excluding third-party use. We assume the abstract supply chain outlined in Figure 1 (Chandra & Tumanyan, 2003), where each arrow represents a contractual relationship.

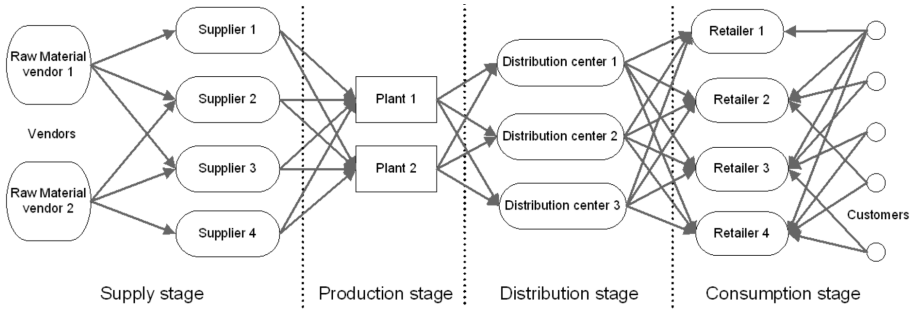


Fig. 1: Generic supply chain (Chandra & Tumanyan, 2003)

The supply chain consists of the stages Supply, Production, Distribution, and Consumption. Each stage contains participants with the same role, so they are likely to be in a competitive situation.

In terms of signatures, for example, there are many signatures between the plant and the supplier. This begins with the delivery contracts, which have to be signed by both parties; other departments within one party may also have to give their signatures, e. g., the management. In addition, many delivery bills are involved, which must be signed by the recipient, or invoices, which must be signed by the creator or sender.

3 Legal regulations for electronic signatures

Since this concept is intended to comply with the North American and European minimum requirements at the legal level, the regulations of the eIDAS of the European Union (EU) and the ESIGN Act and the UETA Act of the USA are taken into account.

3.1 European Union

The European regulation on electronic identification and trust services for electronic transactions in the internal market (eIDAS) sets basic requirements for electronic signature systems. In this context, a distinction is made between simple, advanced,

and qualified electronic signatures. The latter are, according to the regulation, equal to handwritten signatures in their legal effect.

Furthermore, the regulation states that neither the legal effect nor admissibility in court proceedings may be denied on the basis of the electronic form or the lack of qualification of an electronic signature (Parliament & the Council of the European Union, 2014).

Simple electronic signatures, which include scanned signatures or footers in emails, are not taken into account due to their lack of probative value. For advanced electronic signatures, Article 26 of the regulation defines four key criteria:

- Clear assignment to the signer
- Enabling the identification of the signer
- Creation using signature creation data that is subject to the signer's control (i. e., passwords, PINs, or keys)
- Connection with the signed data to detect changes (i. e., through cryptographic hash values or timestamps)

Qualified signatures require qualified digital certificates (issued by qualified trust service providers) and may not be achievable with a DLT-based system without the involvement of certification authorities. However, advanced electronic signatures are sufficient for the vast majority of business processes.

3.2 United States of America

In the USA, there are two laws in particular that need to be considered; first, the Electronic Signatures in Global and National Commerce Act (ESIGN Act) of 2000, and second, the Uniform Electronic Transactions Act (UETA Act) of 1999.

The ESIGN Act, similar to eIDAS, states that a signature, record, or contract signed with it may not be denied validity or the ability to be enforced merely because an electronic signature was used. The term "electronic signature" under the ESIGN Act means an electronic sound, symbol, or process attached to or logically associated with a contract or other record that is executed or adopted by a person with the intent to sign the record (Electronic Signatures in Global and National Commerce Act, 2000).

The UETA Act gives a very general definition of an electronic signature, so that (almost) the only criterion is that the signer must be aware of the legal effect and the fact of signing. The intention to sign is what defines the signature. Furthermore, the Act stipulates that the signature must be able to be assigned to the signer, even if this is performed by a human or electronic representative (Uniform Electronic Transactions Act, 1999).

3.3 Summary of the Legal Basis

The requirements of the various regulations can be summarized as follows:

The electronic signature must be unique to the signer, assignable to the signer, and created using data and systems unter the sole control of the signer. The signed object/document must be linked to the signature and protected from subsequent changes. The signer must be identifiable and aware of performing the action.

4 State of the Art

The options for electronic signatures are fairly limited. Not because they are particularly difficult to implement or because of the lack of suitable tools (Blythe, 2005), but because the usability of those systems is limited (Monzón, Tupia, & Bruzza, 2020; Zefferer, Krnjic, Stranacher, & Zwattendorfer, 2014) and innovation in legally binding processes is slow (Roßnagel, 2006). In the following, we describe system types which are current contenders for electronic signature systems.

4.1 Cloud-Based Electronic Signature Services

One solution used by businesses which require their customers to sign a contract is the use of cloud-based services. The need for a service like this might arise for any transaction otherwise requiring the mailing of a paper contract back and forth, e. g., buying a car or hiring a new employee. As those solutions are browser-based they are equally easy to use for consumers and businesses. However, as the contracts are processed by third parties, who accumulate many signed documents, those systems can pose security and data protection risks at the expense of general usability.

4.2 Public-Key Infrastructure

When it comes to digital signatures the solutions are usually based around public-key infrastructure (PKI) and asymmetric cryptography. One of the two main approaches for building a PKI involves hierarchical structures. In those, a certificate authority (CA) issues certificates that bind a specific entity to its public key and digitally signs a document with its private key. Other participants can validate the certificate by checking the digital signature against the CA's pre-distributed public key. The main standard for PKI certificates is X.509v3, which defines how certificates should be created and what additional metadata, e. g., expiration date or area of use, should be included.

The other approach is based on the "web of trust" concept. In contrast to the hierarchical structure, it establishes trust in the validity of the association between an entity and its public key by having it recognized, signed, and published on a database by other

entities. For example, with PGP/GPG this allows users to check pairs of email addresses and public keys by personal connections.

Both systems have advantages and disadvantages, but are capable of fulfilling the legal requirements described previously. From a technical point of view, the security of the hierarchical approach is determined by the choice of CAs. The web of trust system suffers from scalability, impersonation, and general acceptance issues. For regular non-technical users the hierarchical systems "just works", but understanding the trust and business relations between service providers and CAs is difficult. This results in superficial usability, where the significance of certificate warnings and "secure" connections is not understood. Similarly, web of trust systems overwhelm their users with secondary tasks, e. g., checking public key fingerprints.

4.3 Self-Sovereign Identity

An alternative to hierarchical or web of trust PKI is introduced with DLT. Besides its use-case for digital cryptographic currencies, DLT can also be utilized for identity and access management and the concept of SSI. One of the most consistent contenders in this relatively new field of research is being developed by the Hyperledger Foundation and called Hyperledger Indy. The focus of SSI is proving an entity's attributes to services. Without centralized control, this proves to be difficult especially for revocation and with keeping private data off any DLT.

For self-sovereign identities to be stored in wallets, there is a W3C standard called Decentralized Identifiers (DID). These DIDs are more or less an address to a DID document that represents the identity. This address indicates the method or source by which it is obtained. Verifiable Credentials can be issued to this DID to confirm certain properties, permissions or attestations. Ownership of the data promises a high degree of control over it. However, when pseudonymous information is used for authentication, DIDs offer little advantage over traditional approaches, and once personal data associated with an individual has been shared, that information is owned by the recipient and can be used for further purposes (Brunner, Gallersdörfer, Knirsch, Engel, & Matthes, 2020).

5 System approach

Our approach aims to utilize the advantages of SSI but focuses on digital signatures. The main difficulties are data protection due to the blockchain, a decentralized solution, and despite all this, a positive user experience.

This section is divided into the subsections *Objective*, *Technical Requirements*, *Architecture*, *Governance*, *Data Structure*, *Smart Contract*, *API*, *Usage*, *Requirements Fulfillment*, and *Prototype*.

5.1 Objective

As mentioned at the beginning, a highly user-friendly approach is intended which does not require any prior knowledge or deeper technical understanding on the part of the end user or the recipient of the signature. Consideration must be given to the way in which the signature is associated with the object or document and how it can be readable and verifiable without discarding the usability condition.

For this reason and for a use as universal as possible, we have made the design decision to enable a signature in the form of a URL that can be placed on the document or object. The URL is ideally also provided as a QR code for the case that the document is in paper form. It should have the following structure: **https://..../*signatureID*/*signatureKey***. The *signatureID* represents the Universally Unique Identifier (UUID) of the signature within the blockchain. The *signatureKey* is the key required to decrypt the personal data in the signature and is only present within this URL. More details follow in Section 5.3.2.

The use of HTTPS-based URLs makes the signatures associated with the document accessible to any recipient through the browser without the need for prior knowledge of the system. The user does not require a special application to read the signature(s), which allows any entity presented with the document (or item) digitally or physically to read and verify the signature.

The system used to create this signature should not be owned and operated by a central provider, but by a consortium, which results in the use of DLT. For the consideration of supply chains, an approach using a private permissioned blockchain is chosen, in which few instances are involved; here, the end user (employee) must trust one of them, e. g., his/her organization.

To ensure maximum trust and independence, the digital identities required for signing should be as self-sovereign as possible.

5.2 Technical requirements

The identity or wallet must generate or manage signature creation data that is unique to the wallet, must be able to attach the signer's name to the signatures, and must be able to cause deliberate actions to be performed. The wallet application must secure the identities it contains against unauthorized access and possible loss. This requires encryption and a backup solution.

The consortium must be decentralized, as some of the participants may be in competition with each other and therefore no single party may have dominion over the signatures and therefore the truth. Hence, the use of a blockchain results. However, the blockchain itself must never be at risk of obtaining personal data. It may only store IDs, hashes, and other non-problematic parameters.

The web service must connect to a database that is responsible for storing the personal data and protect this data. It consists of an API and a frontend to enable use via web and mobile applications, so that no potential user is technologically excluded. In addition, the front end is intended to provide information about the procedures used in an easy-to-understand form.

5.3 System architecture

Figure 2 shows the instances and applications involved, which are explained in more detail in the following.

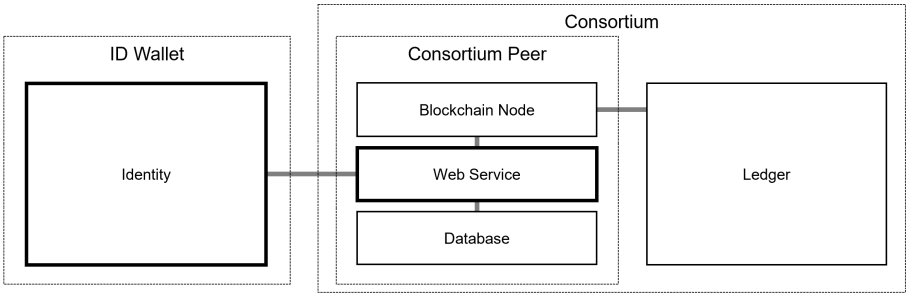


Fig. 2: System overview

5.3.1 Identity

An individual (or object) can have one or more digital identities. These identities are stored together with their properties, keys, and connection data in an application (wallet). This wallet can be a mobile application, desktop application, application extension, or similar.

The identity is defined by a public/private key pair of an asymmetric cryptographic algorithm. However, crypto-agility should be taken into account here. If a public key is transmitted, the corresponding string is given a prefix that defines the algorithm, e.g., `RSA:MIIECgKCB...`. This enables the future use of quantum-safe algorithms once they are standardized and widely available in software libraries. This key pair, which defines the identity, is used for signing.

An identity can register with different services. Separate key pairs are generated for each service so that the identity's activities cannot be tracked. The identity always determines the algorithm and key length.

An identity contains attributes, each consisting of a keyword and a value. Optionally, these attributes can also be signed (URL, see Section 5.1). Attributes such as "FirstName" and "LastName" are important for creating signatures and should ideally be signed by a trusted instance.

These identities must additionally be portable between wallets, and backups should be possible.

5.3.2 Consortium Peer

The composition of the consortium has already been roughly outlined. A limited number of participants is conceivable, ideally a consortium between organizations within specific supply chains that agree to use this technology. The admission requirements and the associated process as well as the assignment of rights must be regulated from the beginning and mapped in the smart contract of the private permissioned blockchain. Also, possible changes to the rules (or the smart contract) must be regulated.

A certain number of participants should provide a web service that serves as a communication and data interface for the creation and retrieval of signatures and/or enables the issuance of signed attributes including the necessary identification procedure. Those web services are interchangeable interfaces to the DLT backend, and for reliability and trust reasons there should be more than one. The service chosen by the user holds the encrypted personal data associated with his/her signatures and can be changed at any time. The communication between the wallet and this web services should be based on the same communication protocol as the communication between the wallet and other web services (Section 5.7), using the system for authentication/authorization, but extended by the necessary functions.

By providing multiple services within the consortium and the defined structure of the signature URL, the signature can be verified on multiple instances. The use of this type of web services integrates the advantages of cloud-based systems in terms of usability.

A database is connected to the web service and used to store the data of a signature, which must not enter the blockchain for data protection reasons. This data is stored in encrypted form and the associated *signatureKey* is only stored in the signature itself (in the URL). The data contained in the database can be validated by the signature data in the blockchain.

If a privacy issue ever arises with a signature, the associated record can be removed from the database, eliminating the user-friendly validation but maintaining judicial provability based on the other non-personal data such as timestamps, hash codes, and digital signatures. The reader of the signature can then no longer see the name of the signer, but still his/her digital signature, time stamp etc., and for verification then needs to know his public key.

5.4 Governance

Another important issue is blockchain governance. The admission requirements and associated process and the assignment of rights must be regulated from the beginning and mapped in the smart contract of the private permissioned blockchain.

Governance within the consortium is an important topic that also addresses issues such as onboarding new participants, changing the chaincode, or handling unexpected problems. In a broad sense, blockchain governance can be regarded as the integration of norms and culture, the laws and code, the people and institutions that facilitate coordination and together determine a given organisation (Aron & Valiente, 2021).

In addition to signatures (without personal data) and references to signable objects, the ledger/blockchain also contains information about public non-personal identities, in particular IDs and public keys of the organizations with authorization to confirm identities and sign the associated attributes.

5.5 Data Structure

Now that the different participants and applications of the system have been named, the most important properties (Figure 3, not conclusive) and their consequences are explained in this section.

5.5.1 Identity

As described, an identity is defined by a key pair from an asymmetric cryptographic algorithm such as RSA or ECC.

Each identity owns a set of attributes. These consist of a keyword, value, and optionally a signature URL that makes the value verifiable. Attributes can be properties, authorizations, or custom data. Proof of vaccinations or permitted driving classes from the driver's license are also imaginable.

Connection information to services is also stored within an identity. This consists in particular of the globally unique ID (maybe similar to an app ID) and the public key of the service as well as a key pair for the user (the identity). In addition, a symmetric key is stored/changed after each login, which is used for communication between the identity and the service. This key is an additional hurdle to the existing encryption of the HTTPS protocol. As a special feature, the ID wallet can log which data was transmitted to which service and when, and can also query the service directly to find out which data is stored on it.

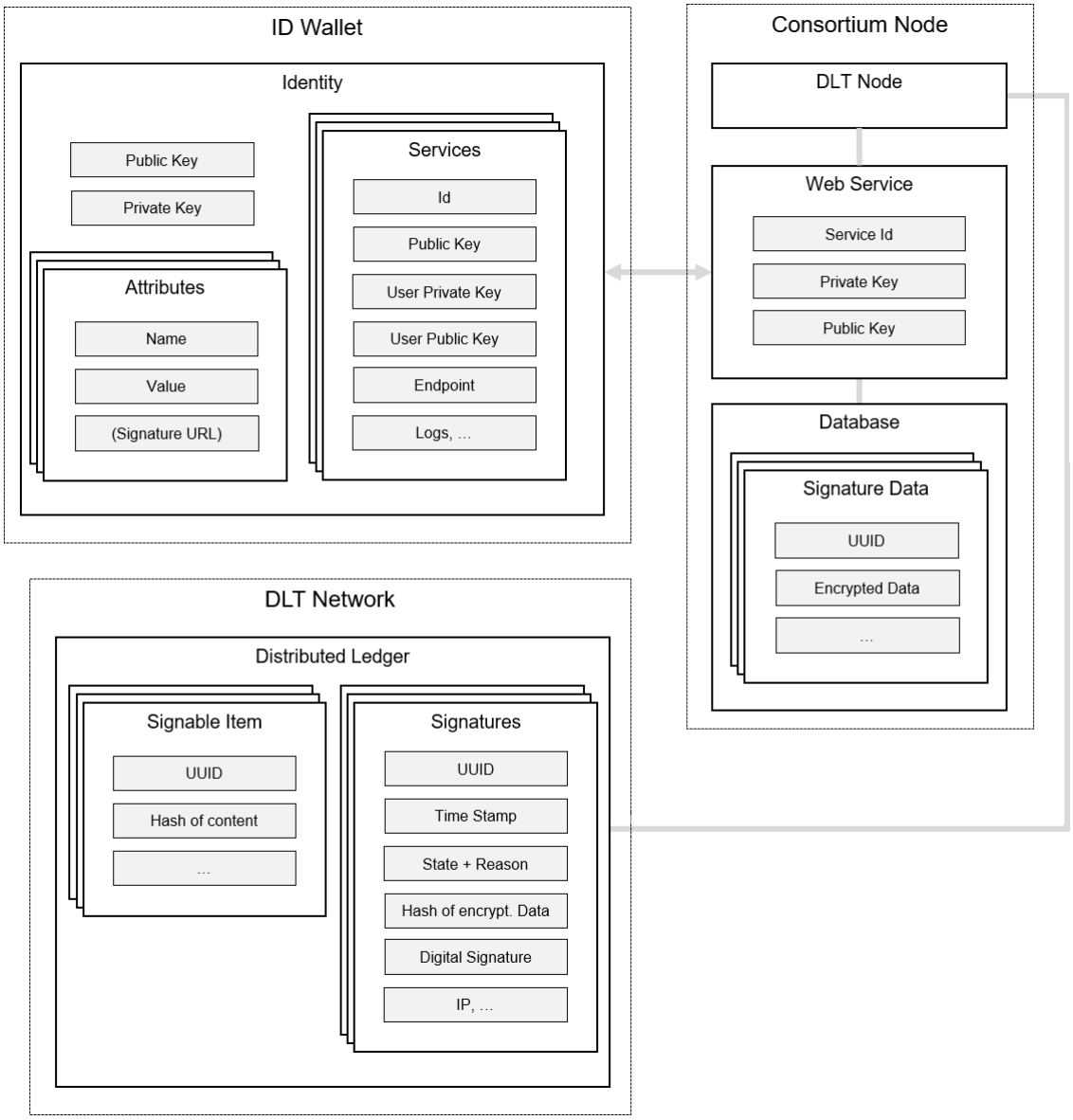


Fig. 3: Data structure with basic properties

5.5.2 Service / Node

In addition to the ID and own key pair, the service must store the public keys and IDs of the users. In the case of signatures, the signature data is stored, which may contain personal data or other tracking-relevant information. To do this, a data object containing this data is encrypted and stored in the database using the UUID of the associated blockchain entry. Only a hash value of the encrypted struct is passed on to the blockchain for checking or verification. Due to the prior encryption, it is not possible to determine the content of the data object from the hash, including by guessing.

However, a member of the consortium can also simply operate a blockchain node and dispense with the web service as an interface.

5.5.3 Ledger / Blockchain

Besides the signature-relevant data, the blockchain should also provide the public identities of the entities that have permission to identify identities and set and sign their attributes. These would be organizations within the supply chain.

In the blockchain, the signatures themselves consist of the ID (UUID), timestamp, status, the hash of the data, and a cryptographic signature of the hash value, generated with the private key of the identity.

In addition, reference objects to real objects, such as documents or contracts, can be added to the blockchain. A signature can then refer to such an object by its UUID. Such a reference object then consists of ID, name, and a hash value of the real content, e.g., a file. This enables the following scenario: A reference object for a contract is created and the URL/QR code is inserted in the contract file. As long as no hash code has been added to the object and no signature has been assigned, the hash code can be added later, e.g., after exporting the document as a PDF. This PDF can then be sent to all relevant recipients and signed using the QR code or URL. The transmission can take place in all conceivable ways (email, paper) and time-delayed. The web service must provide easy-to-understand functions for checking and use.

This system allows a high degree of independence regarding the applications and communication channels used.

5.6 Smart contract

The blockchain must provide a number of functions via smart contracts. Hyperledger Fabric is used in the implementation of this design. Here, the smart contracts are also called chaincode. The chaincode of this system provides e. g., the following functions:

Create signable objects:
Little information needs to be supplied to create signable objects. ID and timestamp are set by the chaincode. Result is an empty object as placeholder whose ID is necessary as return value to generate the corresponding URL.

Adding hashes to objects:
As long as no signature and no hash value has been assigned to a signable object, the hash value can be added. This is necessary so that the corresponding URLs can be generated before the associated file is completed.

Create signatures:
The most important function is storing signatures in the blockchain. Also here, ID, timestamp, etc. are generated by the chaincode. Optionally, an ID of a signable object can be provided.

Retrieving objects and signatures:
Of course, the chaincode also allows different queries, e. g., retrieving all signatures for a specific object.

5.7 Communication API

The communication between ID wallet and service is URL- and web-based. If the service has a request to the ID wallet, the user receives a URL and a QR code. This URL has a specific scheme (prefix) that allows the operating system to call the associated application, as long as the scheme was registered during application installation. The URL then contains the request data, which may be symmetrically encrypted and cryptographically signed. Signing the request ensures that no one else can impersonate this service. The URL is used in practice either by clicking on it (application on the same machine), or by scanning the QR code (application on a different device).

The identity responds to such a request via HTTPS. The endpoint for this is stored in the connection information of the service. The response is also symmetrically encrypted and signed with the identity's private key.

The cryptographic signatures and symmetric encryption pose a number of hurdles that make many attack vectors very difficult at best to execute.

5.8 Usage

In this section, the overall process is outlined in discrete steps from the user's perspective.

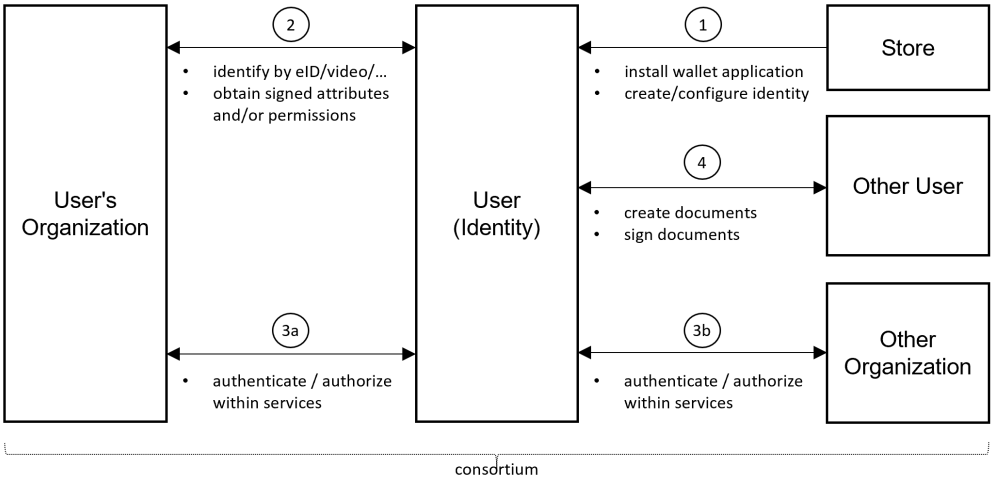


Fig. 4: User workflow

5.8.1 Identity creation

First, the end user must install the wallet application and then create an identity. This generates the cryptographic keys that define this identity.

5.8.2 Obtaining attributes

Next, the user can register and identify him-/herself with the web service of his/her organization. Identification can take place via eID, video ident, or other methods. Once the user is identified, certain attributes are set and signed by the organization's web service. These attributes can include data such as name and department, but also permissions.

5.8.3 Use of services

Once the identity including the required attributes is set up, the user can use it for authentication/authorization within other services within the consortium that implement this system. These could be cloud services or business applications.

5.8.4 Creation of document references and signatures

Once the user has created a document to be signed by one or more people, he or she can use the front end of the web service to register the document, receiving the URL and

the corresponding QR code and inserting these into the document. Once the document is finally saved/exported, the user can send the fingerprint (hash) of the file via the frontend and the timestamp is set. Now the user and others can use the URL/QR code to add their signature.

6 Validation

Finally, the fulfillment of the previously described legal and practical criteria must be checked.

6.1 Legal requirements

The assignment to the signer and the signer's identification is possible on the basis of the verifiable attributes (e.g., first name and last name) in the encrypted part of the signature. The private key that is used to cryptographically sign the whole thing again is only accessible to the signer, and hash codes of the files to be signed ensure that subsequent changes can be detected.

The signing process is mapped via the ID wallet and is confirmed by an active action on the part of the user, which means that the signer must be aware of the fact of signing. Furthermore, the signature is unique, not only because of its own keys, but also because of the signer's attributes.

To ensure compliance with the requirements within the EU, advanced electronic signatures can be generated with this process. In order to be able to create qualified electronic signatures and thus be equal to handwritten signatures, two basic requirements would have to be met; first, a qualified certificate would have to be used to generate an identity or its keys and attributes. Second, the blockchain consortium and/or the wallet application would have to be considered a qualified electronic signature creation unit according to eIDAS.

6.2 Data security requirement

Data protection in terms of the blockchain has been solved, and the use of URLs and web services creates a familiar user experience. Also, offering multiple services and blockchain nodes provides a decentralized and trustworthy solution.

6.3 Trust requirement

Also, offering multiple services and blockchain nodes provides a decentralized and trustworthy solution. By using DLT, no competitor has control over the truth, and every supply chain participant and its members, as well as third parties, have access to all the necessary technology.

6.4 Technical requirements

The identities are stored in encrypted files and are only loaded into memory for a short time when used. Backups must be solved by the wallet application, but since each identity is also a file, it can easily be backed up as well. Since a private permissioned blockchain is used in this scenario, no participant has data sovereignty and the data structure is built in a way that no personal data can get into the blockchain and thus no mechanism for tombstoning is necessary. The off-chain database for personal data protects it through prior encryption, and the web application is designed to provide a intuitive experience and requires nothing more than a standard web browser.

7 Prototype

At the time of submission of this paper, an early prototype of the system has been implemented. The implementation serves as a tech demo as well as a proof of concept and is part of an iterative development process. It provides experience that reveals which aspects of the architecture cause problems in implementation or use. One example is QR codes that are too large to be reliably captured by the camera due to the amount of data they contain.

7.1 Wallet application

The implementation involved the development of a mobile ID wallet using Microsoft's .NET MAUI (Multi-Platform App UI). The user has the ability to generate identities including the key pair. Since these are stored as a file on the device, they must be protected from access by other applications. The user has the option of encrypting this file with a password or a long randomly generated key, which is stored on an external chip (RFID/NFC) and must be held up to the device each time it is used. The application is automatically invoked when using URLs with the corresponding scheme.

7.2 Web service

A web service including the functions for signature creation was also implemented. ASP.NET was used for this, but for the moment all data is kept in RAM instead of a blockchain. The application allows users to register and log in, as well as register documents and generate signatures. In addition, the application is called by the URLs on the documents and presents the associated data (timestamp, hash, etc.) and signatures with their data (name, ID, timestamp, digital signature, etc.). In the further process, a Hyperledger Fabric network on a computing cluster will be installed and deployed as the underlying blockchain.

8 Conclusion

In this work, we have presented a system that can manage identities and signatures according to the SSI paradigm. A private permissioned blockchain is used to increase the sovereignty of the organizations participating in the system and their employees. Thus, we achieve a high level of **digital sovereignty** on the **individual** and **organizational** layers.

The legal requirements were met to the extent that at least advanced electronic signatures are possible. The data protection requirements were also met through the use of hashes and off-chain databases.

9 Future Work

This concept represents an initial state of work and has so far only been worked out in theory, together with the early prototype. The next step will be a complete implementation of the wallet and a signature service in conjunction with a working blockchain to test practical viability. For this purpose, a server cluster of the parent research project will be used, which makes it possible to test the blockchain under realistic conditions.

Furthermore, we will examine variations of this concept that rely on a public blockchain and modern cryptographic methods. Herein lie further challenges in terms of data protection and usability, but also the opportunity to contribute to digitalization processes.

Acknowledgment

This work originated from the LIONS research project. LIONS is funded by dtec.bw — Digitalization and Technology Research Center of the Bundeswehr, which we gratefully acknowledge. dtec.bw is funded by the European Union — NextGenerationEU.

References

Aron, F., & Valiente, M.-C. (2021). Blockchain governance. In *Internet policy review: Journal on internet regulation* (Vol. 10, p. 1-10).

Blythe, S. (2005, January). Digital Signature Law of the United Nations, European Union, United Kingdom and United States: Promotion of Growth in E-Commerce With Enhanced Security. *Richmond Journal of Law & Technology*, *11*(2), 6. Retrieved from `https://scholarship.richmond.edu/jolt/vol11/iss2/3`

Brunner, C., Gallersdörfer, U., Knirsch, F., Engel, D., & Matthes, F. (2020). Did and vc: Untangling decentralized identifiers and verifiable credentials for the web of trust. In *3rd international conference on blockchain technology and applications.*

Chandra, C., & Tumanyan, A. (2003). Supply chain system taxonomy: development and application. In *Proc. 12th annual industrial engineering research conference ierc-2003, portland, oregon, usa* (Vol. 313).

Electronic Signatures in Global and National Commerce Act. (2000). 15 U.S.C. Chapter 96.

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., & Wendeborn, T. (2022). Towards a layer model for digital sovereignty: A holistic approach. In *International conference on critical information infrastructures security* (pp. 119–139). Cham, Switzerland: Springer.

Laatikainen, G., Kolehmainen, T., & Abrahamsson, P. (2021). Self-sovereign identity ecosystems: benefits and challenges. In *Scandinavian conference on information systems.*

Monzón, F. H., Tupia, M., & Bruzza, M. (2020). Security Versus Usability in E-Government: Insights from the Literature. In Á. Rocha, M. Paredes-Calderón, & T. Guarda (Eds.), *Developments and Advances in Defense and Security* (pp. 29–42). Singapore: Springer Nature. doi: 10.1007/978-981-15-4875-8_3

Parliament, T. E., & the Council of the European Union. (2014). *European regulation on electronic identification and trust services for electronic transactions in the internal market.* Official Journal of the European Union.

Roßnagel, H. (2006). On Diffusion and Confusion – Why Electronic Signatures Have Failed. In S. Fischer-Hübner, S. Furnell, & C. Lambrinoudakis (Eds.), *Trust and Privacy in Digital Business* (pp. 71–80). Berlin, Heidelberg: Springer. doi: 10.1007/11824633_8

Srivastava, A. (2012). Electronic signatures: Legislative developments and acceptance issues. In *Electronic signatures for b2b contracts* (pp. 31–59). Springer.

Uniform Electronic Transactions Act. (1999). National Conference of Commissioners on Uniform State Laws. 1999.

Zefferer, T., Krnjic, V., Stranacher, K., & Zwattendorfer, B. (2014). Measuring Usability to Improve the Efficiency of Electronic Signature-Based e-Government Solutions. In M. P. Rodríguez-Bolívar (Ed.), *Measuring E-government Efficiency: The Opinions of Public Administrators and Other Stakeholders* (pp. 45–74). New York, NY: Springer. doi: 10.1007/978-1-4614-9982-4_4

Designing a Reputation Evaluation System for More Resilient IT Supply Chains

Razvan Hrestic[1], Maximilian Greiner[2], Andreas Fink[3], Ulrike Lechner[4], and Karl Seidenfad[5]

Abstract: The challenges in developing complex software systems become increasingly apparent as globally more and more cyber-incidents become known in which software supply chains and, increasingly, hardware supply chains are being targeted. Organizations in the industry and public sector alike are focusing on devising and implementing effective countermeasures, on increasing resilience, and on creating strategies to be able to act in a self-determined manner in the digital sphere. We attempt to address the issue in a holistic manner by integrating the supply chain perspective with information from the software development process and the IT security perspective in a reputation management system, and simulate the effect of using reputation information on decisions and the market. In simulation, we address the related questions of how organizations and their software and hardware supply chains (subsumed here as IT supply chains) can become more resilient and increase their digital sovereignty. The research-in-progress method proposed in this paper uses a combination of design science research (DSR), digital reputation systems, and agent-based simulation. The proposed system can also be used as governance support to aid sustainable IT ecosystems with fair conditions for not only larger organizations, but also smaller actors, freelancers, and open source communities.

Keywords: Reputation Systems, Supply Chain, Simulation, Supplier Selection

1 Introduction

Motivation. Globalization and the increasing complexity of supply chains pose major challenges for the resilience of organizations and require more efficient monitoring, controlling, coordination, and collaboration with suppliers. Cybersecurity threats and strategic considerations about sourcing are factors to be considered. The political concept of digital sovereignty is gaining momentum in political and scientific discourse (Fries et al., 2022). The supply of digital products and services, which we refer to as the "IT supply chain (ITSC)", is essential for modern society and its digital sovereignty.

One aspect of digital sovereignty is the ability to carefully screen suppliers while efficiently choosing from a wider pool of potential candidates (Baker, Kaye, & Terry, 2016). Facilitating fairness and inclusion of all kinds of actors, such as freelancers, open source communities, and companies, is important in times of digitization and a shortage of IT-experienced workforce.

1 University of the Bundeswehr Munich, Neubiberg, razvan.hrestic@unibw.de
2 University of the Bundeswehr Munich, Neubiberg, maximilian.greiner@unibw.de
3 Helmut Schmidt University, Hamburg, afink@hsu-hh.de
4 University of the Bundeswehr Munich, Neubiberg, ulrike.lechner@unibw.de
5 University of the Bundeswehr Munich, Neubiberg, karl.seidenfad@unibw.de

Our work specifically addresses challenges in supply chains for software, hardware, or cyber-physical systems. Organizations developing such systems as products face challenges from their customers, who not only want quality, but also have increased security ("Is my mobile device using weak wireless encryption?") and data protection requirements ("Is my child's drone collecting data without my knowledge or consent?"). They face the challenges of patching security vulnerabilities or compliance with, e.g., the NIS2-directive, which stipulates that certain components of critical infrastructure must be provided by trustworthy suppliers. It is essential to not only react to changes and disruptions, but also to plan possible reactions to them.

In the context of our research, we use specific meanings of key concepts as follows:

The capability to react is captured by **resilience**. Resilience is a multi-faceted concept where many dimensions emerge: (shock) absorption, recovery, and bouncing forward (Chandler & Coaffee, 2017). Thus, when working in the context of organizations and supply chains as systems of organizations, we use the following definition from the United Nations: "The ability of a system, community or society exposed to hazards to resist, absorb, accommodate, adapt to, transform and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions through risk management" (United Nations, 2016). For organizations, we use the availability in the market as a specific measure of resilience.

Digital Sovereignty describes the ability to act in a self-determined, secure, and independent manner on all digital issues. This may be understood from a political, but also from an economic organizational perspective (Pohle & Thiel, 2021). Digital sovereignty takes effort to be able to execute options to use or further develop digital goods and services. In an ITSC context, one can argue that increasing digital sovereignty means supporting the supply chain goal of creating a functional product with the desired functional and non-functional (e.g., security) requirements while not creating hard dependencies on any single supplier (Fries et al., 2022). It means fostering a market with enough capacity and quality to react or develop further. The capacity of the market and quality indicators of suppliers are our measurements.

Supply chain coordination and formation is a complex challenge, usually represented as a centrally coordinated process in the scientific discourse (Sucky, 2013; Vosooghidizaji, Taghipour, & Canel-Depitre, 2020). While centrally coordinated supply chains may have lower transaction costs (low-hanging fruit) as the issue of trust is delegated to one entity, distributed supply chains could be fairer and more resilient. The latter type must deal with trust and information asymmetry issues, thus warranting the use of reputation systems (Hendrikx, Bubendorfer, & Chard, 2015).

The main objective of the design research is to combine the advantages of a customizable reputation system with a simulation-based approach to gauge its fitness within the context of an IT supply chain scenario. The goal of this design is to increase resilience and digital sovereignty of the IT market. We use simulation as a method to experiment with the

design of such a reputation system. This work is embedded in a larger research effort in which field and design studies are conducted to use blockchain technologies to transfer quality and security information across the IT supply chain. This work is motivated by modern approaches such as DevSecOps that automate software engineering and that eventually deliver some of the content of such a reputation system in a tamper-free way across the ITSC. Our work is furthermore motivated by the Software Bill of Materials initiative (BSI, 2024) stating that all critical software products need to come with a bill of materials that details the implemented hardware and software components. Our research interest is to enhance this information with easily available information, like the documentation of DevSecOps tools, to have more information with the goal to react more efficiently in case of security incidents (resilience) and to plan more strategically for the future (digital sovereignty). The scenario-based simulation runs provide a first evaluation of the design science-guided approach of the overall research project and a first step towards the design of an analysis and prediction tool.

Based on the described challenges and our objective, the following guiding questions drive this article, each of them representing a different perspective upon the same phenomenon within a sum-part relationship:

1. How to assist organizations in improving decentralized reputation-based supplier selection systems for IT supply chains under the aspects of fairness, resilience, and digital sovereignty? (organization-centric view)
2. What consequences for the market result from increased resilience and digital sovereignty? (market-centric view)

Our unique contribution in the current research stage is in the instantiated reputation system combined with the agent-based simulation, while considering specific factors with respect to digital sovereignty and resilience. We thus attempt to find a more appropriate way of assessing risk for organizations or quasi-organizational entities (e.g., full-time freelancers) instead of sourcing this decision to a third-party platform or information system.

This article is structured as follows. First, we present the research design, followed by the scenario with the stakeholder analysis and related work. Section 5 presents the model design and implementation. Finally, the results and a discussion (Section 6) lead to the conclusion and outlook (Section 7).

2 Research Design

We follow the design science research methodology based on Peffers, Tuunanen, Rothenberger, and Chatterjee (2007) to design our artifact, consisting of a flexible reputation system for supplier selection and a simulation of a digital ecosystem. We also distinguish between three roles of the simulation in this contribution: (a) as an output of the design process, i.e., an artifact which can be used by organizations to achieve their goals for resilience and digital sovereignty, (b) as a means of evaluating the effects

of using specific methods for computing reputation on the entire ITSC, and (c) as a basic means of evaluation to improve the reputation model itself. The simulation was implemented in AnyLogic.

Our design procedure entails the following phases: problem identification and motivation, objectives of a solution, design and development, and demonstration and a first means of evaluation (Peffers et al., 2007). For the first phase of problem identification and motivation, we opted to use a scenario (Section 3) created in collaboration with an industrial research partner and in a dialog with industry. This collaboration is also an entry point for real-world use of our concept as a future system needs to be concretely described (Rosson & Carroll, 2012). For the second phase, to identify the problem

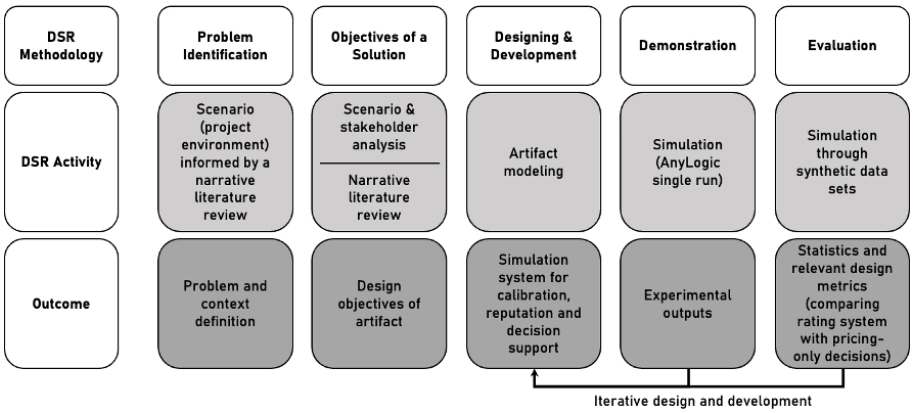


Fig. 1: Research design according to Peffers et al. (Peffers et al., 2007)

space, we describe stakeholders, needs, goals, and requirements (following (Maedche, Gregor, Morana, & Feine, 2019)) for our artifact as they emerge from our scenario.

The third phase covers the design and development of our simulation for the ecosystem and for the reputation model. Rein (Rein, 2005) provides guidance on reputation system design. The design of a simulation model is often based on a combination of deductive and inductive reasoning about the entities (agents) and processes (interactions). We derive the proposed simulation model mainly deductively, based on the literature on reputation systems and supplier selection, yet we also draw on insights gained from analyzing a tendering platform in the public sector. Due to the characteristics of the studied system, which is distinguished by interacting autonomous agents, we pursue an agent-based simulation approach. This serves to observe the behavior of the agents and the evolution of the system state over time depending on the respective assumptions and purposeful variations of the configuration and parameters (Macal, 2016). For the reputation system, we relate our design decisions to the general requirements for reputation systems described in (Vavilis, Petković, & Zannone, 2014).

Finally, our demonstration and evaluation phases are based on the simulation being run in the appropriate mode. For demonstration purposes, the single simulation run is used. For evaluation of the simulation, a set of parameter scenarios (Venable, Pries-Heje,

& Baskerville, 2012) have been chosen as the simulation model must be verified and validated as well (Sargant, 2005; Balci, 2003). To demonstrate the artifact effectiveness, we apply single simulation runs outlining statistics and relevant design metrics. In this type of simulation, one time-continuous run starting from given parameters and a random seed is executed. Thus, it is possible to check the outputs and perform basic plausibility checks. Finally, we run the simulation for set scenarios by varying simulation parameters, the results of which are discussed in Section 6.

To ensure alignment between theory and practice with respect to our artifact, we define the objective of our solution to the problem by implementing a rigor cycle with the underlying literature as knowledge base (Hevner, 2007). Relevance is derived from our real-world scenario and other involved industry partners.

3 Scenario and Stakeholder Analysis

A vacuum cleaner robot's value proposition is – at the time of the purchase decision – to clean the rooms over the next few years. Additional considerations to the value proposition are that it will not cause any hassle, that it will not film or record and send the material to foreign parties, and that it will not be active in any sense when switched off; updates may include optimizations of all algorithms, or the product may perform its duties and updates to keep it compliant with regulations as it needs, for exmaple, be compliant with night sleep time or recharge when power is cheapest. All in all, even for a simple product over a couple of years, many updates and optimizations may occur. The owner and the company producing the robot must ensure that they can do the necessary work to meet compliance, security, and functionality requirements.

Such a vacuum cleaner robot is a fairly simple product, and in our scenario, we assume a small digital ecosystem. To illustrate the IT supply chain of cyber-physical products, we introduce the robot producer: a fictional organization (HHH). This organization develops parts of the software in-house, but specifies and outsources the construction of hardware and other pieces of software for its product. The robot has two hardware components and three software components. We distinguish between off-the-shelf components (including open source) and custom software projects. Security properties are critical: the robot must not violate privacy by recording voice or images, transferring them online, or accepting a command to harm furniture, pets, or people. End customers and organizations want to be able to apply patches or develop functionality further (e.g., improved algorithms) over a defined period.

The organization wants to reduce costs and thus attempts to outsource components too expensive to produce itself, without threatening its resilience and risking introducing software vulnerabilities or lower-quality components. Thus, the resilience goal of the organization is conditional on the resilience goal of the entire supply chain. A similar consideration exists for digital sovereignty in the sense introduced at the beginning of this paper, i.e., sufficient suppliers are available for changes and functionality extensions during the planned lifetime without loss of quality or security.

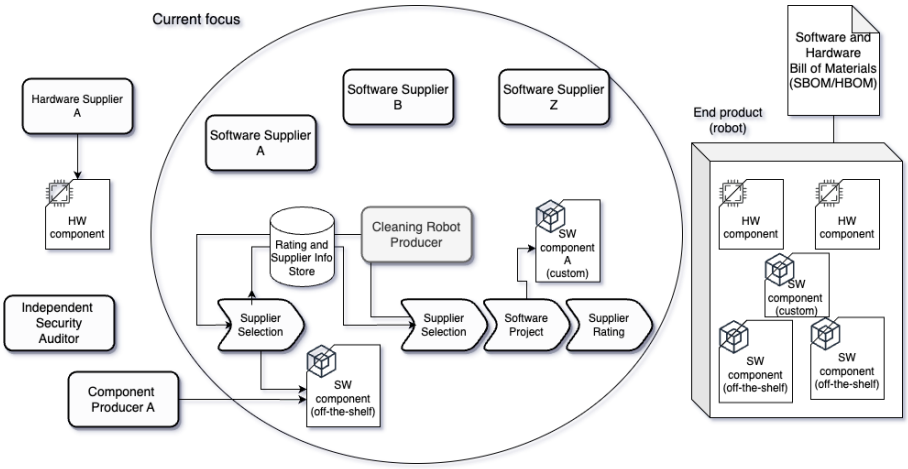


Fig. 2: Overview of the vacuum cleaner robot scenario

In our scenario, we are presented with a multifaceted organizational challenge involving several key stakeholders, each with distinct goals, needs, and requirements:

The procurement department, acting on behalf of the contractee, is tasked with identifying suppliers who not only meet the organization's criteria, but also contribute to a robust and reliable supply chain. Their challenge lies in acquiring current and trustworthy data for supplier selection and tracking project origins. They require easy access to contractor information from different sources, a simplified supplier ranking system, and the ability to integrate risk management and business rules into the supplier evaluation process (SR1).

The risk management department, also a contractee, aims to mitigate organizational risks by ensuring that the procurement team does not engage with potentially harmful suppliers. They need to establish minimum risk requirements for procurement processes to safeguard the organization (SR2).

The supply chain manager from a contractee organization is tasked with evaluating the business impact of forming and maintaining a supply chain. This involves analyzing current, sometimes partial or anonymized data, from both their organization and potential partners in the market. They require means to assess the impact of various business rules, pricing strategies, and other parameters on their organization and other entities within the supply chain (SR3). Additionally, they need to evaluate digital sovereignty and resilience based on specific operational criteria, ensuring supply chain robustness and adaptability in the face of changing market conditions and technological advancements (SR4).

In the upper tier of management (CxO of any organization), there is a concentrated effort on evaluating business potential and market dynamics. Their objective is to

comprehend and maneuver through the market configuration in a strategically aligned way, i.e., ensuring that decisions made at various levels are in line with the long-term vision and goals of the organization (SR5).

From a contractor's perspective, specifically in capacity management, the goal is to optimize the allocation of human resources within the organization. This requires access to current and reliable data on the size and scope of various projects, facilitating informed decision-making for bidding and resource allocation (SR6).

For all actors, an overall goal is to minimize information asymmetry, thus detecting bad or inefficient other actors. In the context of the supply chain, each actor has the goal of being more resilient while also maintaining the resilience of the supply chain. For the reputation system itself, we have used requirements set forth in (Vavilis et al., 2014).

4 Background and Related Work

Dishonest actors can disrupt or even destroy a market. As Akerlof wrote: "Dishonesty in business is a serious problem in underdeveloped countries" (Akerlof, 1970, p. 495). This is related to information asymmetries and resulting transaction or agency costs, which may lead to poor-quality products. Thus, one of the main goals of providing a reliable data source for rating market participants is the reduction of information asymmetries. Considering the digital space, where there is little or no developed legislation and even less enforcement, one could state that dishonesty in business is a serious problem in underdeveloped digital spaces, and thus that the more digitally sovereign a digital space is, the better it is able to promote and enforce honesty and fairness. In the context of the considered IT supply chain ecosystem, this necessitates reasonable organizational arrangements that govern the multi-party sourcing within project-based IT development work. From a more general perspective, such issues are also discussed in the literature under terms such as "crowdsourcing", "gig economy", and "digital labor markets/platforms" (Wagner, Prester, & Paré, 2021; Prester & Wagner, 2021; Benson, Sojourner, & Umyarov, 2020; Gussek & Wiesche, 2023). Related approaches on the one hand often consider simple task types ("gigs") with low task-specificity and low transaction costs under consideration of the challenges of quality assessment and behavioral uncertainty in the context of information asymmetry. On the other hand, there is increasing interest in means for the sovereign sourcing and control of digital work that involves knowledge-based tasks (e.g. high-skilled knowledge work) and respective challenges of matching supply and demand of labor with high-skill requirements (e.g., contributing to software artifacts). This leads to the necessity to assess the capability of delivering knowledge-based digital work, which may be facilitated by measuring reputation. For example, (Claussen, Khashabi, Kretschmer, & Seifried, 2018) study the drivers of agency costs and the assessment of performance within distributed and project-based collaborations in the context of online labor markets; they highlight the importance of managing performance based on trustworthy quality signals in order to mitigate problems of hidden information about the type of a potential contractor.

In a broader context, the term "reputation economy" has been coined, suggesting that reputation is a seminal asset (Fertik & Thompson, 2015; Gandini & Gandini, 2016). General requirements for reputation systems are described in (Vavilis et al., 2014). According to (Hendrikx et al., 2015) and (Ruohomaa, Kutvonen, & Koutrouli, 2007) a reputation system functions by facilitating the ability to gather, aggregate, and disseminate information about an entity, which may be used to forecast that entity's future behavior. This in turn reduces information asymmetry for market participants. Reputation systems are often based on ratings and feedback from others who have already had experience with respective suppliers and may assess various criteria such as delivery quality, price, delivery time, and customer service (Resnick, Kuwabara, Zeckhauser, & Friedman, 2000). According to (Liu & Munro, 2012), reputation systems can help to minimize potential risks in supplier selection. Further, in (Irissappane, Jiang, & Zhang, 2012) a testbed for reputation systems is proposed which offers an evaluation solution with respect to unfair rating detection. A similar perspective on reputation can be found in peer-to-peer (P2P) research such as the one described by (Griffiths, 2005). Specifically, the issue of multi-criteria trust can be framed in the context of autonomous agents with a set of beliefs regarding the trustworthiness of other agents. Lin et al. (Lin, Liu, & Viswanathan, 2018) study the effects of reputation in online labor markets where contractors provide clients with customized products such as computer software. Security issues are addressed in (Rauf, Lopez, Tun, Petre, & Nuseibeh, 2023) and (Rauf, Petre, Tun, Lopez, & Nuseibeh, 2023) for online freelance software development ecosystems.

Agent-based modeling and simulation have been shown to be useful in analyzing and optimizing complex systems of interactions in business between suppliers, manufacturers, and distributors as well as customers (Onggo & Foramitti, 2021; Rouzafzoon & Helo, 2016) and also specifically in IS Research (Beese, Haki, Aier, & Winter, 2019). By endowing agents with certain behaviors, different scenarios and decisions can be simulated to evaluate the impact on efficiency, cost, and customer satisfaction. This can support identifying bottlenecks, minimizing risks, and improving supply chain processes (Ghadimi & Heavey, 2013). For example, (Li, Lim, Chen, & Tan, 2015) address supplier selection through agent-based simulation. The models that are developed based on a set of defined criteria and the scenarios of a distribution supply chain reveal the impact of supplier profiles on key performance indicators, which in turn support the decision-making process of organizations. Further, (Ghadimi & Heavey, 2013) provide a review of the use and development of agent-based modeling and simulation for supplier selection. The approach of (Swaminathan, Smith, & Sadeh, 1998) demonstrates the usefulness of examining supply chain dynamics, using a multi-agent approach that provides a reusable base of domain-specific primitives enabling rapid development of customized decision support tools. More recently, the selection of sourcing strategies in supply chains has been investigated under consideration of multi-factor criteria (Mohammed et al., 2022; Torres-Sanchez, Saucedo-Martinez, Marmolejo-Saucedo, & Rodriguez-Aguilar, 2023; Rajesh & Ravi, 2015). This fundamental problem has been extended to consider supplier resilience under disruption as illustrated in (Wissuwa, Durach, & Choi, 2022) using the COVID-19 crisis as an example. Disruptions could

also be caused by attackers using software vulnerabilities in code (Gasiba, Lechner, Albuquerque, & Fernandez, 2020) and in cloud infrastructure deployment tools (Iosif, Gasiba, Zhao, Lechner, & Albuquerque, 2022). In the context of evaluation, agent-based simulation has been used to analyze complex systems and to investigate potential implications of changes or interventions (Ferreira & Borenstein, 2011).

5 Model Description and Simulation Design

In this section, we describe our reputation model and simulation system and relate these to the requirements for reputation systems described by Vavilis et al. (Vavilis et al., 2014) (R1 through R13), and also to our own requirements based on the chosen problem and scenario definition introduced in Section 2 (SR1 through SR6).

In fulfillment of requirements R3, R4, R5, R6, R7, R11 and R12, we employ a metric store component which immutably stores all transactions and rating information. We also employ a data store component for the multiple data source requirement for multi-factor decision making in supplier selection. For the data store, we need the additional assumptions of reliability and availability as they are critical in allowing contractees to compute ratings for contractors. The data sources for the data store should be heterogeneous and provide verifiable data points. For the scope of this paper, we assume these properties to be true. Further, in fulfillment of R5, R8 and R13 as well as SR1 and SR2, the reputation system is designed to be flexible in the metric computing method and in the selection of specific metric data. Regarding R9, only partial conformity can be guaranteed in a simulation, as measures to prevent fraud (e.g., actors with bad ratings could simply exit and re-join the ecosystem as "new" actors in order to reset their rating) rely on identity verification mechanisms which are relevant in the actual working system. We thus assume that all actors are identity-verified and are not able to commit rating fraud in the way described above. SR3 and SR4 are specifically addressed by the simulation artifact, as it is possible to create scenarios with multiple reputation models as starting points and see their impact on project success. Contractors using the simulation may use it in forecasting capacity requirements in fulfillment of SR6. By using a similar approach, members of management (SR5) can also be presented with different scenarios of how the ecosystem develops within given parameter ranges.

We model three-layered supply chains consisting of non-consumer customers (businesses, governmental or non-governmental organizations, etc.) on the demand side and professional suppliers (freelancers, businesses, other governmental or non-governmental organizations) on the supply side. The supply chains are not assumed to remain static but are re-formed after each epoch, so, for example, a supplier can work on the project of another customer.

Using the reference model from (Rein, 2005), we represent our reputation system as such with two modifications. Firstly, multiple communities (contexts) have not been considered as we assume a common platform. Secondly no reputation information system

(RIS) is assumed; Instead, the function of the RIS is taken over by the combination of the data store, metric store, and flexible reputation computing model. For the reputation system itself, we could identify three components: (a) the metric store, (b) the data store, (c) the metric computing method, and (d) the individual (weighted) rating model.

5.1 Conceptual Model

At the beginning of the simulation, the world is populated with a number of contractors and contractees given through model parameters. Each contractor then attempts to secure a project as fast as possible according to its preferences. In each round, a specific number of artifacts (e.g., code) is delivered and must undergo a quality check from the contractee. A project is thus comprised of one or more artifact deliveries corresponding to the fulfillment of project requirements. A product consists of multiple projects – two to four in accordance with our scenario.

An overview of the modeled process is given in Figure 3.

We use two main software-specific properties as the basis for our metrics. Namely, we consider the number of security vulnerabilities and code quality metrics to be predictive of the supplier's ability to deliver functionality and security software in the future. There are multiple public and private repositories of security vulnerabilities, such as the MITRE database (MITRE, 2024), and there has been research in the area of predicting software vulnerabilities (e.g., by Gasiba et al. (Espinha Gasiba, Lechner, Pinto-Albuquerque, & Méndez, 2021) or feature-based prediction (Neuhaus, Zimmermann, Holler, & Zeller, 2007)) based on other software characteristics. In the context of the larger research project, this scenario is embedded in a DevSecOps process and uses blockchain technology to share this information in a tamper-free manner. Part of the reputation is hereby linked with the security-lifecycle management. The second metric used is a software quality rating metric, a value based on an assessment such as a code review. The reputation system is, however, not limited to these two metrics; any organization can add or remove these in their reputation computation model.

Two agent types are central to the simulation: the customer organization (contractee) and the supplier organization (contractor).

Figure 4 shows a representation of our data model used in the simulation.

The contractor agent's capacity is currently limited to working on one project at a time. We thus necessarily limit the applicable scenario to the case where organizations (both the demand and supply side) do not opt for a single market, e.g., the fictional project platform in our model, but rather pilot this new option with a minimum capacity in order to evaluate it before giving up tried-and-tested business practices.

Agents have their behaviors driven by strategies. We assume rational agents in a non-cooperative game as defined by (Nash, 1950) so that each player has an individual strategy and utility function. We chose to cluster the options spectrum to three main

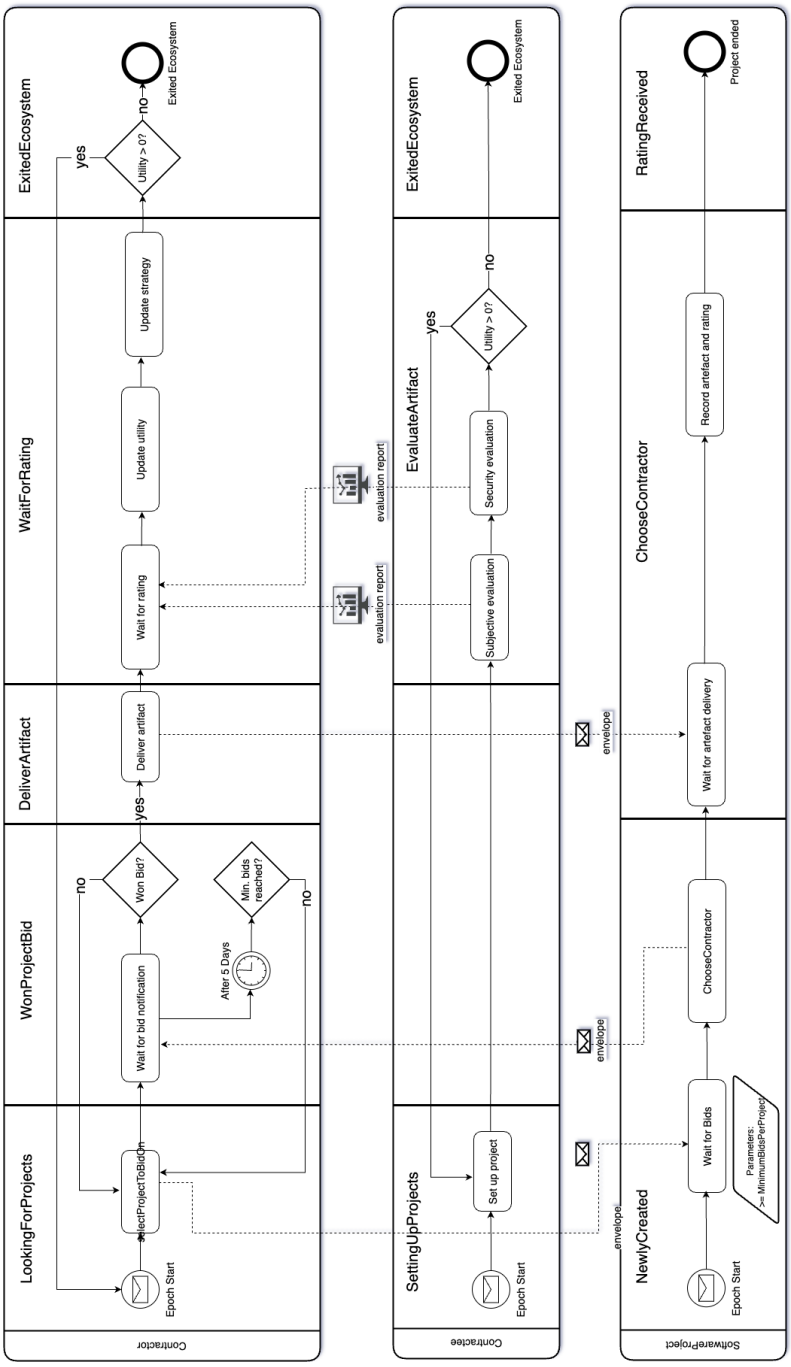


Fig. 3: Process view of a simulation cycle (epoch)

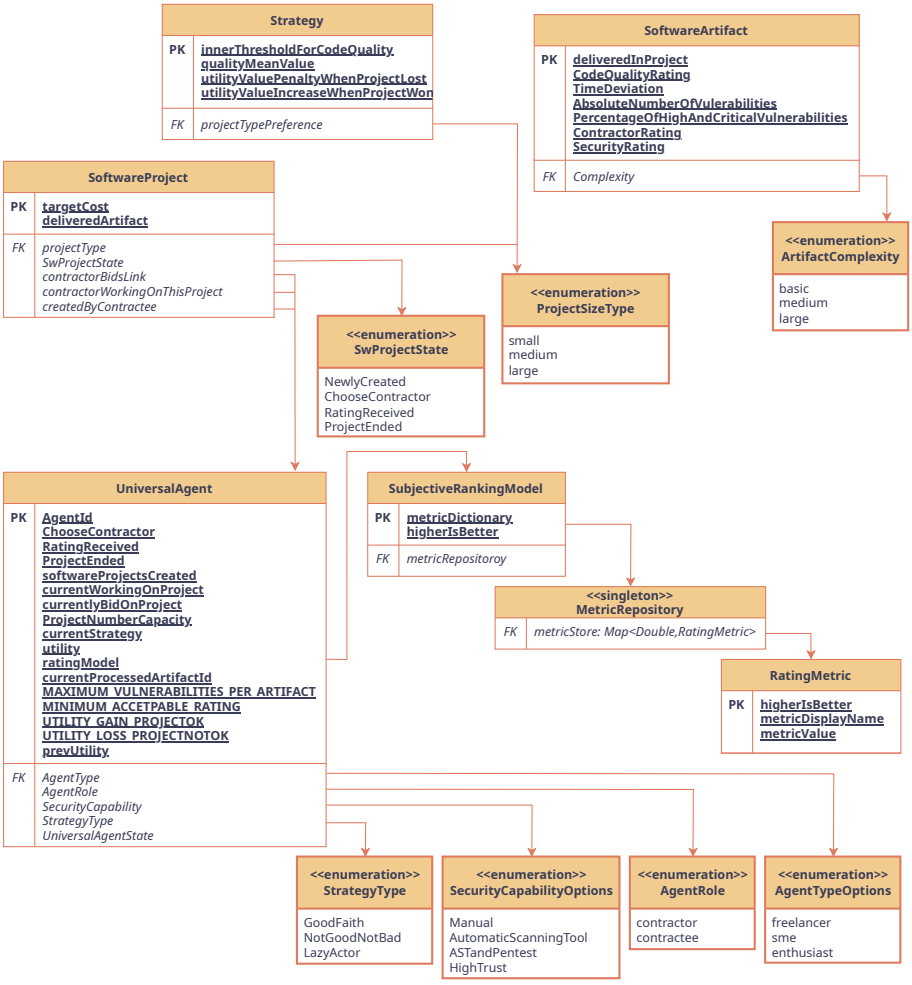


Fig. 4: The simulation data model

strategies. The initial allocation of these strategies as well as other agent characteristics are given by statistical distributions and can be adapted freely to suit the environment to be modeled. It is possible to endow both agent types with strategies. The current implementation allows strategy choice only for contractors. These can currently employ three strategies:

1. Good faith strategy – The supplier tries its best to deliver maximum quality and tends to focus more on quality assurance and good practices.
2. Balanced strategy (not good, not bad) – The supplier tries to deliver the required quality of the artifact without too much investment. It tries to optimize its own profit margin by not overachieving, and thus maximizing its utility value. This would be the equivalent of a rational agent.
3. Bad actor strategy (lazy) – The supplier tries to game the system in attempting to do as little work as possible while maintaining good reputation. Even if some projects are lost, this has little influence.

Our model assumes a software project market with an initial number of contractors and contractees able to build supply chains up to three levels deep. In the simulation, there are three operation modes regarding the supplier selection method: price only, reputation system with one metric (quality rating), and reputation system with two metrics (quality rating and security rating). In the case of the price only method, a price offer within given limits is generated which is either accepted or rejected. Time and materials models, also common in software development projects, are not implemented in this version.

It is assumed that the strategy parameters are fixed. The experimenter may, however, choose to introduce other strategies as required. The evaluation model can be used on top of an economic simulation layer which provides additional data for each organization, e.g., profit, cash reserves, bankruptcy events, etc. This is why some advanced parameters have been included so they can be provided by another model or manually if needed. This makes it easier to accommodate different researcher backgrounds and tools while keeping the model as simple as possible.

6 Discussion and Evaluation of the Results

For the evaluation of our simulation, we have considered multiple approaches as discussed in (Janssen & Ostrom, 2006) and (Venable et al., 2012). For this artifact implementation we decided to use different experiments in order to check the correctness of the used simulation model (artificial ex post method). This method allows efficacy and usefulness to be evaluated quickly, with the downside of a higher risk for false positives. This risk is acceptable and can be further reduced by using a second means of evaluation. We plan to do this with real participants as recommended in (Janssen & Ostrom, 2006). For our experiments, indicators were selected which are mapped to stakeholder goals and to the definitions put forward in Section 1 as follows:

Resilience is expressed as a measure of market availability which is affected by disruptions (e.g., cybersecurity incidents). The latter is computed for each agent in the simulation as a percentage of total simulation time and then averaged for the entire ecosystem. We allow for multiple types of disruption events and implement a first type of disruption in the form of discovered critical software vulnerabilities.

Digital sovereignty is expressed as a measure of how much choice a contractee has in terms of suppliers. Here, we use the proxy indicator of average project assignment time. Assumption here: The more suppliers are available, the less time it takes to find a suitable one.

Fairnesss In our context, the term is synonymous with ensuring a fair and balanced distribution of resources, opportunities and risks (Arrow, Sen, & Suzumura, 2010). It is expressed as the mix of contractor types on average per simulation run. In some cases it is highly desirable to include not only the established (big) players, but also SMEs, freelancers, or startup companies. In this interpretation a more evenly distributed picture regarding roles would be better, i.e., it provides better access to opportunities for all contributor types.

The analysis of the above aspects combined with financial measures of success can, e.g., reveal the need for an additional metric or simply an exaggerated expectation. This issue is not meant to be solved computationally, but rather serve as a trigger for further governance-related talks or even automated negotiations among the participating organizations.

For the first version evaluation of the model, a synthetic dataset was created with a set of parameters partly informed by a public dataset of EU tenders (Directorate-General for Internal Market & SMEs, 2021).

A series of parameter variation experiments was conducted. In Figure 5, the indicators for resilience, digital sovereignty, and fairness were evaluated in the course of varying the initial number of agents. Initial tests show that there is a significant (t-Test, $p = .000002$) increase in quality when using code quality metrics (experiment E2) in addition to price (experiment E1). Adding another metric (security rating, experiment E3) boosts this improvement on this dimension (t-Test, $p = .021$). Regarding the total number of completed products, we see a drop from E1 to E2 ($p = .0009$), which in relation to quality can be seen as a compromise. However we notice a larger drop when using both metrics (E1 to E3, $p = -.000005$), a situation which can be plausible when the two metrics do not correlate and contractees set higher quality requirements as in our scenario. Finally, we see very small differences in average assignment times with large standard deviations so this result needs to be further investigated. Low difference assignment times may indicate low “costs” for using more advanced metrics in terms of digital sovereignty.

From a design science perspective, we argue that through this first cycle, we have made progress towards a design theory or meta-design as described by (Venable, 2006). Specifically in the areas of the design method, i.e., how to use simulation to refine a

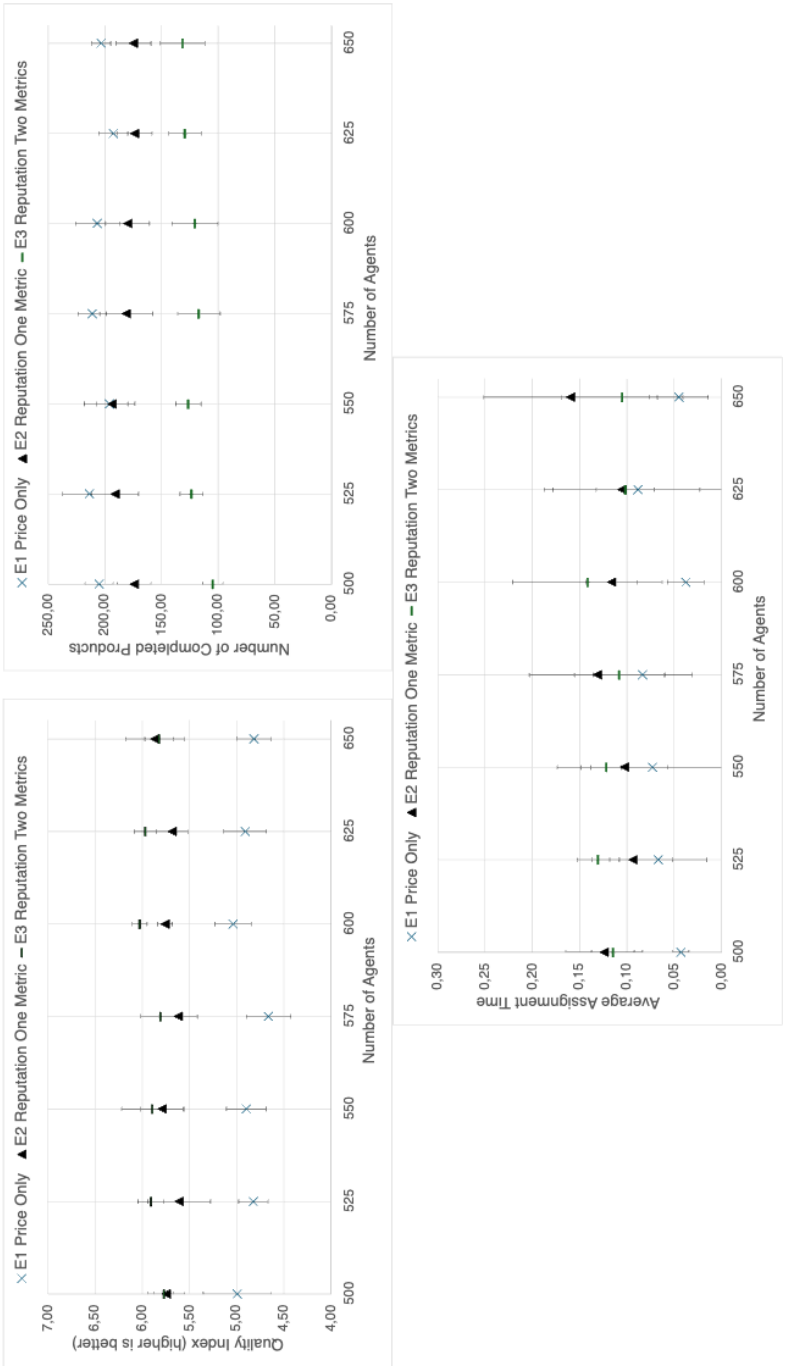


Fig. 5: Comparing execution runs E1 (using price only selection), E2 (using price and code quality) and E3 (using price, code quality, and security rating) while varying the number of agents. Time is expressed in days. All other parameters are fixed.

reputation system and by providing testable product hypotheses, i.e., the use of a specific instance of a reputation system, leads to more digital sovereignty of the organization/the entire supply chain. In further iterations and evaluations, the design and requirements can be further generalized with varying usage scenarios and kernel theory mixture.

7 Conclusion and Outlook

To answer the first research question, we set out to design a reputation evaluation system for IT supply chains with special attention paid to the aspects of digital sovereignty and resilience. The first design science research iteration presented here resulted in an artifact that can be used by stakeholders in our scenario to gradually refine their supplier selection and risk management strategies, while avoiding wasted time and costs by actually attempting to do this refinement directly in the market.

Regarding the second research question it is possible to use the simulation in a different way: Outside of our scenario, the artifact could also be used to find out market optimal conditions for, e.g., specific products or even more restrictive digital sovereignty conditions. This could be used by supply chain managers in a consortium, government, or academia.

One unexpected side effect during the research was the scope of the artifact. In the beginning we had assumed that the reputation system itself was the artifact, with the simulation only being there as a means of evaluation. We have since realized that the simulation is an essential part of the artifact, as it helps guide the refinement process for the reputation system.

While we are attempting to create an artifact useful for organizations, at this stage it is only usable with expert guidance. More work needs to be invested in refining guidance and usability if organizations are to use it directly. The issue of representing goals such as digital sovereignty, resilience, or fairness in operational terms is complex. The indicators used in this paper represent one means of doing this and may need to be adjusted after further evaluation in real-world scenarios. Also, because we are considering multiple research areas, these must be refined and more clearly separated.

Further work needs to be done in validating the model using other methods such as experiments or role-playing games, to reduce bias and avoid hidden variables. Exploratory interviews with participants with a professional background in procurement could further drive the evolution of the artifact, and may help to identify means for bridging between classical supplier management (e.g., for hardware components), and software supply chains. Some extensions are being considered for more realism, such as support for combinations of components (hardware and software), time-and-materials project work models, and usual procurement constraints such as target time frame and project budget expectations. Currently only three selection types are included in the model. Further complexity can be later added and encapsulated in specific sub-components, e.g., for representing specific legal frameworks or economic models.

The next steps in model development include developing additional simulation aspects such as technical, social, ethical, and psychological viewpoints.

Acknowledgments

This work originated in the LIONS research project. LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

etoolbox

References

Akerlof, G. A. (1970). 4. the market for 'lemons': quality uncertainty and the market mechanism. *Market Failure or Success*, *66*.

Arrow, K. J., Sen, A., & Suzumura, K. (2010). Chapter thirteen - Kenneth Arrow on social choice theory. In *Handbook of Social Choice and Welfare* (Vol. 2, pp. 3–27). Elsevier.

Baker, D. B., Kaye, J., & Terry, S. F. (2016). Governance through privacy, fairness, and respect for individuals. *eGEMs*, *4*(2).

Balci, O. (2003). Verification, validation, and certification of modeling and simulation applications. In *Winter Simulation Conference* (Vol. 1, pp. 150–158).

Beese, J., Haki, M. K., Aier, S., & Winter, R. (2019). Simulation-based research in information systems: epistemic implications and a review of the status quo. *Business & Information Systems Engineering*, *61*, 503–521.

Benson, A., Sojourner, A., & Umyarov, A. (2020). Can reputation discipline the gig economy? experimental evidence from an online labor market. *Management Science*, *66*(5), 1802–1825.

BSI. (2024). *Technical Guideline TR-03183: Cyber Resilience Requirements for Manufacturers and Products - Part 2: Software Bill of Materials (SBOM).* Retrieved from `https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TR03183/BSI-TR-03183-2.html`

Chandler, D., & Coaffee, J. (2017). *The Routledge Handbook of International Resilience*. Routledge.

Claussen, J., Khashabi, P., Kretschmer, T., & Seifried, M. (2018). Knowledge work in the sharing economy: What drives project success in online labor markets? *SSRN 3102865*.

Directorate-General for Internal Market, E., Industry, & SMEs. (2021). *Tenders Electronic Daily (TED) (csv subset) – public procurement notices [DataSet].* Retrieved 21.01.2024, from `https://data.europa.eu/data/datasets/ted-csv`

Espinha Gasiba, T., Lechner, U., Pinto-Albuquerque, M., & Méndez, D. (2021). Is secure coding education in the industry needed? an investigation through a large scale survey. In *2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)* (p. 241-252). doi: 10.1109/ICSE-SEET52601.2021.00034

Ferreira, L., & Borenstein, D. (2011). Normative agent-based simulation for supply chain planning. *Journal of the Operational Research Society*, *62*(3), 501–514.

Fertik, M., & Thompson, D. (2015). *The reputation economy: How to optimise your digital footprint in a world where your reputation is your most valuable asset*. Hachette UK.

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., & Wendeborn, T. (2022). Towards a Layer Model for Digital Sovereignty: A Holistic Approach. In *Proceedings of the 17th International Conference on Critical Information Infrastructures Security (CRITIS 2022).* Springer. doi: inpreparation

Gandini, A., & Gandini, A. (2016). Reputation, the social capital of a digital society. *The Reputation Economy: Understanding Knowledge Work in Digital Society*, 27–43.

Gasiba, T., Lechner, U., Albuquerque, M., & Fernandez, D. (2020, 12). Awareness of secure coding guidelines in the industry - a first data analysis. In *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)* (p. 345-352). doi: 10.1109/TrustCom50675.2020.00055

Ghadimi, P., & Heavey, C. (2013). A review of applications of agent-based modelling and simulation in supplier selection problem. In *2013 8th EUROSIM Congress on Modelling and Simulation* (pp. 101–107).

Griffiths, N. (2005). Task delegation using experience-based multi-dimensional trust. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems* (p. 489–496). New York, NY, USA: Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/1082473.1082548` doi: 10.1145/1082473.1082548

Gussek, L., & Wiesche, M. (2023). IT Professionals in the Gig Economy: The Success of IT Freelancers on Digital Labor Platforms. *Business & Information Systems Engineering*, *65*(5), 555–575.

Hendrikx, F., Bubendorfer, K., & Chard, R. (2015). Reputation systems: A survey and taxonomy. *Journal of Parallel and Distributed Computing*, *75*, 184–197.

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*(2), 4.

Iosif, A.-C., Gasiba, T., Zhao, T., Lechner, U., & Albuquerque, M. (2022, 01). A large-scale study on the security vulnerabilities of cloud deployments. In *International Conference on Ubiquitous Security (UbiSec 2022* (p. 171-188). Springer. doi: 10.1007/978-981-19-0468-4_13

Irissappane, A. A., Jiang, S., & Zhang, J. (2012). Towards a comprehensive testbed to evaluate the robustness of reputation systems against unfair rating attack. In *UMAP Workshops* (Vol. 12).

Janssen, M. A., & Ostrom, E. (2006). Empirically based, agent-based models. *Ecology and society*, *11*(2).

Li, Z., Lim, L., Chen, X., & Tan, C. S. (2015). Supplier selection decision-making in supply chain risk scenario using agent based simulation. In *2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* (pp. 900–904).

Lin, M., Liu, Y., & Viswanathan, S. (2018). Effectiveness of reputation in contracting for customized production: Evidence from online labor markets. *Management Science*, *64*(1), 345-359. doi: 10.1287/mnsc.2016.2594

Liu, L., & Munro, M. (2012). Systematic analysis of centralized online reputation systems. *Decision Support Systems*, *52*(2), 438–449.

Macal, C. M. (2016). Everything you need to know about agent-based modelling and simulation. *Journal of Simulation*, *10*, 144–156.

Maedche, A., Gregor, S., Morana, S., & Feine, J. (2019). Conceptualization of the problem space in design science research. In *Extending the Boundaries of Design Science Theory and Practice: 14th International Conference on Design Science Research in Information Systems and Technology, DESRIST 2019, Worcester, MA, USA, June 4–6, 2019, Proceedings 14* (pp. 18–31).

MITRE. (2024). *CVE Vulnerability Database.* Retrieved 21.01.2024, from `https://www.cve.org/`

Mohammed, A., Lopes De Sousa Jabbour, A. B., Koh, L., Hubbard, N., Chiappetta Jabbour, C. J., & Al Ahmed, T. (2022, December). The sourcing decision-making process in the era of digitalization: A new quantitative methodology. *Transportation Research Part E: Logistics and Transportation Review*, *168*, 102948. Retrieved 2023-02-27, from `https://linkinghub.elsevier.com/retrieve/pii/S1366554522003258` doi: 10.1016/j.tre.2022.102948

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36*(1), 48–49.

Neuhaus, S., Zimmermann, T., Holler, C., & Zeller, A. (2007). Predicting vulnerable software components. In *Proceedings of the 14th ACM conference on Computer and communications security* (pp. 529–540).

Onggo, B. S., & Foramitti, J. (2021). Agent-based modeling and simulation for business and management: a review and tutorial. In *2021 Winter Simulation Conference (WSC)* (pp. 1–15).

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.

Pohle, J., & Thiel, T. (2021). Digitale Souveränität-Von der Karriere eines einenden und doch problematischen Konzepts. In *Der Wert der Digitalisierung: Gemeinwohl in der digitalen Welt* (pp. 319–340). Bielefeld: transcript Verlag.

Prester, J., & Wagner, G. (2021). Contracting decisions on digital markets for knowledge work services: A qualitative systematic review. In *ICIS 2021 Proceedings. 6.*

Rajesh, R., & Ravi, V. (2015). Supplier Selection in Resilient Supply Chains: A Grey Relational Analysis Approach. *Journal of Cleaner Production*, *86*, 343–359.

Rauf, I., Lopez, T., Tun, T., Petre, M., & Nuseibeh, B. (2023). Security in online freelance software development: A case for distributed security responsibility. *arXiv preprint arXiv:2307.06066*.

Rauf, I., Petre, M., Tun, T., Lopez, T., & Nuseibeh, B. (2023). Security thinking in online freelance software development. , 13–-24. doi: doi.org/10.1109/ICSE-SEIS58686.2023.00008

Rein, G. L. (2005). A Reference Model for Designing Effective Reputation Information Systems. *Journal of Information Science*, *31*, 365 - 380. Retrieved from `https://api.semanticscholar.org/CorpusID:10008098`

Resnick, P., Kuwabara, K., Zeckhauser, R., & Friedman, E. (2000). Reputation systems. *Communications of the ACM*, *43*(12), 45–48.

Rosson, M. B., & Carroll, J. M. (2012). Scenario-Based Design. In J. A. Jacko (Ed.), *The Human-Computer Interaction Handbook* (pp. 1105–1124). CRC Press.

Rouzafzoon, J., & Helo, P. (2016). Developing service supply chains by using agent based simulation. *Industrial Management & Data Systems*.

Ruohomaa, S., Kutvonen, L., & Koutrouli, E. (2007). Reputation management survey. In *The Second International Conference on Availability, Reliability and Security (ARES'07)* (pp. 103–111).

Sargant, R. (2005). Verification and validation of simulation models. In *Proceedings of the 2005 Winter Simulation Conference (WSC)* (pp. 130–143).

Sucky, E. (2013). *Koordination in Supply Chains: Spieltheoretische Ansätze zur Ermittlung integrierter Bestell-und Produktionspolitiken*. Springer.

Swaminathan, J. M., Smith, S. F., & Sadeh, N. M. (1998). Modeling supply chain dynamics: A multiagent approach. *Decision Sciences*, *29*(3), 607–632.

Torres-Sanchez, E. M., Saucedo-Martinez, J. A., Marmolejo-Saucedo, J. A., & Rodriguez-Aguilar, R. (2023). Multi-criteria Decision-Making for Supplier Selection Using Performance Metrics and AHP Software. A Literature Review. In *Smart Applications with Advanced Machine Learning and Human-Centred Problem Design* (Vol. 1, pp. 735–743). Cham: Springer International Publishing. doi: 10.1007/978-3-031-09753-9_56

United Nations. (2016). *Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction* (Tech. Rep.). United Nations. Retrieved from `https://www.preventionweb.net/files/50683_oiewgreportenglish.pdf`

Vavilis, S., Petković, M., & Zannone, N. (2014). A reference model for reputation systems. *Decision Support Systems*, *61*, 147–154. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0167923614000256` doi: https://doi.org/10.1016/j.dss.2014.02.002

Venable, J. (2006, 01). The role of theory and theorising in design science research. *First International Conference on Design Science Research in Information Systems and Technology*.

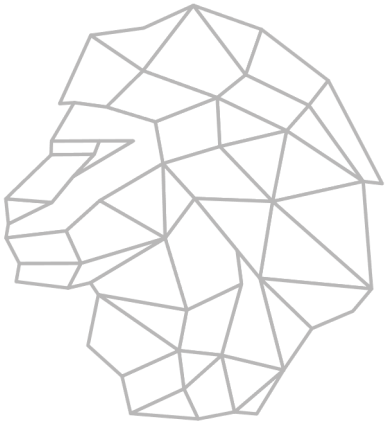
Venable, J., Pries-Heje, J., & Baskerville, R. (2012). A Comprehensive Framework for Evaluation in Design Science Research. In K. Peffers, M. Rothenberger, & B. Kuechler (Eds.), *Design Science Research in Information Systems. Advances in Theory and Practice* (pp. 423–438). Berlin, Heidelberg: Springer Berlin Heidelberg.

Vosooghidizaji, M., Taghipour, A., & Canel-Depitre, B. (2020). Supply chain coordination under information asymmetry: a review. *International Journal of Production Research*, *58*(6), 1805–1834.

Wagner, G., Prester, J., & Paré, G. (2021). Exploring the boundaries and processes of digital platforms for knowledge work: A review of information systems research. *The Journal of Strategic Information Systems*, *30*(4), 101694. Retrieved from `https://www.sciencedirect.com/science/article/pii/S096386872100041X` doi: https://doi.org/10.1016/j.jsis.2021.101694

Wissuwa, F., Durach, C. F., & Choi, T. Y. (2022, November). Selecting resilient suppliers: Supplier complexity and buyer disruption. *International Journal of Production Economics*, *253*, 108601. Retrieved 2023-02-27, from `https://linkinghub.elsevier.com/retrieve/pii/S0925527322001840` doi: 10.1016/j.ijpe.2022.108601

Digital Sovereignty as a Field of Learning



LIONS Media Education

Proposing a Research Design for Investigating the Transfer of Digital Sovereignty from Serious Games to Lifelike Scenarios

Kai Weeber[1] and Manuela Pietraß[2]

Abstract: This article investigates digital sovereignty from a media pedagogical perspective. Even though digital sovereignty as a term is used more commonly in informatics and politics discourse, parallels can be drawn to media competence as a foundation of sustainable education. It is shown why the LIONS project aims to raise digital sovereignty through serious games, combining interdisciplinary perspectives, and describes the conditions to be met. We elaborate primarily on the transfer learning between serious game and real-world scenarios as well as the crucial function of framing according to Goffman (1974). This results in the recommendation of a case vignette design that may enable further insights into how digital sovereignty may be enhanced through transfer learning. The overall qualitative research design with pre-test and post-test are described as well as useful recommendations for the design of case vignettes as a stimulus.

Keywords: Media Pedagogy, Transfer Learning, Framing, Qualitative Research, Case Vignette

1 Introduction

Gaining digital sovereignty has the traits of a media educational task that is both new and old at the same time: Media education supports skills for dealing with the diverse demands of the media, and at the same time it touches on limits where their technic loses accessibility for the users. This is particularly the case where people deal with complex technologies that are difficult to understand, even for experts. In this respect, the development of digital sovereignty is a task for those who develop technical systems and for those who deal with them. The LIONS project makes it possible to achieve precisely this dual role in an interdisciplinary consortium that brings together computer science and its sub-discipline of business informatics as well as educational science and psychology. The media education sub-project addressed in this contribution is located at the interface between educational science and business informatics. The main tasks are 1) the educational interpretation of digital sovereignty on a theoretical level and 2) the possibility of a sustainable development of digital sovereignty:

Ad 1) This is a question of the investigation of concepts: Where do concepts come

[1] University of the Bundeswehr Munich, Neubiberg, kai.weeber@unibw.de
[2] University of the Bundeswehr Munich, Neubiberg, manuela.pietrass@unibw.de

from, what are the theoretical conceptions on which they are based? Where do these terms overlap? How can business informatics benefit from a pedagogical expansion of its concept of digital sovereignty? The first part of this article deals with this question.

Ad 2) The question of "sustainable" knowledge, i.e., knowledge that can be transferred from the learning context to the application context, is at the heart of empirical research. The particular challenge lies in finding a suitable method that makes it possible to investigate transfers in action and not just changes in knowledge and attitudes.

Analog and digital serious games form the application context for the acquisition of digital sovereignty in the LIONS project. They serve as a method for collecting data and imparting knowledge in the area of digital sovereignty, so that the guiding question of the media education sub-project is in the area of knowledge building with serious games and their didactic design for the development of digital sovereignty.

2 Digital Sovereignty and Media Competence – Two Terms With Different Traditions

The term "digital sovereignty" is being used more and more frequently in academia and by the public. It aims to show that, and how, the use of digital media can be carried out despite their systemic superiority, e.g., by attaching an invisible, digital shadow to us. The demands placed on this competence are therefore high, as it requires the consistent inclusion of the technical side, which is associated with the term "digital sovereignty." The term initially refers to the overarching level of a state's strategic digital autonomy:

The "term digital sovereignty [...] almost always refers to the digital dimension of strategic autonomy, i.e., the ability to decide and act autonomously on the essential digital aspects of our longer-term future in the economy, society, and democracy. This concerns the use and structuring of digital systems themselves, the data produced and stored in them, and the processes depicted as a result" (Moerel & Timmers, 2021, p. 8).

However, the term is also used to refer to a digitally sovereign individual and is thus defined as the "ability to act and make decisions in a self-determined manner in the digital space" (Bundesministerium für Wirtschaft und Klimaschutz, 2017, p. 6-7). The overarching reference to both governmental/international institutions and organizations as well as to individuals, while their different aspects of self-determination are expressed in the "Layered Model of Digital Sovereignty" by Fries et al. (2023) (see Chapter 1 in this volume).

Couture and Toupin already bring all levels together: "Whereas state sovereignty relates to a collective structure, the use of sovereignty may refer to an abstract 'we' of

civil society, while it can also relate to the individual such as the one that can be called upon to use free and open-source software or encryption technologies to protect oneself" (Couture & Toupin, 2019 p. 2316).

In the LIONS project, the term "digital sovereignty" is transferred from computer science to pedagogy, where the non-technical aspects of media competence play a role, and therefore the discourse on media competence, in which digital sovereignty can be interpreted in terms of its subjectivity. It should not only enable users to understand the technical and systemic aspects of media, but also to use media and its content for their own orientation in everyday life and personal development in a self-determined way and to use media productively: in other words, to use it for their own communication needs.

Of course, the technical dimension of skills requirements also plays an important role in the media competence discourse. It always becomes dominant when new media technologies appear. For example, the questioning of the relationship between media reality and the real world was a central topic in the case of television. The film image, in its perceived resemblance to reality, had to be understood as an image of the world, not as its identical reproduction: were pictures an identical reproduction of the world, they would lose their capability to tell about the world, and would only show it. To this day, one component of active media work is to be able to recognize how reality can be simulated in the moving image by acquiring the camera technology and studying how it is used for building meaning (cf. Lutz et al., 2023).

In historical retrospect, teaching how to use a film camera seems like an easy educational task. However, film cameras were initially expensive and difficult to operate and filming required expert knowledge. A future retrospective of the present day may prove to be similar.

Based on critical theory, the possibility of teaching media literacy also came into focus in the discourse of pedagogy with regard to its limits: The media are part of an economically and technically operating system. Pedagogy encounters disciplinary boundaries here, and therefore has to conduct interdisciplinary research. Moreover, it has to be located on the edge of professional practice in order to learn from it what should be taught. Media ethics also came onto the scene, as producers also have a responsibility. Applied to the digital medium of the internet, the General Data Protection Regulation and the obligation of platforms to remove problematic content are measures to regulate use where competence is no longer sufficient. Applied to the area of IT security, this means, for example, that employees in companies outside of IT departments cannot be expected to comprehensively prevent all security threats or repair damage. Rather, the area of responsibility of these specialists relates to operational tasks that also contribute to the company's success.

The LIONS project offers an ideal context for pedagogy to approach and understand computer science expertise in order to carry out interdisciplinary studies on the development of a concept of digital sovereignty and its communication. Fields of application are commercial enterprises and governmental organizations where awareness of a security-oriented approach to digital technology is essential. In a narrower sense, this concerns critical infrastructure facilities and companies that have particularly complex IT systems due to their digital business model. In a broader sense, however, interdependencies in supply chains mean that almost all companies are involved in IT security issues.

In order to develop individual digital sovereignty for the individual use of digital technology in the workplace, the LIONS project is researching serious games as a medium for imparting skills, but also for gaining knowledge about security gaps.

Similar to the above-mentioned professional use of cameras in film, IT specialists in companies implement technical solutions against security risks that are beyond the competence of non-technical employees. At the same time, IT systems can never eliminate all threats, which requires employees to act competently with regard to IT security. IT security awareness training measures are aimed at realizing these individual competencies. According to Hänsch and Benenson (2014), security awareness comprises firstly: the perception and recognition of potential sources of danger, secondly: knowledge of technical concepts and processes with regard to security incidents, and thirdly: the learning of behavioral patterns that are appropriate for preventing and responding to security incidents. Security awareness is therefore an important prerequisite for digital sovereignty: on the one hand for the individual, who should be able to make self-determined decisions in the workplace, and on the other hand for the company, whose sovereignty would be restricted by IT security incidents. This is where the LIONS project comes in with serious games as innovative educational measures.

3 Foundations of Learning with Serious Games: The Inter-World Transfer of Knowledge from the Game to Reality

The research objective of LIONS meets the growing research interest in the use of serious games in vocational training to promote IT security (Prümmer et al., 2024). Various explanatory approaches speak in favor of the effectiveness of games in educational contexts. From the participants' perspective, learning processes take place in a more authentic learning environment compared to classroom settings, and knowledge is acquired in a situated way (Dishon, 2021). Game elements would also increase motivation to learn (Plass et al., 2015). In this context, further reference should be made to the work of Klimmt (2009), which provides a comprehensive overview of the theoretical explanatory relationships between serious games and learning success.

However, the application contexts in which "serious games" can meet this educational requirement must be carefully documented. Ultimately, the promotion of security awareness, and thus autonomous action in the sense of digital sovereignty, will not be achieved if people do not use serious games in such a way that educational processes reach into the world of their everyday thoughts and actions, in everyday working life, and thus lead to valuable experiences. This is a problem that has already been discussed in detail in media education and where there is still a need for further development at both a theoretical and methodological level, namely the question of the transfer of newly acquired knowledge to contexts that differ from the learning situation. In existing studies, transfer predominantly methodologically deviates from theoretical concepts and is limited to objective differences in knowledge and subjectively perceived differences in attitude and intention (Okunlula, 2023). For this reason, there is a need for educational science methods that are suitable for capturing the breadth of transfer, such as the concept of security awareness. At a theoretical level, frame analytic approaches provide promising explanations for this, which are presented in the first part. In a second step, existing methodological approaches are enriched on the basis of the frame analytic understanding of learning transfers. In a final step, a method will be proposed that can be used in concrete studies to better understand learning transfers in the context of serious games.

The reference to reality of serious games can be derived from analogous characteristics of the concept of games in general. From the fact that games fulfill a double representational function by iconically referencing real-world objects on the one hand and expressing a certain event or activity as a whole on the other, there is a connection to sign systems outside the game (Salen & Zimmerman, 2004). However, despite this relationship to reality, games do not take on an informative, depictive function. This property of games to refer to meanings beyond reality has already been dealt with transdisciplinarily. Juul (2011) describes games as "half-real" in the sense that the statements they contain can be described as both true and untrue, depending on the referenced framework. For example, a video game character who is named as an American does not represent a real person from the USA; within the game world, however, the statement about this character is true. Erving Goffman (1974) describes the delimitation of game actions as the "keying" (p. 45) of primary frameworks. Primary frameworks are not socially contextualized, while all other frames are contextualized in typical social communication situations (Pietraß, 2012). According to Goffman, frames are introduced by cues that mark the beginning and end of an interaction as "taking place in this frame," but which can also be used within frames and serve as a cue to the frame edge.

Johan Huizinga (1987) also delineates game as a world of its own, and Salen and Zimmerman assigned the term "magic circle" (2004) to this concept. As a consequence, "possibility spaces" open up within this clearly defined play area (Pietraß, 2018): Individuals can practice contingency under conditions that deviate from reality

and are usually clearly identifiable. However, in reality, i. e., under different situational conditions, tried-and-tested possible actions are initiated by different intentions, and lead to different consequences, than in the game. Beyond the functional limitation of the transfer of game actions, individuals are able to use “mediality awareness” (Pietraß, 2018) to establish relationships between game and reality and to reflect on the transfer possibilities of knowledge and behavior from game to the everyday world. The fact that the transfer from game to reality can take place is linked to extensive conditions and must be distinguished from learning transfers within a game. In the context of his game pedagogical transfer model, Fritz (2005) refers to the former as "intra-world transfers" and the latter as “inter-world transfers.” Only when games have structural links to real-world contexts in which their use is embedded, such as personal interests or experiences of players or social conditions in the form of norms or economic demands, may inter-world transfer occur (Fritz, 2011). With regard to learning opportunities and didactic activities, Dohn (2021) has identified five different levels of situational characteristics on the basis of which individuals perceive learning situations and application situations and make decisions regarding the extent to which knowledge can be applied to other contexts and the way in which the transformation of this knowledge is required:

- Domain level: thematic reference of the activity, content knowledge
- Activity level: activities to be carried out at a concrete level
- Life-setting level: surrounding life context that determines the purpose of the activity at activity level
- Societal structure level: organizational unit to which the named life framework level is assigned and which makes social demands/norms
- Cultural practices level: manners, cultural norms, and archetypal understandings that apply across organizations and may be anthropologically determined

Going beyond Fritz’s approach, these relationships between reality and game can be differentiated according to different levels of competence in moral judgment (Pietraß 2018), a question that could play a particularly important role in the transfer of knowledge to reality. This is because knowledge of inner perpetrators can be learned while playing (Hofmeier, 2024), which could make it possible to harm a company.

However, the first question to ask is how it is even possible to create transfers. A serious game deviates from reality in that the players experience differences on a life-setting level. An example related to the topic of serious games from LIONS: Although security-related concepts (e.g. passwords) may be relevant in a game as well as in a real-life situation and are related to similar activities (e.g. securing end devices), the meanings of the concept of “password” and securing IT devices as well as the associated intentions of the actors differ in the game and in reality. According to Engle (2006), differences in meaning can be bridged through targeted forms of interaction,

e.g., references to time periods before and after the game or the transfer of responsibility for the effective use of gaming experiences in other contexts, which is referred to as "expansive framing" (Engle, 2006). For example, learners understand that the artifact of a USB stick, which exists in both the real world and the game world, can have a harmful effect in both worlds and not just in the game. In addition, learners are given responsibility for translating this knowledge into action in the real world.

Given the interrelated framing of serious games and real application contexts, the question now arises as to how the transfer processes to be expected in this case are to be recorded and thus observed. According to the understanding of inter-world transfers, the following processes could be recorded: "We could investigate what feelings the player had before the game, in other words with what feelings the player entered the virtual world. In a second step we would examine what transformation processes these feelings were exposed to in the virtual world. Finally, in a third step we would try to establish what feelings were transferred from the virtual world into the real world, in what manifestation they appear here and what transformation processes they are subject to" (Fritz, 2005, p. 96). From this, it can be deduced that in order to determine inter-world transfers, it is necessary to record focused constructs – in the case of the LIONS research project, for example, IT security awareness – before and after a game intervention on the one hand, and to observe framing-related activities during the serious game itself on the other. However, the latter method proves to be challenging. If the participants are observed during the game, their game actions have consequences for reality outside the game. Initially, this also seems to apply to expansive framing. However, the quality of the relationship between game actions and consequences for reality outside the game differs: while participants should experience empowerment in other situations through game experiences by means of expansive framing, which has positive connotations from a subjective perspective, the observation and evaluation of game actions can also lead to the experience of negative emotions such as shame, as in the context of a subsequent analysis. As a result, the participants would no longer perceive the serious game as a game in the sense of a space of possibility. But even without observing the game, the other part of the procedure described – the application-oriented recording of "centers of focus," i.e., "the features, regularities, properties, or conceptual objects to which individual[s] [...] attend" (Lobato et al., 2012, p. 439) – can provide information about inter-world transfers. Obviously, not all changes in thinking, perception, etc. can be attributed to participation in the game. Rather, there must be a clear substantive link between the recorded transformations and the player's life. The addition of the construct of "transformative experience" according to Pugh et al. (2017) is helpful here. Based on the concept of experience according to Dewey (Pugh, 2011), a completed learning transfer is reflected in the application in everyday situations. This application happens:

- from motivated use, i.e., participants would draw on educational experiences during the game without other people prompting them to do so,

- by expansion of perception, i.e., participants recognize new safety-relevant aspects, for example,
- due to experiential value, i.e., participants value the self-motivated application of new patterns of perception as meaningful, for example through benefits for their professional activities or social recognition.

If participants were to exhibit an expanded perception of IT security awareness in work-related situations, associate these with game experiences without being asked, and rate them as subjectively valuable, this would be a clear indicator of completed transfers from the serious game. This speaks for the methodological appropriateness of recording game-related transformation processes using a qualitative pre-test/post-test design. Although the approach according to Pugh et al. (2017) offers impulses for the development of the research design and operationalization, the concrete methodology of this study implies challenges with regard to feasibility. A delayed post-test several weeks after the intervention is significantly more time-consuming for participants and is met with resistance, particularly from contact persons from the business world. However, as Lobato et al. (2012) demonstrate, transfer processes can also be recorded with measurement points shortly after the learning activity. It remains important that the interview situation is open to interpretations by the participants, including those that deviate from the intended learning objectives (Lobato, 2008). In the case of safety awareness and safety-relevant behavior, this openness of the survey situation is additionally influenced by the moral significance of this topic area. The fact that safety awareness in particular is subject to strong behavioral norms can be empirically proven and justified by the fact that a deviation from safety standards can, under certain circumstances, result in the failure of essential structures and high consequential costs for a collective (Herath et al., 2018; Myyry et al., 2009)

4 Qualitative Case Vignette Design as a Research Method

Consequently, there is a need for an instrument that depicts situational aspects of the professional context, that enables the open-ended depiction of transfer processes that have taken place, taking into account any socially desirable answers, and at the same time demands an acceptable number of resources for both participants and researchers. In the social sciences, the case vignette method is used for such phenomena in order to obtain contextualized statements from people on a morally relevant topic (Finch, 1987). According to Skilling & Stylianides (2019), case vignettes are fictitious descriptions of realistic situations that are intended to stimulate statements on perceptions, opinions, decisions, etc. from the people addressed. Since the people in a case vignette are not observed in a real situation themselves, but rather maintain a certain distance by talking about a fictitious, realistic situation, they are more likely to deviate from socially desirable statements (Herskovits, 1950; Skilling & Stylianides, 2019; Spalding & Phillips, 2007). Transferred to the lived world model according to Fritz

(2005), the transfer of perception patterns and attitudes from the serious game to a case vignette would not mean a transfer from the game world to the real lifeworld, but would mean a transfer to the mental world. By processing such a case vignette before and after participation in a serious game, inter-world transfers could be recorded.

How should case vignettes be designed and implemented? According to Anselmann and Mulder (2022), case vignettes always contain a description of the situation, which is often narrative in nature, but without being explicitly formulated subjectively and thus anticipating interpretations. Further central quality criteria from the existing literature that a case vignette must fulfil are:

- Realism – representations must appear internally consistent to the participants and correspond to their perception of their lifeworld (Anselmann & Mulder, 2022; Converse et al, 2015; Jasinski et al, 2021; Sheringham et al, 2021; Skilling & Stylianides, 2019)
- Openness – while presentations should contain sufficient information, they should be formulated generically enough to allow for individual understanding by participants and thus variance in performance (Anselmann & Mulder, 2022; Lichand et al., 2023; Skilling & Stylianides; 2019)
- Debriefing – the research topic should be kept as secret as possible until the end of the interviews in order to minimize desired effects and reactance (Lichand et al, 2023; Sheringham et al, 2021)
- Visualization – purely textual representations emphasize already valent situational features externally and thus exert too strong an influence on the subjective interpretation of the participants (Jasinski et al., 2021; Sheringham et al., 2021; Skilling & Stylianides, 2019)

In addition to these quality characteristics, which should be fulfilled in any case so that the recorded data can be validly evaluated, further design decisions must be made, which depend on specific research questions. Especially in the early evaluation phase of the game, in which the changed behavior of the participants is to be described openly, open interview questions to the participants are more suitable than closed response specifications (Anselmann & Mulder, 2022; Skilling & Stylianides, 2019). Even if this increases the evaluation effort, the criterion validity of the interview data is higher, as decisions can be mapped more closely than in work contexts (Converse et al., 2015). Since patterns of perception and behavioral intentions are particularly relevant with regard to the serious games in the LIONS project, the instructions and any follow-up questions should also focus on what the participants notice about the situation presented and what they would do in such a situation.

In addition, it must be decided whether the participants should work on the same case vignette before and after the game, or on different ones. According to Sheringham et

al. (2021), specific details have a stronger influence on participants' judgments if only one case vignette or, in experimental designs, one vignette per content feature is used. On the other hand, changes can be observed more directly with repeated measurements using the same case vignette (Skilling & Stylianides, 2019). However, it is important to consider carry-over effects, i.e., which changes are caused by repeated exposure alone. For this reason, similar to Lobato et al. (2012), triangulation of the interview data with the game should take place in order to identify which changes go beyond idiosyncratic occurrences.

Linked to the use of the same or different vignettes is also the question of how many case vignettes a person should process. One argument in favor of working on several case vignettes is that this makes it possible to identify variances in the judgment of individual people that are related to the situation descriptions and not to the game (Sheringham et al., 2021). However, it is warned that if too many case vignettes are presented in succession, the participants' attention wanes and the reception of the case vignettes is time-consuming, which may shorten the practicable duration of the interview (Jasinski et al., 2021; Renta-Davids et al., 2020). From a research economics perspective, the number of case vignettes should be considered in relation to the number of interviewees. The more data generated by additional case vignettes, the less capacity remains for the evaluation of additional participants. For this reason, Lichand et al. (2023) recommend using variants of case vignettes rather sparingly, especially in early surveys, and instead conducting 10 to 15 interviews with different people per characteristic of interest in order to achieve sufficient variance in personal characteristics. An alternative to using several case vignettes would be to use a longer narrative that contains several events or situations. Firstly, contextual information does not have to be repeated several times, which avoids redundancies and supports the participants' attention (Skilling & Stylianides, 2019). Secondly, dynamic representations, i.e., evolving narratives, appear more authentic and plausible to participants (Jasinski et al., 2021).

Ultimately, case vignettes can also be written from different narrative perspectives. Skilling and Stylianides (2019) recommend using not the participant's perspective as the deictic center for sensitive or morally charged topics, but instead the perspective of a fictional character or a neutral narrator. However, it cannot be assumed that all participants will perceive the case vignette as equally morally salient. One possible solution to this would be to combine several perspectives in one vignette, as suggested by Spalding and Phillips (2007). During the interview, the participants could first evaluate the case on behalf of the fictional character. Afterwards, the person conducting the interview could ask what the participants themselves would think and do in the situation. However, even if participants share their personal perspective on the case vignette without any recognizable inhibitions, it should always be borne in mind that this perspective does not necessarily correspond to the real-life perspective (Jasinski et al., 2021). The reason for this is that the researcher's perspective influences the content of the vignette, and thus also indirectly influences the participants'

applied perspective.

Accordingly, the existing literature strongly recommends careful validation of the designed case vignettes (Anselmann & Mulder, 2022; Converse et al, 2015; Jasinksi et al, 2021; Lichand et al, 2023; Sheringham et al, 2021; Skilling & Stylianides, 2019). When determining the content focus, a theoretical connection between the constructs of interest and the events and situational characteristics in the case vignettes should be comprehensible. In our case, this would mean that although the situation should contain elements such as IT security precautions (e.g. signatures) and potential threats (e.g. phishing mail), these should be homogeneously embedded in operational work processes in order to adequately address the construct of awareness. In addition, the case vignette should also be based on experience reports. This can be done through the involvement of practical professionals or, for example, through resources like records or forum and social media posts. Finally, the carefully designed vignette should be subjected to pretests in which particular attention is paid to the participants' judgments of realism. Feedback loops contribute to the successive improvement of the instrument.

A qualitative pre-test/post-test design with carefully designed case vignettes represents a promising innovation both for educational transfer research and for information systems research on protection and security, as implied by the successful use of case vignettes in other disciplines such as medical didactics (Haan & van der Voort, 2018). In further developments, the research design developed as part of LIONS can contribute to generating ecologically valid findings on educational measures, also in organizational contexts in the future.

Acknowledgments

This work originated in the LIONS research project. LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

References

Anselmann, V. & Mulder, R. H. (2022). Using the Vignette Technique to Increase Insight into Professional Development at Work. In M. Goller, E. Kyndt, S. Paloniemi & C. Damşa (Eds.), *Professional and Practice-based Learning. Methods for Researching Professional Learning and Development* (Vol. 33, pp. 71-86). Springer International Publishing. https://doi.org/10.1007/978-3-031-08518-5_4

Bundesministerium für Wirtschaft und Klimaschutz (2017). *Kompetenzen für eine digitale Souveränität* [translation]. FZI Research Center for Information Technology, Accenture GmbH & Bitkom Research GmbH (Ed.). https://www.bmwk.de/Redaktion/DE/Publikationen/Studien/kompetenzen-fuer-eine-digitale-souveraenitaet.html

Converse, L., Barrett, K., Rich, E. & Reschovsky, J. (2015). Methods of Observing Variations in Physicians' Decisions: The Opportunities of Clinical Vignettes. *Journal of general internal medicine*, *30*(Suppl 3), 586-594. https://doi.org/10.1007/s11606-015-3365-8

Couture, S., & Toupin, S. (2019). What does the notion of "sovereignty" mean when referring to the digital? *New Media & Society*, *21*(10), 2305-2322. https://doi.org/10.1177/1461444819865984

Dishon, G. (2021). The designability paradox: rethinking authenticity and situatedness in educational video games. *Educational Technology Research and Development*, *69*(2), 497-513. https://doi.org/10.1007/s11423-021-09992-5

Dohn, N. B. (2021). Conceptualizing knowledge transfer as transformation and attunement. *Frontline Learning Research*, *9*(3), 13-30. https://doi.org/10.14786/flr.v9i3.733

Engle, R. A. (2006). Framing Interactions to Foster Generative Learning: A Situative Explanation of Transfer in a Community of Learners Classroom. *Journal of the Learning Sciences*, *15*(4), 451-498. https://doi.org/10.1207/s15327809jls1504_2

Finch, J. (1987). The Vignette Technique in Survey Research. *Sociology*, *21*(1), 105-114. https://doi.org/10.1177/0038038587021001008

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., Wendeborn, T. (2023). Towards a Layer Model for Digital Sovereignty: A Holistic Approach. In B. Hämmerli, U. Helmbrecht, W. Hommel, L. Kunczik, & S. Pickl (eds): *Critical Information Infrastructures Security. CRITIS 2022. Lecture Notes in Computer Science* (119 - 139). Springer. https://doi.org/10.1007/978-3-031-35190-7_9

Fritz, J. (2005). How Virtual Worlds Affect Us: On the Structure of Transfers from the Media World to the Real World. In G. M. Buurman (Ed.), *Total Interaction* (pp. 95-121). Birkhäuser-Verlag. https://doi.org/10.1007/3-7643-7677-5_8

Fritz, J. (2011). *Wie Computerspieler ins Spiel kommen: Theorien und Modelle zur Nutzung und Wirkung virtueller Spielwelten* [translation]. Vistas.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harper & Row.

Hänsch, N. & Benenson, Z. (2014). Specifying IT Security Awareness. In *2014 25th International Workshop on Database and Expert Systems Applications* (pp. 326-330). IEEE. https://doi.org/10.1109/DEXA.2014.71

Herath, T., Yim, M.-S., D'Arcy, J., Nam, K. & Rao, H. R. (2018). Examining employee security violations: moral disengagement and its environmental influences. *Information Technology & People*, *31*(6), 1135-1162. https://doi.org/10.1108/ITP-10-2017-0322

Herskovits, M. J. (1950). The Hypothetical Situation: A Technique of Field Research. *Southwestern Journal of Anthropology*, *6*(1), 32-40. http://www.jstor.org/stable/3628688

Hofmeier, M. (2024). *Operation Digital Butterfly – Ein Serious-Game-basierter Ansatz zur Identifikation und Analyse von Intentionalen Bedrohungen durch Innentäter und Innentäterinnen (Malicious Insider Threats)* [dissertation, translation]. https://athene-forschung.unibw.de/148672

Huizinga, J. (1987). *Homo ludens: Vom Ursprung der Kultur im Spiel* [translation]. Rowohlt Taschenbuch Verlag.

Jasinski, L., Nokkala, T. & Juusola, H. (2021). Reflecting on the value of vignettes in higher education research: towards a preliminary typology to guide future usage. *European Journal of Higher Education*, *11*(sup1), 522-536. https://doi.org/10.1080/21568235.2021.1999841

Juul, J. (2011). *Half-real: Video games between real rules and fictional worlds* (First MIT Press paperback edition). The MIT Press.

Klimmt, C. (2009). Serious games and social change: Why they (should) work. In U. Ritterfeld, M. J. Cody & P. Vorderer (Eds.), *Serious games: Mechanisms and effects* (pp. 248-270). Routledge.

Lichand, G., Serdeira, A. d. P. & Rizardi, B. (2023). Best Practices in Testing Behavioral Mechanisms. In G. Lichand, A. d. P. Serdeira & B. Rizardi (Eds.), *Behavioral Insights for Policy Design* (pp. 115-124). Springer International Publishing. https://doi.org/10.1007/978-3-031-33034-6_9

Lobato, J. (2008). Research Methods for Alternative Approaches to Transfer: Implications for Design Experiments. In J. Y. Baek, A. E. Kelly, & R. A. Lesh (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching* (pp. 167-194). Routledge. https://doi.org/10.4324/9781315759593-20

Lobato, J., Rhodehamel, B., & Hohensee, C. (2012). "Noticing" as an Alternative Transfer of Learning Process. *Journal of the Learning Sciences*, *21*(3), 433-482. https://doi.org/10.1080/10508406.2012.682189

Lutz, K., Schemmerling, M. & Reißmann, W. (2023). Editorial: Aktive Medienarbeit in Zeiten ihrer Entgrenzung [translation]. *MERZ - Medien + Erziehung* (3). https://www.merz-zeitschrift.de/alle-ausgaben/details/2023-03-kritische-aktive-medienarbeit/

Moerel, L., & Timmers, P. (2021). Reflections on digital sovereignty. *EU cyber direct, research in focus series.* https://ssrn.com/abstract=3772777

Myyry, L., Siponen, M., Pahnila, S., Vartiainen, T. & Vance, A. (2009). What levels of moral reasoning and values explain adherence to information security rules? An empirical study. *European Journal of Information Systems*, *18*(2), 126-139. https://doi.org/10.1057/ejis.2009.10

Okunlola, J. O. (2023). Learning Transfer in the Workplace: An Insight Into the Missing Link in the Education and Training of Employees. *Studies in Learning and Teaching*, *4*(2), 349-354. https://doi.org/10.46627/silet.v4i2.241

Pietraß, M. (2012). Rahmentheorie [translation]. In B. Schäffer & O. Dörner (Eds.), *Handbuch Qualitative Erwachsenen- und Weiterbildungsforschung* (pp. 153 - 165). Barbara Budrich.

Pietraß, M. (2018). *Formen von Medialitätsbewusstsein: Relationen zwischen digitalem Spiel und Wirklichkeit am Beispiel moralischer Entscheidungen* [translation]. Nomos.

Pietraß, Manuela (2020). Das "mögliche Unmögliche" in digitalen Spielwelten: Die Hervorbringung von Wirklichkeit in neuen Zeichenkonfigurationen [translation]. In: Iske, Stefan; Fromme, Johannes; Verständig, Dan & Wilde, Katrin (eds.), *Die Kunst der Zahlen – Digitale Transformationen des Ästhetischen.* Wiesbaden: VS Springer Verlag, pp. 141-156.

Plass, J. L., Homer, B. D. & Kinzer, C. K. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, *50*(4), 258-283. https://doi.org/10.1080/00461520.2015.1122533

Prümmer, J., van Steen, T. & van den Berg, B. (2024). A systematic review of current cybersecurity training methods. *Computers & Security*, *136*, 103585. https://doi.org/10.1016/j.cose.2023.103585

Pugh, K. J. (2011). Transformative Experience: An Integrative Construct in the Spirit of Deweyan Pragmatism. *Educational Psychologist*, *46*(2), 107-121. https://doi.org/10.1080/00461520.2011.558817

Pugh, K. J., Bergstrom, C. M., & Spencer, B. (2017). Profiles of Transformative Engagement: Identification, Description, and Relation to Learning and Instruction. *Science Education*, *101*(3), 369-398. https://doi.org/10.1002/sce.21270

Renta-Davids, A.-I., Camarero-Figuerola, M. & Tierno-García, J.-M. (2020). Assessment of the Quality Education Awareness Competence of Pre-Service Educators Using Vignettes. *Sustainability*, *12*(23), 10203. https://doi.org/10.3390/su122310203

Salen, K. & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. The MIT Press.

Sheringham, J., Kuhn, I. & Burt, J. (2021). The use of experimental vignette studies to identify drivers of variations in the delivery of health care: a scoping review. *BMC medical research methodology*, *21*(1), 81. https://doi.org/10.1186/s12874-021-01247-4

Skilling, K. & Stylianides, G. J. (2019). Using vignettes in educational research: a framework for vignette construction. *International Journal of Research & Method in Education*, *43*(5), 541-556. https://doi.org/10.1080/1743727X.2019.1704243

Spalding, N. J. & Phillips, T. (2007). Exploring the use of vignettes: from validity to trustworthiness. *Qualitative health research*, *17*(7), 954-962. https://doi.org/10.1177/1049732307306187

The Relevance of the Facets of Technology Commitment for Dealing with Digital Media and Security Precautions

Isabelle Marie Sophie Haunschild[1] and Bernhard Leipold[2]

Abstract: The use of computers and digital media is well established in both private and professional everyday life and demands a sovereign approach to the new technologies. In addition to merely using the technology, successful handling should also include increased competence in its use and compliance with security guidelines to avoid risks such as malware, fraud, identity theft, etc. We therefore investigated the connections between facets of technology commitment (technology acceptance, technology competence, technology control) and the use of digital media (social media, internet, and emails) and computers, and security precautions during use. A total of 661 adults ($M = 38.60$; $SD = 14.01$) participated in a cross-sectional questionnaire study. Alongside the age variable, technology acceptance was a significant predictor of social media use (frequency, preference), while technology competence convictions were a significant predictor for security precautions. Technology acceptance and technology control moderated the negative correlation between age and social media use.

Keywords: Technology Commitment, Technology Acceptance, Technology Competence Convictions, Digital Media, Security Precautions

1 Introduction

In numerous areas of life, the use of computers and digital media is intended to support us in our private and professional tasks and help make everyday life easier. The use of digital processes (e.g., digital appointment scheduling systems to reduce waiting times for patients and optimize workflows for medical staff) is so integrated into daily life that digital use is hardly optional and switching to analog-only processes is not easily feasible. Given its relevance, being equipped with a basic knowledge of how digital technologies work (e.g., registering and logging on to an online platform or connecting a device to Wi-Fi) is important. Because these technologies are constantly evolving, one must remain up to date to ensure sovereign and continuous use (lifelong learning; Beier, 2021). A starting point is the level of awareness of gaps in one's own knowledge, attitudes, and uncertainties in dealing with technologies. The model of technology commitment (Neyer et al., 2012) describes a successful interaction with digital technologies based on attitudes such as open-mindedness towards technology, technology-related self-efficacy, and the skilled use of technology. In addition, success is also related to frequency of, and affinity for, use as well as compliance with

[1] University of the Bundeswehr Munich, Neubiberg, isabelle.haunschild@unibw.de
[2] University of the Bundeswehr Munich, Neubiberg, bernhard.leipold@unibw.de

security guidelines. In the following, we examine the three facets of *technology commitment – technology acceptance, technology competence convictions, and technology control convictions* – and their associations with the *use of digital media* (social media, internet, and emails) and *security precautions*.

The term *security precautions* refers to the self-reported precautionary measures taken when dealing with digital media, such as using passwords, carrying out updates, or being cautious browsing the web (Egelman & Peer, 2015), and not the behavior put into practice.

Social media refers to the use of websites and applications for communication, interacting, and sharing contents (e.g., the use of social networks, chat programs, or video communication programs), whereas the term *emailing/surfing* refers to writing emails and browsing the web. *Digital media* is used when both *social media* and *emailing/surfing* are being referred to.

1.1 Commitment to Technology and its Facets

As a theoretical framework in information systems research, the Technology Acceptance Model (TAM; Davis, 1989) proposes that the use of technology is influenced by two primary factors: perceived usefulness and perceived ease of use. Both of these affect the attitude towards technologies and thus their actual practical use. Extended theories (Venkatesh & Davis, 2000) include several other factors, such as social influence processes (job relevance, output quality, result demonstrability, and perceived ease of use) and cognitive instrumental processes. The underlying assumption that positive attitudes towards technologies are relevant to actual behavior provides an understanding of how and why individuals decide to use or reject technology. Neyer et al. (2012) introduced the concept of technology commitment, which includes *technology acceptance*, competence beliefs, and subjective perception of control as the three main components. While technology acceptance relates to personal interest in new technologies, *technology competence convictions* refer to an individual's self-confidence and belief in their personal ability to use technology effectively, and *technology control convictions* address an individual's perception of the extent to which technology is controllable.

Research shows a relationship between computer-related self-efficacy and safety behavior (e.g., Branley-Bell et al., 2022) as well as a relationship between technology commitment and technology use (Neyer et al., 2012). Empirical studies have shown computer self-efficacy to be negatively associated with anxiety and positively associated with willingness to use computers (Czaja et al., 2006), while computer anxiety was negatively associated with computer skills (Shah et al., 2012). In addition, we have shown that technology commitment is associated with both *security precautions* and *use of social media* (Haunschild & Leipold, 2023). Technology commitment also

moderated the negative association between age and *use of digital media*. Specifically, when technology commitment was high, older respondents used social media almost as much as younger ones. In this case, however, technology commitment was examined as an overall scale. The intercorrelations (Neyer et al., 2012) have shown that the three facets do differ from one another and address different aspects of technology commitment. The correlations ranged between .23 and .42. One aim of the present study is to examine whether they are associated with *security precautions* and *digital media use* to a different degree.

Following Davis (1989), Neyer et al. (2012) define *technology acceptance* as an attitude reflecting one's appraisal of technological progress. This emphasizes the individual's personal connection to modern technologies and primarily reflects a personal interest in new technology. We therefore expect *technology acceptance* to show a significant correlation with *digital media usage* and *security precautions*.

According to the concept of competence beliefs (Krampen, 1991), *technology competence convictions* are defined as one's anticipation of being able to handle situations that involve technology. They reflect the experience gained with familiar technologies over the course of an individual's life and the anticipated ability to adapt to new technological innovations, representing a self-concept of personal skills (Neyer et al., 2012). Neyer's items, however, suggest anxiety and a perceived lack of competence in dealing with new technology (e.g., "I'm afraid of damaging new technology rather than using it correctly"). Validation studies showed that *technology competence convictions* were negatively related to neuroticism, with scores ranging from $r = -.27$; $p < .01$ to $r = -.39$; $p < .001$. We therefore expect *technology competence convictions* in particular to be associated with anxiety-related *security precautions awareness*.

Neyer et al. (2012) define *technology control convictions* as individual anticipation of outcomes from handling technology (see also Beier, 1999; Krampen, 1991). They reflect the extent to which technology is perceived to be controllable, hence showing the expected personal influence and control over technology processes and their impact. The items used here tend more towards self-efficacy in handling new technology (e.g., "Whether I am successful in using modern technology mainly depends on me."). We thus expect *technology control convictions* to be associated with *security precautions*.

1.2 Age and the Cautious Use of Digital Media

As mentioned above, age can play a role in the association between *technology commitmen*t and both *security precautions* and *use of social media* (Haunschild & Leipold, 2023). Even though many older adults use digital media and new technology, there are age differences in terms of usage. As a generation of digital natives (Prensky, 2001), that is, growing up as part of an environment that involves computers and their

use, the younger adults tend to use digital media more frequently. In addition, the fast adoption of digital media among younger generations is further facilitated by the implementation of new technology in educational contexts. In contrast, older adults may show lower motivation, less openness to innovation, or a lack of confidence in their skills (Branley-Bell et al., 2022; Nimrod, 2017; Tyler et al., 2020). Not only have older adults been shown to be more anxious about using technological innovations and to have less confidence in their ability to learn about them (Marquié et al., 2002): they also have lower media participation or media literacy than younger adults (Chang et al., 2015; Kubicek, 2023). Thus, we expect age to be negatively associated with the *use of digital media*. We also expect *technology control convictions* to moderate the relationship between age and *digital media usage*: namely, that older people tend to use digital media more often if they show higher values in *technology control convictions*. Due to its relation to interest in new technologies, we similarly expect *technology acceptance* to moderate the relationship between age and *digital media usage*.

As to *security precautions*, studies have shown that older adults are more cautious in different domains of life (Greve et al., 2018; Rolison et al., 2014), including the use of digital technology (Branley-Bell et al., 2022). Older users have been found to protect their data more actively and to report a significantly higher awareness of online privacy than younger ones (Zeissig et al, 2017). Similarly, self-reported compliance with computer security advice (Egelman & Peer, 2015) was positively correlated with age and *technology commitment* (Haunschild & Leipold, 2023), possibly also because risk-taking behaviors vary with age (e.g., Rolison et al., 2014). Thus, we expect a positive association between *security precautions* and age.

Overall, this study will first explore whether the three facets of *technology commitment* are positively associated with the *use of digital media* and *security precautions*, i.e., the components involving the use of technology. We will then examine whether *technology acceptance*, *technology competence convictions,* and *technology control convictions* can predict these components, taking age into account. Age correlations will also be examined more closely – for example, whether older adults use digital media less than younger people and are more cautious in their *use of digital media*. Finally, we will examine whether interactions can be identified. We assume that age differences in the *use of digital media* will be smaller if *technology acceptance* and *technology control convictions* are higher among older people.

2 Materials and Methods

2.1 Participants and Procedure

Initially, 1040 participants were recruited for this cross-sectional study. They took part via seminars with students and online panel platforms (Cint AB, Bilendi GmbH)

and received either course credits or financial compensation for their participation. Of the original 1040 participants, 36 were excluded because they did not pass the quality check. The quality check consisted of a control question ("Please check the box with the number 2") embedded in the questionnaire to ensure that the participants had read the instructions properly. Fifty-two participants were removed due to extremely rapid completion (TIME_RSI > 2, indicating extremely fast completion). Of the remaining 952 participants, 176 were excluded due to their job situation (e.g., if they were unemployed, looking for work, on work disability or retired); 102 participants were excluded because they belonged to the first period of data collection, which did not include both questions concerning computer use. From the remaining total of 674 participants, 13 cases with extreme z scores (> |+- 3|) were found to be univariate outliers and were deleted. The final sample consists of 661 participants aged 19 to 69 (M_{age} = 38.60, SD_{age} = 14.01). Gender ratio was balanced (47% female). Sixty-one percent were employed full-time and 22% part-time; 17% were students. School education was mostly high: approximately 67% had a high level of education with 12 or more years of schooling (German Abitur), 26% had a medium educational level with 10 years of schooling, and almost 7% had a lower educational level, with 9 or fewer years of schooling. Slightly less than 41% had completed vocational training, and 38% had a university degree.

Prior to the analysis, we tested deviations from normality using graphical methods as well as skewness and kurtosis values. For *security precautions, technology commitment*, *use of social media*, and *emailing/surfing*, values of skewness and kurtosis were between -1 and +1 and indicate no strong deviations from normal distribution. However, a value of kurtosis of -1.25 for *age* indicate a lack of normal distribution.

Before taking part, participants were informed about the study procedure and data privacy. They had to confirm that they were at least 18 years old and give their consent to take part in the study by checking a box. Participation was voluntary. This study was approved by the ethics committee of the University of the Bundeswehr Munich.

2.2 Measures

Security precautions: The *Security Behavior Intentions Scale* (SeBIS) was used to measure security intentions (Egelman & Peer, 2015). SeBIS consists of 16 items that measure the frequency of precautionary behavior such as password generation, software updating, being cautious when surfing, and device securement. Participants were asked to rate on a five-point Likert scale (1 = "never", 5 = "always") how frequently they use each of the security precautions. Of the four subscales, the subscale device securement was not included in the analyses because of its low reliability (α= .53). The reliabilities of the other three subscales were between .60 and .65. Hence, the total mean value was calculated from the remaining twelve items, as the reliability was better (see Tab. 1).

Variable	*M* (*SD*)	Range	Cronbach's Alpha
1. Security precautions	3.49 (.63)	1.58-5.00	.77
2. Commitment to technology	3.71 (.62)	1.83-5.00	.85
Technology acceptance	3.39 (.97)	1.00-5.00	.89
Technology competence convictions	4.08 (.84)	1.50-5.00	.88
Technology controll convictions	3.67 (.67)	1.75-5.00	.73
3. Use of social media	3.40 (.82)	1.00-5.00	.77
4. Emailing/surfing	3.83 (.66)	2.00-5.00	.71
5. Age	38.60 (14.01)	19-69	-

Note. $N = 661$. M = mean, SD = standard deviation.

Tab. 1: Descriptive Statistic of the Main Variables

Commitment to technology: We used the *Short Scale for Measuring Technology Commitment (Technikbereitschaft, TB)* by Neyer and colleagues (2012). It consists of three subscales, each with four items, that measure positive attitudes towards technology (*technology acceptance*) as well as *technology competence convictions*, and *technology control convictions*. Participants rated the extent to which the statements ("I am very curious about new technical developments") applied to them on a five-point Likert scale (with the response options from 1 = "not true at all" to 5 = "completely true"). The items in the *technology competence convictions* scale were negative statements and had to be inverted. The Cronbach's Alphas for the three subscales were between .73 and .89. Mean scores were calculated for all three facets. A higher mean value reflects a higher degree of technology commitment.

Use of digital media: The *use of digital media* was measured using five items: Private chat (e.g. WhatsApp, Signal, Telegram), social networks (e.g., Facebook, Instagram), video calls (e.g. Zoom, Skype, Webex, Jitsi), emails, and internet surfing. Participants were asked on a five-point Likert scale how often they used each option and how much they enjoyed using it (see Nikstat et al., 2018). Factor analyses indicated two dimensions: *use of social media* (private chat, social networks, video calls) and *emailing/surfing*. Frequency and preference were highly correlated (r's = .78 and .60), so we calculated mean values for *social media use* and for *emailing/surfing*. Descriptive statistics and reliability of the scales can be found in Tab. 1.

Control variables: Age was included as a control variable because of its significant associations with central variables of the study.

2.3 Data Analysis

We performed correlative analyses using SPSS to examine the bivariate relationships between the central variables. Multiple regression analyses were used to control for age and to examine the unique prediction of each of the facets of technology commitment. Variables were z-standardized for interaction analyses (Aiken & West, 1991).

3 Results

3.1 Bivariate Correlations between the Facets of Technology Commitment, Use of Digital Media, and Security Precautions

In line with previous findings (Neyer et al., 2012), intercorrelations between the three facets ranged between $r = .26$ and $r = .38$ ($p < .01$). As expected, all three facets of *technology commitment* showed a positive correlation with *security precautions,* but among the three facets, *technology acceptance* showed the highest correlation with *security precaution*s *(*$r = .26$; $p < .01$). Furthermore, *technology acceptance* was positively associated with *use of digital media*, namely with *use of social media* ($r = .28$; $p < .01$) and *emailing/surfing* ($r = .22$; $p < .01$). As expected, age was positively correlated with *security precautions (*$r = .36$, $p < .01$) and negatively with the u*se of social media* ($r = -.27$; $p < .01$). The association with *emailing/surfing* was, however, positive ($r = .32$; $p < .01$). The results can be found in Tab. 2.

Note. $N = 661$.
** $p < .01$; * $p < .05$.

Variable	1.	2.	3.	4.	5.	6.
1. Security precautions						
2. Use of social media	-.14**					
3. Emailing/surfing	.23**	.18**				
4. Technology acceptance	.26**	.28**	.22**			
5. Technology competence convictions	.24**	.02	.10**	.38**		
6. Technology control convictions	.23**	.10*	.06	.37**	.26**	
7. Age	.36**	-.27**	.32**	-.05	-.13**	-.00

Tab. 2: Descriptive Statistics and Correlations

3.2 Control Analysis

We then tested with regression analyses whether the hypothesized bivariate correlations remained significant when the three facets of *technology commitment* and age served as predictors (see Tab 3, step 1). All three facets of *technology commitment* predicted *security precautions* significantly, although they became smaller for *technology acceptance* and *technology control convictions*. *Technology competence convictions* was the strongest predictor of *security precautions* (β = .20; $p < .001$).

The expected associations between *technology acceptance* and *digital media use* (β = .31; $p < .001$) and *emailing/surfing* (β = .22; $p < .001$) remained significant. *Technology competence convictions* and *control convictions* had no unique effects on *digital media use*, except for a significant (but small and negative) association between *technology competence* and *social media use*.

The positive correlation between age and *security precautions* remained significant, as did a negative correlation with *social media use* and a positive correlation with *emailing/surfing*.

Criterion	Social media use			Emailing/Surfing			Security precautions		
Predictor variables	β	*t*	*p*	β	*t*	*p*	β	*t*	*p*
Step 1	$R^2 = .16$; $\Delta R^2 = .16$, $p < .001$			$R^2 = .16$; $\Delta R^2 = .16$, $p < .001$			$R^2 = .26$; $\Delta R^2 = .26$, $p < .001$		
Age	-.27	-7.58	< .001	.34	9.46	< .001	.40	11.79	< .001
TA	.31	7.63	< .001	.22	5.52	< .001	.16	4.27	< .001
TComp	-.13	-3.42	< .001	.07	1.87	.06	.20	5.45	< .001
TCont	.01	0.37	.71	-.05	-1.17	.24	.12	3.17	< .01
	F (4, 656) = 31.12 (p<.001), $R^2 = .16$			F (4, 656) = 31.84 (p<.001), $R^2 = .16$			F (4, 656) = 58.29 (p<.001), $R^2 = .26$		
Step 2	$R^2 = .20$; $\Delta R^2 = .04$, $p < .001$			$R^2 = .17$; $\Delta R^2 = .01$, $p = .08$			$R^2 = .27$; $\Delta R^2 = .01$, $p = .16$		
Age	-.27	-7.62	< .001	.34	9.40	< .001	.39	11.51	< .001
TA	.29	7.30	< .001	.22	5.35	< .001	.17	4.37	< .001
TComp	-.12	-2.98	< .01	.07	1.71	.09	.19	5.05	< .001
TCont	-.02	-0.45	.65	-.05	-1.38	.17	.13	3.44	< .001
Age x TA	.10	2.46	< .05	.04	0.95	.34	-.06	-1.61	.11
Age x TComp	-.08	-1.95	>.05	.04	0.89	.37	.07	1.83	.07
Age x TCont	.18	4.74	< .001	.05	1.17	.24	-.03	-0.71	.48
	F (7, 653) = 23.82 (p<.001), $R^2 = .20$			F (7, 653) = 19.26 (p<.001), $R^2 = .17$			F (7, 653) = 34.15 (p<.001), $R^2 = .27$		

Note. TA = Technology acceptance, TComp = Technology competence convictions, TCont = Technology control convictions.

Tab. 3: Regression Analyses

3.3 Moderation Analyses

In a final step, we tested the interaction effects. We expected that *technology acceptance* and *technology control convictions* would moderate both the age-related differences in *use of social media* and *emailing/surfing*. We conducted moderated regression analyses and computed interaction models (by using *z*-transformed scores) and controlled for the first-order terms of age, *technology acceptance*, and *technology control convictions* (Tab. 1, step 2). As expected, the moderation effects of *technology acceptance* ($\beta_{\text{Age} \times \text{technology acceptance}} \rightarrow$ Social media use = .10; $p < .05$) and *technology control convictions* ($\beta_{\text{Age} \times \text{Technology control convictions}} \rightarrow$ Social media use = .18; $p < .001$) were significant after controlling for the first-order terms. As for *emailing/surfing*, the interaction effects between *technology acceptance* and age as well as between *technology control convictions* and age were not significant.

The results of the significant interactions are depicted in Fig. 1. Higher values of *technology acceptance* (Fig. 1A) and *technology control convictions* (Fig. 1B) dampened the negative correlation between age and *use of social media*.

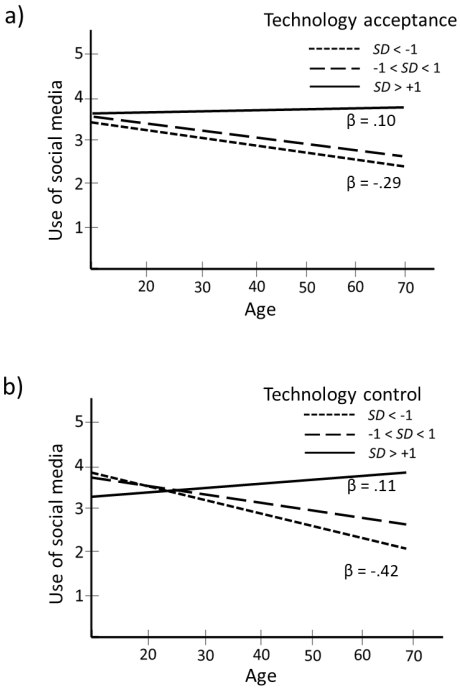


Fig. 1: Use of Social Media as a Function of Age and a) Technology Acceptance, b) Technology Control Convictions

4 Discussion

One aim of the present study was to examine whether the three facets of *technology commitment* are associated with *security precautions* and *digital media use*. All three facets of *technology commitment*, as well as age, were positively correlated to *security precautions*.

As expected, *technology acceptance* was positively correlated with *use of digital media*, consisting of both *social media usage* and *emailing/surfing.* It still remained the strongest predictor for *social media usage* and *emailing/surfing*, even after controlling for background variables, age, and interactions. These findings possibly indicate that an individual's personal interest in new technology is highly relevant for the actual *usage of digital media*, whereas the awareness of and intentions to practice secure behavior while using new technology goes along with the personal anxiety involved in their use.

As expected, *technology competence convictions* were associated with *security precautions*, in particular when age and the other facets of *technology commitment* were controlled. Due to the correlative nature of the data, one can speculate about the underlying dynamics, that is, whether a high degree of feeling competent leads to security-related awareness and behavior, or whether taking security precautions results in experiencing more competence. The role of *technology control convictions* is not so clear. This variable showed only a small, albeit expected, association with *security precautions* when control variables were included.

In accordance with our expectations, age was negatively correlated with *use of social media*. The association with *emailing/surfing*, however, was positive. This result possibly reflects that writing emails or browsing the internet could already be well known among older people, while social media is still perceived as relatively new technology which requires more effort (e.g., the installation of apps) and motivation (Tyler et al., 2020). It is also possible that younger adults communicate more with their peers via social media or via messenger services on their mobile devices, which makes the need for emails redundant.

The correlation between age and *security precautions* is in line with results from other studies (Branley-Bell et al., 2022; Zeissig et al., 2017) and emphasizes the protective function of cautious behavior for older adults. It has been shown that the behavioral component of fear, and to a lesser extent the affective or cognitive components, showed an increase with age (Greve et al., 2018).

As expected, the interaction terms for age and both *technology acceptance* and *technology control convictions* were shown to be significant predictors for *social media usage*, which indicates the relevance of the two facets in connection with age. Due to

its relation to interest in new technologies, we expected *technology acceptance* to moderate the relationship between age and *digital media usage*. This was the case only for the *use of social media* and not for *emailing/surfing*, which matches the positive correlation between *emailing/surfing* and age. For *emailing/surfing, technology acceptance* seems not to be important, probably because older adults may already be familiar with these activities; it is no longer new, and is therefore equally familiar to all generations. However, especially when it is used by older people, they may be more easily susceptible (e.g., to phishing mails) than younger people. It is therefore important for older people in particular to be informed about the potential risks. On the other hand, younger people could also be easily targeted, as they are not used to dealing with potential risks in this area or engage more in risk-taking behavior (Rolison et al., 2014).

In contrast, older adults used social media more frequently if they showed higher levels of *technology acceptance*. In short, age no longer mattered as long as participants were interested in new technologies. The same applies to *technology control convictions*, which moderated the association between age and *use of social media*, but not between age and *emailing/surfing*. Again, older adults used social media more often if their *technology control convictions* were more pronounced, which indicates the importance of feelings of self-efficacy in handling new technology.

3.4 Limitations

The cross-sectional data used in this study does not allow conclusions to be drawn about causal directions regarding the correlations. Interaction analyses should still be validated in further research. Another aspect is that our data rely on self-assessed information from the participants. Our data may thus be subject to biases such as socially desirable responses, or may not represent actual behavior; this is in contrast to studies that examined actual behavior like the task of (not) responding to phishing emails (Halevi et al., 2013) or creating a secure online password (Maraj et al., 2019) (e.g., in the case of *security precautions*).

3.5 Conclusion

Earlier findings have shown that *technology commitment* is associated with both *security precautions* and *use of social media* (Haunschild & Leipold, 2023) and also moderates the negative association between age and *use of digital media*. This study expands the findings in showing that *use of social media* is closely associated with *technology acceptance* (i.e., one's personal interest), whereas *security precautions* is associated with *technology competence convictions* (i.e., less anxiety in dealing with new technology). In addition, age was positively correlated with *security precautions* and *social media usage*. Finally, age was no longer important in terms of *social media*

usage if older adults showed higher levels of *technology acceptance* or *technology control convictions*. Taking this into account can provide approaches to promote and encourage (a cautious) use of new technologies across all age groups.

Acknowledgments

This work originated in the LIONS research project. LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – Next GenerationEU.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage Publications.

Beier, M. E. (2021). Life-Span Learning and Development and Its Implications for Workplace Training. *Current Directions in Psychological Science, 31*, 56–61. https://doi.org/10.1177/09637214211003891

Branley-Bell, D., Coventry, L., Dixon, M., Joinson, A., & Briggs, P. (2022). Exploring Age and Gender Differences in ICT Cybersecurity Behaviour. *Human Behavior and Emerging Technologies, 2022,* 2693080. https://doi.org/10.1155/2022/2693080

Chang, P. F., Choi, Y. H., Bazarova, N. N., & Löckenhoff, C. E. (2015). Age Differences in Online Social Networking: Extending Socioemotional Selectivity Theory to Social Network Sites. *Journal of Broadcasting & Electronic Media*, *59*(2), 221–239.

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (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), pp333–352. https://doi.org/10.1037/0882-7974.21.2.333

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS quarterly*, pp. 319-340. https://doi.org/10.2307/249008

Egelman, S. & Peer, E. (2015). Scaling the Security Wall: Developing a security behavior intentions scale (SeBIS). *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, New York*, pp. 2873–2882. https://doi.org/10.1145/2702123.2702249

Ferizaj, D., Perotti, L., Dahms, R., & Heimann-Steinert, A. (2023). Technologienutzung im Alter: Zusammenhänge zwischen Akzeptanz, Kompetenz, Kontrolle, Interesse und sozialen Indikatoren bei Personen über 60 Jahre. *Zeitschrift für Gerontologie und Geriatrie,* pp. 1–7. https://doi.org/10.1007/s00391-023-02225-9

Greve, W., Leipold, B. & Kappes, C. (2018). Fear of Crime in Old Age: A Sample Case of Resilience? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *73*, 1224–1232. https://doi.org/10.1093/geronb/gbw169

Halevi, T., Lewis, J., & Memon, N. (2013). A pilot study of cyber security and privacy related behavior and personality traits. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 737–744). doi: 10.2139/ssrn.2383427

Haunschild, I. M. S. & Leipold, B. (2023). The Use of Digital Media and Security Precautions in Adulthood. *Human Behavior and Emerging Technologies, 2023,* 1436710. https://doi.org/10.1155/2023/1436710

Krampen, G. (1991). *Fragebogen zu Kompetenz- und Kontrollüberzeugungen (FKK). Handanweisung.* Göttingen: Hogrefe.

Kubicek, H. (2023). Online-Zugang und digitale Teilhabe im Alter: Defizite und Evidenz für eine wirkliche Gestaltung von E-Government aus der Perspektive einer großen Bevölkerungsgruppe. *Verwaltung & Management, 29*(2), 60–72. https://doi.org/10.5771/0947-9856-2023-2

Maraj, A., Martin, M. V., Shane, M., & Mannan, M. (2019). On the null relationship between personality types and passwords. In *2019 17th international conference on privacy, security, and trust* (PST) (pp. 1–7). IEEE. https://doi.org/10.1109/PST47121.2019.8949024

Marquié, J. C., Jourdan-Boddaert, L., & Huet, N. (2002). Do older adults underestimate their actual computer knowledge? *Behaviour & Information Technology*, *21*(4), 273–280.

Neyer, F. J., Felber, J., & Gebhardt, C. (2012). Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft. *Diagnostica, 58*(2), 87–99. https://doi.org/10.1026/0012-1924/a000067

Nimrod, G. (2017). Older audiences in the digital media environment. *Information, Communication & Society, 20*(2), 233–249.

Prensky, M. (2001). Do They Really Think Differently? Digital Natives, Digital Immigrants, Part II. *On the Horizon, 9*(6), 1–6.

Rolison, J. J., Hanoch, Y., Wood, S., & Liu, P.-J. (2014). Risk-Taking Differences Across the Adult Life Span: A Question of Age and Domain. *The Journals of Gerontology: Series B*, 69, 870–880. https://doi.org/10.1093/geronb/gbt081

Shah, M. M., Hassan, R., & Embi, R. (2012). Technology acceptance and computer anxiety. *2012 International Conference on Innovation Management and Technology Research.* Malacca, Malaysia, pp. 306–309. IEEE. https://doi.org/10.1109/ICIMTR.2012.6236408

Tyler, M., De George-Walker, L., & Simic, V. (2020). Motivation matters: Older adults and information communication technologies. *Studies in the Education of Adults*, *52*(2), 175–194.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science 46(2)*, 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Zeissig, E. M., Lidynia, C., Vervier, L., Gadeib, A., & Ziefle, M. (2017). Online privacy perceptions of older adults. In *Human spects of IT for the Aged Population. Applications, Services and Contexts: Third International Conference, ITAP 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part II 3* (pp. 181–200). Springer International Publishing.

Improving Information Security Awareness and Compliance Through Serious Game Participation

Manfred Hofmeier[1]

Abstract: With regard to the digital sovereignty of the individual, empowering individuals to understand risks and use IT securely is a key factor. The LIONS project primarily examines serious games as a means of empowerment. There are several research approaches that investigate the influence of serious games on information security awareness and compliance. The work presented in this paper examines the effect of participation in a serious game on a specific topic area (here: malicious insider threats) on the general information security awareness and information security policy compliance. Therefore, multiple performances of "Operation Digital Butterfly," a serious game about malicious insider threats, were accompanied by surveys to measure changes in information security awareness and information security policy compliance. This paper describes and discusses the results of these surveys.

Keywords: Awareness, Information Security Policy Compliance, Serious Games

1 Introduction

The LIONS project examines digital sovereignty on three layers: state or supranational institution, organization, individual (Fries et al., 2022). With regard to the digital sovereignty of the individual, empowering individuals to understand risks and use IT securely is a key factor. The LIONS project primarily examines serious games as a potential learning method to empower the individual. This involves various learning objectives, with information security awareness and information security policy compliance being of particular interest in the project's research (e.g. Hofmeier, 2024). There are several research approaches that investigate the influence of serious games on information security awareness and compliance. These include, for example, Denning et al. (2013), Rieb et al. (2017), and Sailer et al. (2017). This paper now examines the effect of participation in a serious game on a narrowly defined topic area (here: malicious insider threats) on the general information security awareness and information security policy compliance. The main research questions are:

1. Does participation in a serious game about malicious insider threats (specific topic) improve general information security awareness beyond the topic of malicious insider threats?

[1] University of the Bundeswehr Munich, Neubiberg, manfred.hofmeier@unibw.de

2. Does participation in a serious game about malicious insider threats (specific topic) improve general information security policy compliance beyond the topic of malicious insider threats?

To answer these research questions, multiple performances of "Operation Digital Butterfly", a serious game about malicious insider threats, were accompanied by entry and exit surveys to measure changes in information security awareness and compliance.

2 The Game: Operation Digital Butterfly

"Operation Digital Butterfly" is a tabletop game with an exchangeable game board, game cards, and guidelines. In the game, three to four teams with two to four players per team compete against each other by developing insider threat actor roles, attacks, and countermeasures. The game has two main rounds: After a briefing on the game and the rules and formation of the teams, each team develops motivation, attack, and security measures in a creative process, taking the perspective of malicious insiders. The teams are composed such that the expertise is as mixed as possible, while in each game iteration emphasis is placed on a mix of affiliations and professional roles (and thus real-world insights and competences). Each team presents its results to the others. The second round is about the rating: each team rates the roles and attacks of the other teams and presents the rating. This determines the winning team.

The game board describes the environment in which the insider threats take place. To date, the game has been played with three different game boards: Slaughterhouse and cutting plant, Logistics hub with warehouse, and Travel management in a public authority.

Each team develops an insider role, an attack, and a security measure using a card deck. The discussion in the teams and the presentation of the results are structured by the card deck. The teams are instructed to answer four questions on the role card to guide the creative design of attack measures:

1. Who is the insider (position in the organization)?
2. What does the insider want to achieve (intention)?
3. Why does the insider want that (motivation)?
4. How does the insider justify this to himself/herself (neutralization)?

These role characteristics help to create a plausible insider role that has the potential to give hints about factors that might drive or hamper insider threat actors. In this way, they allow a detailed analysis of the game results in regard to potential countermeasures. Justifications of attacks – in the sense of the neutralization theory (Sykes & Matza, 1957) – have been shown to be particularly useful.

The attack is developed using the scene cards. The filmmaking metaphor is used to make descriptions of attacks easy – also for players not used to formal notations. An attack is represented as a sequence of scenes. This way, each team is able to tell their fictional insider attack by describing a sequence of scenes.

To make the game more fun and also to gain knowledge about countermeasures to insider attacks, each team fills out a security measure card. Teams are instructed to anticipate possible attacks by the other teams (the roles are known) and develop an adequate countermeasure. This measure is valid for the attack plans of all teams, and is then taken into account when rating the attacks.

Tab. 1: Serious game iterations

Iteration	Date	Game board	Players	Attack scenarios
2	May 2020	Meat production	6	2
3	July 2020	Meat production	15	4
4	October 2020	Logistics hub	7	3
5	November 2020	Logistics hub	15	4
6	March 2021	Logistics hub	12	3
7	September 2021	Travel management	12	3
8	February 2022	Travel management	10	6
9	February 2022	Travel management	9	6
10	February 2022	Travel management	9	6
11	July 2022	Logistics hub	13	3

The winning team is determined through a rating system, in which the teams rate each other in three predefined categories: (1) plausibility of the role, (2) plausibility of the attack story, and (3) damage potential. Each team can award up to ten points for each category to the other teams. Note that the most important categories for later analyses are the *plausibility* categories. They ensure that the developed attacks and roles are – to some extent – realistic and fit the profile of the role. The "damage potential" category makes the teams more likely to develop attacks that cause significant damage and therefore are of particular interest in security research.

The game ends with a closing discussion, focusing on possible countermeasures in regard to the attack scenarios developed in the game.

In ten game performances from May 2020 to July 2022 with participants from

research institutions, companies, and public authorities, a total of 40 attack scenarios were developed. The game performances were accompanied by game validation methodology, including tests for changes in awareness and information security policy compliance.

3 Research Methods

Some game sessions (3, 5, 7-11) were accompanied by quantitative surveys of the game participants. In each case, an entry survey was conducted before the game and an exit survey after the game, using an online questionnaire. The participants received the link to the questionnaire by email with the participation information, generally a few days (maximum seven days) before the game. The link to the initial survey was sent to the respondents a few days (one to three days) after the game performance as part of a thank-you email. Participation in the surveys was voluntary and it was also possible to participate in the game without taking part in the survey. In the iterations 7-11, questions on general information security awareness and information policy compliance were included in the questionnaires of both the entry and exit surveys in order to be able to identify a possible change. The individual data sets from the initial and exit surveys were not linked on an individual level (panel), but the mean values from the iterations were compared with each other (trend). This type of comparison was chosen in order to guarantee anonymity of the participants to the research team.

Two subscales from the Security Behavior Intentions Scale (SeBIS) (Egelman & Peer, 2015) were used to measure awareness: Device Securement and Proactive Awareness. The Device Securement scale refers to securing devices and the workplace and was selected because these items play a role in attacks from within the organization. Proactive Awareness is representative of general information security awareness and is intended to show whether there is also an effect on general security awareness outside the area of insider threats. In order to be able to measure a change in attitude towards information security policies, two scales related to information security policy compliance were selected: Information Security Polity Intentions (ISP Intentions) according to Siponen et al. (2010) and Information Security Policy Attitude (ISP Attitude) according to Bulgurcu et al. (2010). The ISP Intentions scale primarily examines the intention to comply with policies, to recommend compliance to others, and to support others in compliance. The ISP Attitude Scale, on the other hand, measures attitudes toward the policies.

4 Analysis Results

The results from the entry and exit surveys on information security awareness and information security policy compliance are examined and discussed in the following.

4.1 Awareness

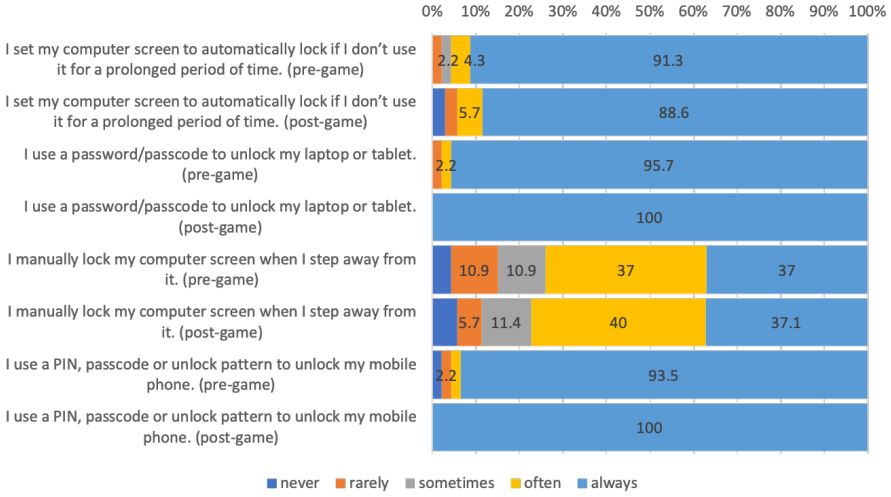


Fig. 1: Device securement (SeBIS) before and after the game

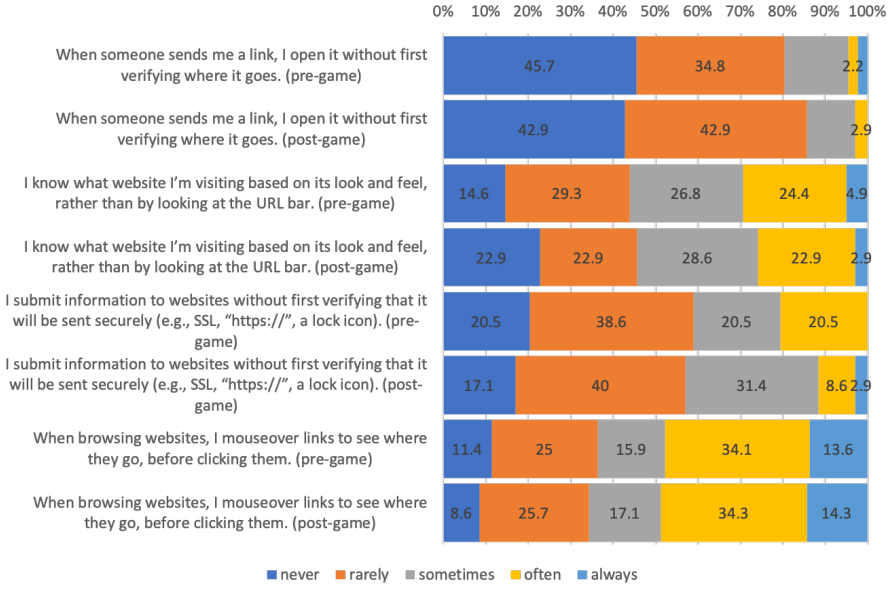


Fig. 2: Proactive awareness (SeBIS) before and after the game

In the case of device securement (Fig. 1), i.e., the protection of IT devices by the user, there are no clear changes after participation in the game. Although there is a positive shift in three of the four items, the shift is negative for the configuration of the automatic screen lock, which is also particularly relevant for the topic of insider threats.

For proactive awareness (Fig. 2), on the other hand, there is a shift in favor of awareness in all items after the game, even though the items are not directly related to the topic of insider threats.

4.2 Information Security Policy Compliance

As can be seen in Figure 3, the intentions in regard of information security policies are higher in all three items after the game than before. The most obvious difference is in those items that affect other individuals: the intention to recommend others to comply with the policies and the intention to help others to comply with the policies.

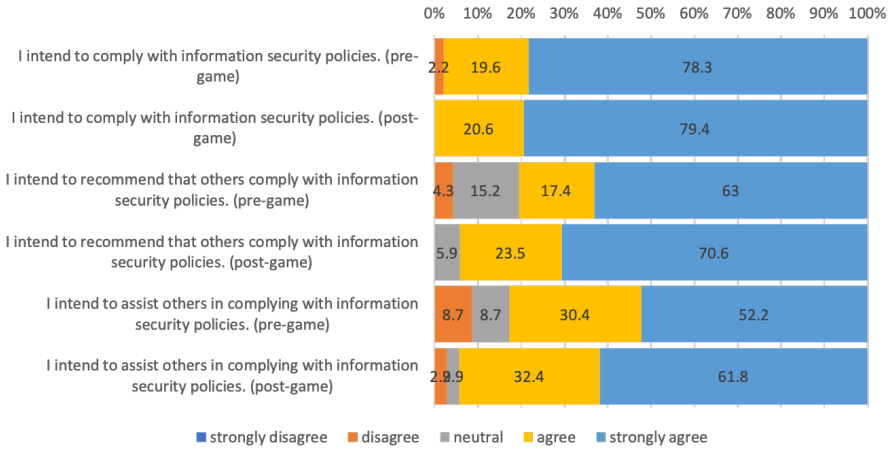


Fig. 3: ISP compliance intentions before and after the game

With respect to attitudes toward information security policies (Fig. 4), the changes are ambiguous. While compliance with the policies is perceived to be more important and useful to the individual after the game, it is also perceived to be less beneficial. When it comes to the perception of necessity, the result is not clear. Overall, game participation still provides an effect on the attitude towards information security policies.

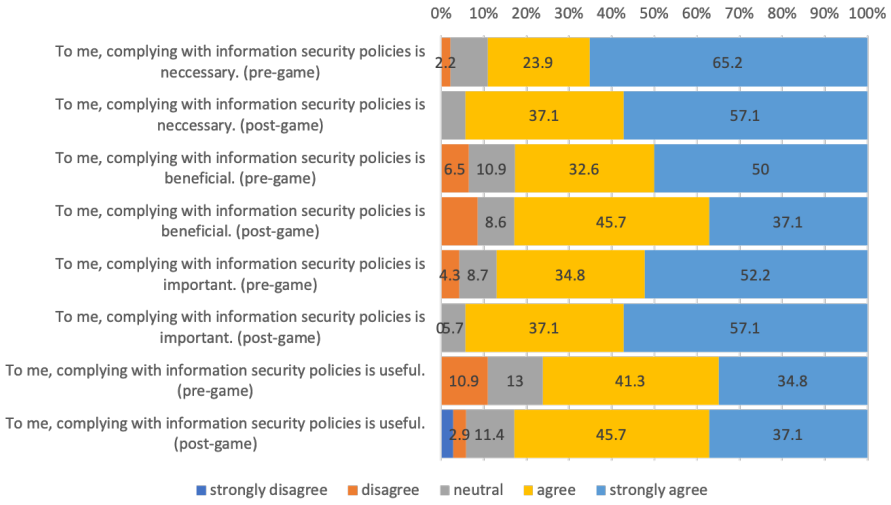


Fig. 4: ISP compliance attitude before and after the game

5 Discussion

Overall, it can be concluded that participation in a serious game on malicious insider threats, a narrowly defined topic, has an impact on general awareness and compliance.

In terms of information security awareness, an overall improvement can be observed as a result of the game, although the relation of the individual items to the topic of insider threats does not appear to be of particular relevance. With regard to information security policy compliance, it is noticeable that the items relating to other individuals showed particular improvement, while the items relating to the individual themselves tended to decline. Similarly, policies are seen as more important and useful, but at the same time less beneficial for the individual. It can be assumed that by participating in the game, people become more aware of the importance of policies and their relationship to individuals, which also raises awareness of the control these policies exert over their own person.

However, there are a few limitations. Testing the long-term effects of game participation has previously shown to be challenging, as the learning experiences of the players are implicit and it is rarely possible to observe the players over a longer period of time after the game. Therefore, the learning was only measured directly after the game participation. Moreover, the present analysis is only a glimpse into the transfer effects of game participation. Future research could examine the transfer effects of game participation in more detail.

Acknowledgments

This work originated in the LIONS research project. LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

References

Bulgurcu, B., Cavusoglu, H. & Benbasat, I. (2010). Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness. *MIS Quarterly 34 (3)*, 523-548.

Denning, T., Lerner, A., Shostack, A. & Kohno, T. (2013). Control-Alt-Hack: The Design and Evaluation of a Card Game for Computer Security Awareness and Education. *Proceedings of the 2013 ACM SIGSAC Conference on Computer Communications Security*.

Egelman, S. & Peer, E. (2015). Scaling the security wall: Developing a security behavior intentions scale (SeBIS). *Conference on Human Factors in Computing Systems - Proceedings*, 2873-2882. doi: 10.1145/2702123.2702249

Fries, I., Greiner, M., Hofmeier, M., Hrestic, R., Lechner, U., Wendeborn, T. (2022). Towards a Layer Model for Digital Sovereignty: A Holistic Approach. *17th International Conference on Critical Information Infrastructures Security*.

Hofmeier, M. (2021). Operation Digital Butterfly. URL: https://github.com/LIONS-DLT /operation-digital-butterfly (visited on October 27, 2023).

Hofmeier, M. (2024). Operation Digital Butterfly – Ein Serious-Game-basierter Ansatz zur Identifikation und Analyse von Intentionalen Bedrohungen durch Innentäter und Innentäterinnen (Malicious Insider Threats). Dissertation, University of the Bundeswehr Munich, https://athene-forschung.unibw.de/148672.

Rieb, A., Hofmann, M., Laux, A., Rudel, S. & Lechner, U. (2017). Wie IT-Security Matchplays als Awarenessmaßnahme die IT-Sicherheit verbessern können. *Towards Thought Leadership in Digital Transformation: 13. Internationale Tagung Wirtschafts-informatik*, 867–881.

Sailer, M., Hoppenz, C., Beckers, K. & Pape, S. (2017). Förderung von IT-Sicherheitsbewusstheit durch spielbasiertes Lernen - eine experimentelle Studie. *Educational Research and Governance (AEPF 2017)*.

Siponen, M., Pahnila, S. & Mahmood, M. A. (2010). Compliance with Information Security Policies: An Empirical Investigation. *IEEE Computer Society 43 (2)*, 64-71.

Sykes, G. M. & Matza, D. (1957). Techniques of Neutralization: A Theory of Delinquency. *American Sociological Review 22 (6)*, 664–670.

Count2zero, a Serious Escape Room Challenge for Cyber Security Training

Design and Evaluation Concepts

Markus Rebhan[1], Jens Holtmannspötter[2], and Ulrike Lechner[3]

Abstract: The escape game "count2zero" is designed to raise awareness for cybersecurity. It is played in an air raid bunker equipped with IoT components, robots, and digital assistance systems. Teams consisting of a maximum of five players are confronted with various puzzles, which must be solved sequentially to escape the room. When solving the puzzles, the players must take care not to violate IT security policies. Each of the puzzles reflects a situation that may occur in general service operations. Players must comply with security policies and a careless action can quickly lead to an avoidable IT security breach when the clock ticks. We present the current game design and results from the first test game. The results indicate that count2zero is a fun factor, and players report that they find IT security puzzles inspiring and that they perceive that the game raises awareness for cybersecurity.

Keywords: Serious Game, Escape Game, IT Security, IoT, Military Scenario

1 Introduction

Increasing digitalization presents the Bundeswehr with new challenges in IT security. With the growing importance of information technology in military operations, soldiers must be aware of IT security when dealing with cyber-physical systems. Social engineering, spear phishing, and insider attacks rely on the fact that individuals make mistakes and do not, or perhaps cannot adhere to security rules: "Our communication systems were not compromised" and "The fact that the conversation could still be intercepted is due to an individual application error" (Spiegel Politik, 2024). This statement, used to comment on a wiretapping affair, illustrates the human factor in cybersecurity and that a military scenario needs dedicated security measures. Awareness campaigns and training courses typically address security awareness in information technology in the "white" area of operational and individual information systems such as office applications, email, and smartphones. The focus of our design is security compliance in case of operational technology, robots, and other cyber-physi-

[1] Bundeswehr Research Institute for Materials, Fuels and Lubricants, Erding, markusrebhan@bundeswehr.org
[2] Bundeswehr Research Institute for Materials, Fuels and Lubricants, Erding, jensholtmannspötter@bundeswehr.org
[3] University of the Bundeswehr Munich, Neubiberg, ulrike.lechner@unibw.de

cal systems in a military context. Training security in the use of operational technologies might be risky and requires special considerations and resources: an operational technology might pose a safety risk to its handler or society when IT security rules are violated. The transfer from training to daily work routines poses a challenge. Serious games are one way to create a memorable experience to facilitate transfer from the tentative actions in game to the real life.

The aim of this research is to develop a serious game to raise awareness of how to deal with cyber-physical systems such as those found in the military environment in a challenging context. The game's target group are members of the Bundeswehr of all age groups, both military and non-military. The evaluations of the game address the perceived usefulness of the game.

"Count2zero" is designed as an escape game. A series of tasks embedded in a story must be solved to escape the room. The design of the game rooms, the plot, and the tasks follows a military theme. The location in an air-raid bunker makes it a memorable experience. Players are confronted with a series of challenging situations. Puzzles must be solved in compliance with generally applicable IT security rules. Security breaches, which are somewhat provoked by the design of the puzzles, need to be avoided as far as possible. Players receive direct feedback from the interactive components within the game environment and IT that is part of the game includes robots and IoT devices. The game is embedded in a process of briefing and debriefing and a data collection for evaluation of the game design. A comprehensive debriefing includes information about the background during the game and points out errors and improvements to the chosen course of action. The game is played by teams of a maximum of five people, which should be heterogeneous; technical understanding is just as important as teamwork and creative thinking for a successful game. The aim of the game is to increase personal IT security awareness by sensitizing participants to IT security risks in an entertaining, memorable, and interactive way.

This article presents the state of research of the serious escape game count2zero after the initial design of the game, and the puzzles and one game for validation purposes.

2 State of the Art

The research interest is to create with count2zero a serious game that raises awareness of cybersecurity for operational technology in a challenging context. This section gives a brief overview of serious game design and evaluation methods, serious games for cybersecurity, and the state of the art in compliance with security policies. A brief evaluation concludes this section.

2.1 Serious Games: Definition and Methods

An early definition of serious games in the literature, which corresponds to today's

usage, was introduced by Clark C. Abt (Abt, 1970). He describes serious games as simulations and games to improve education. As this definition was introduced at a time when the development of video games was still in its infancy, it referred exclusively to pen and paper based games. The computer game he developed, called T.E.M.P.E.R., was also used in a military context (Djaouti et al., 2011).

The article "Success factors for serious games to enhance learning: a systematic review" provides another apt definition of serious games. There, serious games are roughly defined as structured activities that aim to support learning processes by promoting motivation and positive emotions and enabling deeper learning processes through their inherent characteristics (Ravyse et al., 2017).

Blötz defines a comprehensive method compendium and guidelines for designing serious games meant for business purposes (Blötz, 2015).

Criteria from the serious game literature to evaluate serious games are:

- Entertainment factor: The game must be designed to be appealing and fun in order to keep the players' attention. This can be achieved through interesting challenges, reward systems and an exciting storyline (Ferguson, 2007).
- Educational value: The learning objectives clearly defined in advance should be conveyed through the interactive playing of the game. This can involve learning or deepening skills or knowledge (Michael & Chen, 2006).
- Feedback: Immediate feedback on actions helps players reflect on their decisions and improve their skills (Gee, 2005).
- Motivation: Various elements such as rewards, progress indicators, and challenges can be used to generate motivation for the game among players and maintain it throughout the gameplay (Ryan & Deci, 2000).
- Ease of use: Gameplay should be designed to be simple and easy to understand so that players can focus on the task at hand and not waste unnecessary time on complex processes that are difficult to understand. When designing the game, attention should be paid to accessibility to enable all players to play (Connolly et al., 2012).
- Immersiveness: Immersing a player in the gameplay is a crucial factor in maintaining a player's attention over an extended period of time. A compelling story, realistic sounds, and high quality visuals help to increase immersion (Dickey, 2005).
- Ethics: Ethical principles should not be disregarded in a serious game, but respected. This ensures that players are taught positive values (Perron & Wolf, 2008).

Wargames can be seen as a specialized method of serious games in which conflict situations are simulated with the aim of training in tactics and strategy. In the literature, the first applications of wargames can be dated to the nineteenth century. Baron von Reisswitz used this method to train his officers in decision-making (Perla, 1990).

The possible areas of application of wargames include attack scenarios as well as defense and especially prevention scenarios. Business wargames use the game logic of wargames in business contexts (Mark Hope, 2022).

2.2 Escape Games

Escape games from the adventure game genre are a specialized form of serious games. The game idea of an escape game or escape room in terms of entertainment is to solve puzzles and find keys to escape a room in the given time. One of the first successful escape game was developed in Japan in 2004 as an online game called "Crimson Room" (Penttilä, 2018). In Germany, the first live escape games were created in 2013 and called HintQuest (TEAM MANIA GmbH, 2024). Due to the game-based learning approach, escape games are not only part of the entertainment industry, but are also of interest in education, as they offer a mix of excitement, challenge and social experience. (Borrego et al., 2017; Giang et al., 2020).

Among the escape games with an educational and cybersecurity background, very few have been evaluated. One of the few games that have been evaluated is the "Computer-Security-Oriented Escape Room." This was designed to train employees' awareness of computer security. Two scenarios were generated in which the players act as attacker and defender. For evaluation purposes, a questionnaire was distributed to the players after the game in this study. All players confirmed that it was a rewarding and effective experience (Beguin et al., 2019).

A mobile escape room called the Cyber Security Awareness Truck was part of the exhibition at the 2023 ICT Security Conference in Linz. This is the first time that the Austrian Federal Computing Center had used gamification as part of its training and awareness measures (Bundesrechenzentrum, 2023). This indicates the potential of an escape room for training cyber security content.

2.3 Serious Games for Cybersecurity

Serious games are one method for raising awareness of cybersecurity. Shostack maintains a list of serious cybersecurity games online. Most of these games focus on "white" office IT topics and the respective security risks, such as social engineering. Notable serious games of his list include Operation Digital Chameleon and Control-Alt-Hack (Shostack, 2024).

A serious game on IT security in critical infrastructures was designed by A. Rieb with the name "Operation Digital Chameleon" (Rieb & Lechner, 2015). In Operation Digital Chameleon, red and blue teams develop attack and defence strategies for critical infrastructures. This serious game was developed for education and training purposes in critical infrastructures and has been evaluated. The serious game is inspired by the paradigm of open innovation according to Reichwald and Piller and the quality that

comes from participatory designs (Reichwald & Piller, 2009). The game was designed with the hypothesis that employees are often aware of an organization's vulnerabilities and, at a minimum, they should have dealt with threats and vulnerability analysis. Engaging them in a game would raise their awareness of cybersecurity and address vulnerabilities specific to their critical infrastructure. The game provides insight into the level of knowledge on the topic of IT security and novel attack vectors have been identified in the game (Rieb et al., 2017).

2.4 Compliance with IT Security Policies

The purpose of the escape game, which we plan to apply, is to raise awareness of cybersecurity policies, as players have to overcome various challenges in a realistic environment while navigating typical security vulnerabilities and attack vectors and the need to comply with the typical security policies. The literature on cybersecurity policy compliance considers various factors that influence the intention to comply with security policies. The Unified Model of IT security policy compliance (Moody et al., 2018) marks a milestone in the scholarly discussion in the field of information security as it integrates with deterrence theory (Schelling, 1960), neutralization theory (Sykes & Matza, 1957) and important theories and factors. However, the underlying empirical studies seem to have been done for information systems and not for operative technology. Again, there is a lack of understanding and empirical studies for the military domain and its specific technologies and contexts.

3 Research Design

The design of "count2zero" is guided by the design science paradigm according to Hevner (Hevner et al., 2004). Development and evaluation follow an iterative scheme in which refinements, games, and evaluations take place. The game was designed considering current technologies such as IoT and robot technologies as well as the current state of research in the field of serious games and the requirements of the military context. Observation during the game and a pre- and post-survey are used to evaluate the game.

This paper describes the game itself and the status of its development after a test run conducted in 2023 with participants from the University of the Bundeswehr Munich. This trial was to determine the "playability" of the tasks by applying criteria such as time requirement and level of difficulty, as well as the relevance of the individual tasks and the relevance of the storyline.

The actual escape room game is embedded in a process to collect information about the game participants (survey), the processes of obtaining the consent of the participants required by the ethics committee, and a post-game survey with a discussion.

The game design and data collection and processing methods are to be reviewed or

under review by the Ethics Commission of the University of the Bundeswehr Munich and the Data Protection Officer at BAAINBw.

4 The Serious Escape Game "count2zero"

The game "count2zero" was developed as a real-time escape game and takes place in a Bundeswehr air raid bunker. The aim of the game is to recapture the bunker, which is to be regarded as an infrastructure that has been compromised by the enemy, by acting safely without jeopardizing IT security.

4.1 Target Group

Since the development of the game idea, the primary focus of the game has been on members of the Bundeswehr – mainly soldiers, but also civil servants. The game is designed in such a way that the IT affinity of each player does not play a decisive role. However, there should be one person in each game group who has an above-average IT affinity in order to solve one or the other puzzle that requires some in-depth IT knowledge.

4.2 Game Logic

The game begins with a briefing in which the players are informed about the situation and the history of the bunker. The game master informs about the aim of the game and the rules, and in particular on the interaction with the game master during the game. The game master describes the procedures in case of an emergency: how to open the door and escape from the bunker. The players decide on the roles and organization of the team.

Count2zero starts when the front door is closed. From this point on, a countdown runs for 120 minutes. Two game variants are offered: (1) the game may end when the time has elapsed with the number of puzzles solved and security breaches scored up to that point. (2) the players play until the last puzzle has been solved; the number of puzzles solved and security breaches committed up to the end of the time is scored here. Note that the decision as to which game variant is selected can be made before or during the game. A shift towards playing until all puzzles are solved can be seen as a clear indication of the fun factor. In the current configuration, 10 puzzles have been defined, which must be solved sequentially by the players to escape the game. All puzzles, except for the first one require information or items from previous puzzles.

All puzzles must be solved in compliance with the generally applicable IT security policies. Every breach of IT security policies is documented and taken into account for the score.

4.3 Playing Field

The playing field of count2zero is located in an air raid bunker (Fig. 1) on the grounds of the Wehrwissenschaftliches Institut für Werk und Betriebsstoffe (WIWeB) in Erding. The bunker consists of four rooms that can be opened in the course of the game by sequentially solving the puzzles.

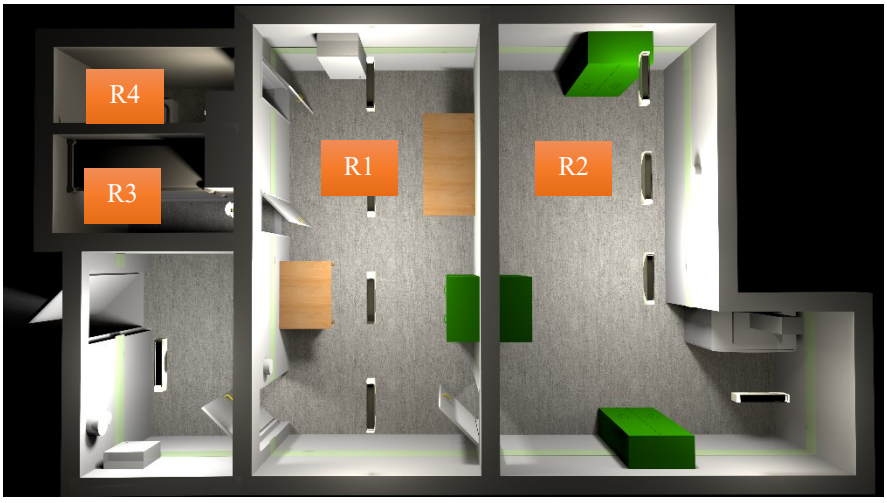


Fig. 1: Playing field “bunker”

The bunker makes the game unique and allows physical stressors to be emulated: dark light, low temperature, and noise. “Locked-in” is a tangible feeling as there is no mobile connection, no daylight, and no noises from outside. Lack of mobile phone connectivity is an advantage for organizing a serious game. The lack of a mobile phone connectivity opens up the possibility of influencing the data connection via a dedicated Wi-Fi connection. The difficulty of the game can be adjusted by controlling, e.g., light intensity in the room. Overall, the atmosphere corresponds with the military character of the game.

An infrastructure was designed and built to make the game interactive and controllable. To this end, a separate power supply was installed in the bunker to supply power to various workstations, lights, video cameras, electronic door locks, Wi-Fi, and audio devices. A dedicated internet connection and a complete network infrastructure with active and passive LAN components was design and installed. Fig. 2 depicts a network plan. Components marked in green are for game management and data collection, and the components marked in red belong to the game process. The room is equipped with IoT components that allow the game to be controlled and monitored remotely.

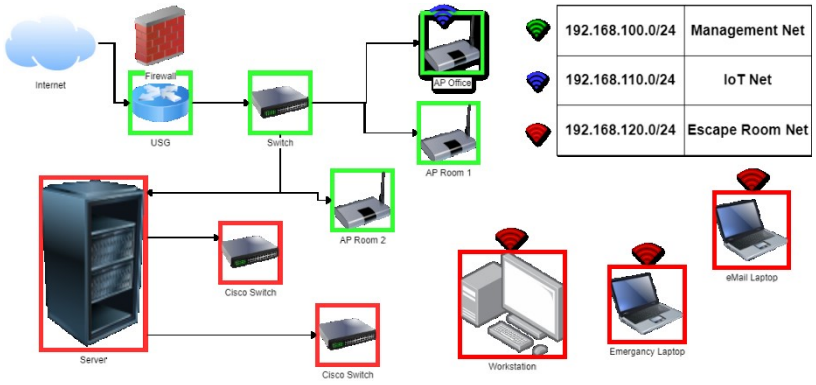


Fig. 2: Network diagram

The game master is not in the room during the game and monitors the players through video surveillance. He can give hints and intervenes when players need help or to clarify rules. In addition to the elements required for the individual puzzles, there is a video bell in the room, which the players can press and contact the game master to receive a clue. A TV screen, as depicted in Fig. 3, visualizes the current game status.

Fig. 3: Screen of the game state

After the second puzzle, the game team receives a cell phone with a Wi-Fi Internet connection and can thus make contact with the Internet, which eventually will pose a risk to them.

4.4 Game Rules and Briefing

The players are informed about the situation they face and their task before the door is locked as part of a briefing. This is as follows:

> "You are the most readily available response team we could find. You have been called together to investigate the following acute IT security incident. Contact with a high-security bunker, which is used as a secret communication node and listening post, has been completely interrupted for four hours. The situation is therefore completely unclear.
>
> Initial assessments suggest that hostile forces have infiltrated and carried out manipulations that are as yet unknown. All attempts to make contact have so far been unsuccessful.
>
> Satellite images taken shortly before communication was lost show two people approaching the bunker, but it is not clear whether they are site support personnel or normal shift workers. An analysis of the time recording system in the building shows that three people are constantly taking turns.
>
> Your task is to investigate the incident on site and restore the interrupted communication as quickly as possible. The bunker's firewall is not responding to requests from outside, so we assume that the active IT infrastructure is without power and or contact to the property access node has been interrupted. All information about the bunker is subject to the highest level of secrecy, we have not yet been able to contact the group of people with the necessary security clearance to obtain further information. However, further information can be expected in the course of the operation.
>
> As time is working against us, you must start the mission immediately. With the help of our AI, we have managed to take control of the intercom in the bunker. Unfortunately, this channel is not bidirectional, so we were only able to establish a voice output without a return channel. We will therefore provide you with new information as the mission progresses. Our AI is still working on remote access to the bunker, so perhaps we will soon be able to transmit image information.
>
> Once you have passed the entrance to the bunker, you are on your own. The walls of the bunker are so well shielded and isolated that no communication is possible apart from the intercom.
>
> We were able to extract an old building plan from the archive showing the bunker and its premises. (Hand over)
>
> Due to the short time available and the unclear situation on site, I cannot give you any further assistance.
>
> If possible, comply with the applicable IT security regulations and restore communication. We do not know what and to what extent the bunker system has been tampered with, so you must connect the following 'Application Layer Gateway' before restoring communication. The task is completed when we regain access to the systems. You will notice this when the power supply in the bunker is restored and the lights come on in all rooms. You can then leave the area by opening the door. If you open the door beforehand, which you are entitled to do at any time, the mission will be aborted prematurely and the mission will have failed."

After the briefing, the players are informed about the escape routes and the symbols in the room. Elements that are not part of the game, such as the network cabinet for the game controls and light switches, are marked with stick-on pictograms. These elements may not be manipulated, and manipulation of these elements may risk the players' safety and would end the game.

4.5 The Game

Count2zero is a team-based escape game, a team consists of a maximum of five players who start in the middle room labelled R1 (Fig. 4). The first puzzle is indicated by a red light above the game zone. The order of the other puzzles is determined by the information and clues acquired during the game. A red light indicates the players to move to the next game zone.

Each puzzle has a specific IT security theme. A puzzle can make the players think about either an action or a technique. For example, the first game has the background of showing how a secure password can be generated with a password card and how quickly it can be possible to crack a password that is too simple with suitable tools. However, it can also provoke "unsafe" behaviour, i.e., violation of IT security guidelines. This is intended to encourage reflection and expand awareness of IT security guidelines and cyber-physical systems. For example, the careless connection of a peripheral device to a computer causes an error that makes the computer temporarily unusable for the players. The astonishment at this unexpected incident causes the players to reflect.

Fig. 4 shows a schematic representation of the bunker and the individual rooms with a flow chart of the individual stations (S1–S10), the numbering corresponds to the chronological solution sequence of the puzzles.

After the players have entered the bunker, the door is closed, and the game master starts the game with a signal tone. A Wi-Fi-controlled countdown timer begins to run down. At the start of the game, the current game zone is indicated by IoT lighting that extends across the entire room. This is intended to help players find their way around the new environment more quickly.

Fig. 4: Game flow chart

At this point in the game, only the room labeled R1 is available to the players. Fig. 5 shows the first station S1, with a standard computer workstation, a monitor, mouse,

and printer. The workstation computer (APC) is locked; the players' first task is to log on to the PC, and they are already faced with the first problem: how do I get valid login data for this workstation? With a little creativity and investigative ideas, this problem should be solved quickly. In detail, the players must recognize that the connected printer still holds a print job that has not yet been printed due to a lack of paper. With this printout and a note under the keyboard, the players can guess part of the login data. For the second part of the credentials they need to explore the desk.

Fig. 5: Workspace game of puzzle 1

After solving a few tasks, the players are able to open and discover another room called R2. This room can be opened with an access card, after the door is opened, a further authentication factor is requested. Specifically, this puzzle emulates a two-factor authentication. If the team cannot find the second factor quickly enough, the team is separated for a brief period of time. The team can decide for itself who stays behind in which room; at least one person must remain in the room. Bright lights, sounds, and time pressure make this situation stressful. Once the players have mastered this situation, they receive support from the humanoid robot "Pepper" (Fig. 6). It interacts with the players and is willing to disclose information about the escape room in exchange for information from them.

Fig. 6: Robot “Pepper”

The game ends when the time runs out, or the players have solved all the puzzles. The final puzzle involves connecting a special device provided at the start of the game to a server located in a locked server cabinet, accessing a website, and entering a reset password. This loads a non-compromised software version, the area can be safely put back into operation, and the escape room is successfully solved.

4.6 Debriefing

A debriefing takes place after the game. At the beginning, the players are asked to fill out a questionnaire about the fun factor, difficulty, comprehensibility, and their cybersecurity takeaways. The players then have the opportunity to ask specific questions and discuss their actions. The game master specifically addresses IT security threats that were simulated in the game. This part of the evaluation is about whether the players were able to identify all security threats as such. The game master asks for suggestions for improvement. As data from the interaction of players on the internet is also collected during the game, a brief demonstration of risks and how a potential attacker could proceed is given.

Furthermore, techniques are discussed with which possible attacks can be recognized and which measures can be taken to counteract them. The debriefing helps to a great extent to better understand and deepen what has been experienced and ultimately to expand one’s own IT security awareness, which results in a significant risk minimization when dealing with IT devices.

5 Evaluation

This section deals with the evaluation of the questionnaires completed by the players before and after the game. At the end of 2023, a first test run of the game was carried out with five players from the University of the Bundeswehr Munich. The aim of this run was to determine the "playability" of the tasks with criteria such as time required and level of difficulty, as well as the relevance of the individual tasks and the relevance of the storyline. At the time of this game run, six of the ten game stations had been completed.

The players were asked to complete a questionnaire before the game. The aim of this questionnaire was to determine the players' level of knowledge and expectations. The composition of the team in terms of IT knowledge was heterogeneous. On a 10-point scale, where one point corresponds to "no IT knowledge" and ten points to "excellent IT knowledge," two people rated their own IT knowledge at three points, one person at eight points and two people at nine points. This corresponds to an average of 6.4 points. The respondents were also asked about the relevance of protecting personal data and their own willingness to comply with IT security rules. In the second part of the questionnaire, questions were asked about experiences with current IT security issues. It was also noticeable that only one question was answered in the affirmative by all participants. All game participants have already had experience with phishing emails, which clearly highlights the presence of this topic. The following table displays some results of the pre-match questionnaires.

Tab. 1 Questions in pre-game questionnaire

Question	Pers.1	Pers.2	Pers.3	Pers.4	Pers.5	Average
Level of own IT knowledge	8	3	9	3	9	6.4 Pt.
Protection of personal data	9	7	10	7	10	8.6 Pt.
Compliance with IT rules	5	9	8	8	9	7.8 Pt.

After the game, the players were again asked to complete a questionnaire about their experience. This questionnaire is structured into two parts; part one consists of 13 questions with a 10-point scale, and part two consists of seven questions with free text.

To assess the immersiveness and entertainment factor of the game, the players were asked whether they enjoyed the escape room. The players gave an average score of 9.4 points, which can be interpreted as positive feedback.

As expected, those players who initially rated their own IT skills as very high tended to answer question two, about the increase in knowledge due to the escape room, with fewer points than players who rated their own IT skills as very low. The average score here was 5.4 points, meaning that players perceived a learning experience in the game.

The players' answers to the open questions enabled significant improvements to the game flow and the individual puzzles. This is reflected in smoother gameplay and a better understanding of the task to be completed by the players.

For example, all players had problems understanding the task in a puzzle with Morse codes; this could be improved with a few targeted hints. Players found the attention to detail, the setting, and the realistic game environment particularly appealing. All puzzles were perceived as exciting and educational.

6 Summary and Future Work

This article presents the serious game count2zero and the experiences gained from the first test run. The first results indicate that the game offers a lot of fun and allows players to improve their personal security awareness. The feedback collected from the questionnaires is positive. Feedback from observation, the survey, and a discussion after the first run contributed to developing more puzzles and refining the existing puzzles.

The game count2zero is meant to motivate players to think about their actions in the game right from the start so that they do not commit security breaches. The complex design of the puzzles makes it challenging to comply with security policies when using the IT devices provided. It is also interesting to see when and whether players are willing to compromise on IT security and what factors contribute to non-compliance with IT security policies. The players perceived the presence and interactions with the humanoid robot "Pepper" as very positive.

Further test runs are planned to improve the gameplay and more puzzles will be developed. The long-term goal is to have a game platform to validate security policies for modern IoT technologies used in the military context and to raise security awareness for these technologies.

Acknowledgments

This work originated in the LIONS research project. LIONS is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr which we gratefully acknowledge. dtec.bw is funded by the European Union – NextGenerationEU.

References

Abt, C. C. (1970). Serious Games. New York: Viking, 1970, 176 pp., L.C. 79-83234. *American Behavioral Scientist*, *14*(1), 129. https://doi.org/10.1177/000276427001400113

Beguin, E., Besnard, S., Cros, A., Joannes, B., Leclerc-Istria, O., Noel, A., Roels, N., Taleb, F., Thongphan, J., Alata, E., & Nicomette, V. (2019). Computer-Security-Oriented Escape Room. *IEEE Security & Privacy*, *17*(4), 78–83. https://doi.org/10.1109/MSEC.2019.2912700

Blötz, U. (2015). *Planspiele und Serious Games in der beruflichen Bildung: Auswahl, Konzepte, Lernarrangements, Erfahrungen – Aktueller Katalog für Planspiele und Serious Games 2015.*

Borrego, C., Fernández, C., Blanes, I., & Robles, S. (2017). *Room escape at class: Escape games activities to facilitate the motivation and learning in computer science* (No. 2). https://doi.org/10.3926/jotse.247

Bundesrechenzentrum. (10.2023). *AUSTRIA CYBER SECURITY CHALLENGE 2023*.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Dickey, M. D. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development*, *53*(2), 67–83. https://doi.org/10.1007/BF02504866

Djaouti, D., Alvarez, J., Jessel, J.-P., & Rampnoux, O. (2011). Origins of Serious Games. In M. Ma, A. Oikonomou, & L. C. Jain (Eds.), *SpringerLink Bücher. Serious Games and Edutainment Applications* (pp. 25–43). Springer London. https://doi.org/10.1007/978-1-4471-2161-9_3

Ferguson, C. J. (2007). The good, the bad and the ugly: A meta-analytic review of positive and negative effects of violent video games. *The Psychiatric Quarterly*, *78*(4), 309–316. https://doi.org/10.1007/s11126-007-9056-9

Gee, J. P. (2005). *Learning by Design: Games as Learning Machines: Interactive Educational Multimedia* (pp. 15–23).

Giang, C., Chevalier, M., Negrini, L., Peleg, R., Bonnet, E., Piatti, A., & Mondada, F. (2020). Exploring Escape Games as a Teaching Tool in Educational Robotics. In M. Moro, D. Alimisis, & L. Iocchi (Eds.), *Advances in Intelligent Systems and Computing. Educational Robotics in the Context of the Maker Movement* (Vol. 946, pp. 95–106). Springer International Publishing. https://doi.org/10.1007/978-3-030-18141-3_8

Hevner, A. R., Park Jinsoo, March, S. T., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1).

Mark Hope. (2022, January 29). *What is a Business Wargame?* https://asymmetric.pro/what-is-a-business-wargame

Michael, D., & Chen, S. (2006). *Serious games: Games that educate, train, and inform*. Course Technology.

Moody, G. D., Siponen, M., & Pahnila, S. (2018). Toward a Unified Model of Information Security Policy Compliance. *MIS Quarterly*, *42*(1), 285–311. https://doi.org/10.25300/MISQ/2018/13853

Penttilä, K. (2018). *History of Escape Games* [Master Thesis]. University of Turku.

Perla, P. P. (1990). *The art of wargaming: A guide for professionals and hobbyists*. Naval Inst. Press.

Perron, B., & Wolf, M. J. P. (2008). *The video game theory reader 2* (2nd ed.). Routledge. https://permalink.obvsg.at

Ravyse, W. S., Seugnet Blignaut, A., Leendertz, V., & Woolner, A. (2017). Success factors for serious games to enhance learning: a systematic review. *Virtual Reality*, *21*(1), 31–58. https://doi.org/10.1007/s10055-016-0298-4

Reichwald, R., & Piller, F. (2009). *Interaktive Wertschöpfung: Open Innovation, Individualisierung und neue Formen der Arbeitsteilung* (2., vollständig überarbeitete und erweiterte Auflage). *SpringerLink Bücher*. Gabler. https://doi.org/10.1007/978-3-8349-9440-0

Rieb, A., Hofmann, M., Laux, A., Rudel, S., & Lechner, U. (2017). Wie IT-Security Matchplays als Awarenessmaßnahme die IT-Sicherheit verbessern können. In J. M. Leimeister & W. Brenner (Chairs), *Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)*.

Rieb, A., & Lechner, U. (2015). *Operation Digital Chameleon.* https://doi.org/10.1145/2957792.2957800

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68.

Schelling, T. C. (1960). *The strategy of conflict: With a new preface* [Nachdr. d. Ausg. 1980]. Harvard University Press.

Shostack, A. (05.2024). *Tabletop Security Games + Cards*. https://shostack.org/games.html

Spiegel Politik (2024). Verteidigungsminister Pistorius spricht von »individuellem Anwendungsfehler«: 05.03.2024. https://www.spiegel.de/politik/deutschland/taurus-abhoeraffaere-verteidigungsminister-pistorius-nennt-individuellen-anwendungsfehler-als-ursache-a-e77aed7d-578a-4a88-951a-2d666332486e (zur Abhöraffäre bei der Bundeswehr).

Sykes, G. M., & Matza, D. (1957). Techniques of Neutralization: A Theory of Delinquency. *American Sociological Review*, *22*(6), 664. https://doi.org/10.2307/2089195

TEAM MANIA GmbH. (2024). *Die Geschichte der Live Escape Games: 03.06.2024*. https://www.exitmania.com/blog/die-geschichte-der-live-escape-games-vom-computerspiel-zum-trend

Digital sovereignty is an essential pillar of the future. Politicians discuss it, businesses strive for it, and public discourse increasingly centers on it. However, what does it take to become digitally sovereign and resilient? What needs to change in society's digitalization strategy and the design and operation of digital infrastructure to achieve this?

In this book, the authors explore these questions, drawing from the LIONS research project led by a consortium of the University of Bundeswehr Munich and the Helmut-Schmidt-University in Hamburg. Through a transdisciplinary approach involving industry partners and the Bundeswehr, perspectives from ethics, psychology, media pedagogy, information systems, and computer science come together in the resulting articles.

The first part, *The Challenge of Digital Sovereignty*, examines the concept of digital sovereignty as a systemic task. The second part, *Designing Sovereign Information Systems*, builds on the ethical perspective and locates it concretely in practical application within the supply chain. The final section, *Digital Sovereignty as a Learning Field*, offers pedagogical and psychological insights into digital competence.

Funded by

dtec.bw
Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr

Funded by the European Union
NextGenerationEU

Logos Verlag Berlin

ISBN 978-3-8325-5834-5