# Designing Digital Involvement in a Datafied Society

Zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der KIT-Fakultät für Wirtschaftswissenschaften des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Carolin Stein

---

| | |
|---|---|
| Referent: | Prof. Dr. Christof Weinhardt |
| Korreferent: | Prof. Dr. Timm Teubner |
| Tag der mündlichen Prüfung: | 01.04.2025 |

Karlsruhe, 2025

## Abstract

In modern societies, datafication has fundamentally reshaped how we comprehend and navigate social, economic, and political systems. Data mediates power and agency giving rise to new innovations, open business and administrative models, and promoting social connectivity and participation. Simultaneously, in reality, datafication has increased societal inequalities and fueled challenges, such as exploitation, privacy invasions, or the spread of disinformation. Digital platforms have become key venues for the positive and negative effects of datafication, mediating the involvement of citizens in political, economic, or even scientific processes. Their design, therefore, goes beyond the mere consideration of technical functionality; it is decisive for maintaining and improving democratic practices. This dissertation examines the design of digital involvement focusing on the complexities of datafied societies. Thereby, it targets two key contributions. First, addressing the increasing convergence of public and private, as well as economic and political spheres, it establishes theoretical and practical foundations to unite traditional research streams on participatory practices under a shared perspective of "digital involvement". Through a two-cyclic Design Science Research project, artifacts are developed that define digital involvement, capture its diversity, and support practitioners in navigating the field. Subsequently, they are applied to the use case of citizen science. Second, the dissertation provides a scientific and practice-oriented foundation for addressing the implications of datafication in formats and platforms of digital involvement. Two experimental studies give insights into the effects of varying data representations and explore the design of tools, simplifying complex participation scenarios. Additionally, three mixed-method studies investigate opportunities and system design for skill-enabling and learning in digital involvement. Overall, the dissertation interweaves multiple areas of information systems research, including crowdsourcing, digital government, e-learning, conversational agents and immersive systems, integrating interdisciplinary knowledge from the fields of citizen science and data literacy. It advances the theoretical and practical discourse through the systematic documentation of literature and platform landscapes, the acquisition of empirical findings, and the development and evaluation of new design artifacts. As a key result of Parts I and II, the dissertation demonstrates interest and applicability of the introduced digital involvement framework in practice and its usefulness in cross-domain and domain-specific contexts. Furthermore, Parts III and IV highlight the need for a nuanced approach to designing and evaluating participation formats and tools in datafied societies and present possible solutions for imparting complex data and data-related skills.

# Acknowledgements

Writing this dissertation on the topic of digital involvement, I am more than thankful for the (non-digital) involvement of so many wonderful people along the way.

First of all, I would like to express my gratitude to Prof. Dr. Christof Weinhardt, without whom this dissertation would not have been possible. Thank you for your guidance in my academic journey, beginning with my first bachelor thesis at IISM, continuing through my time as a junior researcher while based in Berlin, and extending into my PhD studies. Thank you for your always constructive and enriching feedback, your inspiration and creativity, and your advice to embrace exciting (research) questions as they arise, rather than planning everything down to the smallest detail. I would also like to thank my second supervisor, Prof. Dr. Timm Teubner, for his openness to exploring the world of citizen science and his valuable support, from my Master thesis to this dissertation. I am also grateful to Prof. Dr. Kay Mitusch and Prof. Dr. Jella Pfeiffer as part of my examination committee for their valuable feedback and critical discussion of my work.

The completion of this work in all its aspects would not have been possible without the support and cooperation of numerous colleagues and scholars.
First and foremost, I would like to express my gratitude to Dr. Jonas Fegert. Thank you for sharing your enthusiasm for research on digital democracy with me and for supporting me at the FZI from the very beginning. It has been and continues to be a pleasure to work with you at the House of Participation and to head the IMA department. Thank you to my colleagues at the FZI, especially in the Berlin office, for filling my everyday work with joy, inspiring conversations, and enthusiasm for our shared topics and research. I especially want to thank Alicia Wittmer, Gregor Pahlitzsch and Lennard Nicholaisen for their valuable support and collaboration throughout research projects. Working with you has been and continues to be motivating and enriching, and I deeply appreciate your contributions to our joint work. To my colleagues at the IISM - now WIN - thank you for sharing your constructive feedback and ideas in DokSems, and for the wonderful moments we shared on conference trips and retreats, which gave me the comforting feeling of not being alone on this journey. In addition, I am happy and proud to have been able to publish my work with so many other great co-authors - Jessica Jachimowicz, Prof. Dr. Tim Straub, Dr. Anna Soßdorf, Osman Kıvanç Kırıkçı, Lukas Buß, Isabel Bezzaoui, Prof.

Dr. Stefan Morana, and Moritz Müller my thanks also go to you.

Finally, I would like to thank the people who have always reminded me that work is not the most important thing in life. This dissertation is, therefore, dedicated to my family and friends.
Charlotte, Sienna, Faruk, Petra, Julian and Celine, thank you for your friendship over decades and across the country. Thank you for enduring occasional whining and complaining about work and papers, drying my tears in times of failure, and for celebrating every little success with me as if it were a big one. But above all, thank you for all the wonderful times when we didn't think about work at all.
Thank you to my family, for all your kind and loving support, which is simply impossible to describe here in full extent. So I would just like to say by way of example: Mama, Papa, the vocabulary quizzes were worth it in the end; Thank you for never getting tired of encouraging and supporting me with everything I needed. Julia, looking back, your ceiling is still no Picasso, but at least it's the work of a doctor; Thank you for being such an understanding, caring and fun big sister that I could always look up to. Yannik, I solemnly swear that from now on I will remember by myself to eat, drink and step outside for some fresh air again; Thank you for helping me to find solutions for all of life's small and big challenges. You bring peace to my mind, love to my heart, and so much positive energy to my soul, I am grateful for all our time together.

Berlin, April 2025
Carolin Stein

# Contents

# Part I.

# Introduction and Foundations

**Abstract Part I**

*Part I "Introduction and Foundations" of this dissertation introduces the topic of designing digital involvement in a datafied society. It explores the transformative impact of datafication on society, reshaping social, economic, and political aspects of modern life. While datafication offers opportunities for collaboration and participation, it also introduces challenges, such as risks to privacy, exploitation and pseudo-participation, which threaten democratic values. In response, this part emphasizes the need to research the design of digital involvement—a unifying concept for participatory approaches like crowdsourcing, citizen science, and e-participation. The introduction presents key research objectives in addressing the dual challenges of democratic participation and data literacy, and outlines the approaches undertaken in this dissertation. Finally, the part establishes the theoretical foundations for the emerging concept of digital involvement by developing a taxonomy for digital involvement projects. This taxonomy research was conducted in collaboration with Tim Straub, Jessica Jachimowicz and Jonas Fegert, and the contents were presented at the 31st European Conference on Information Systems (Stein et al., 2023b). In this part, titles, tables, and figures from the original publication have been renamed, reformatted, and re-referenced to align with the structure of the dissertation. Additionally, chapter and section numbering were adjusted, and formatting, abbreviations, and references were standardized for consistency.*

# 1. Introduction

*"When we have all data online it will be great for humanity. It is a prerequisite to solving many problems that humankind faces."*, remarked Robert Cailliau (Peninsula, 2009), one of the main contributors to the invention of the World Wide Web in the late 2000s. Yet, the reality of datafication and how we now comprehend and navigate social, economic, and political systems have unfolded with complexities that often challenge, and even threaten democratic values. Data-driven business models and the spread of online disinformation only exemplify two phenomena that threaten modern democracy and make the design of a democratic digital engagement of society a critical, yet challenging pursuit. Participatory approaches such as crowdsourcing, citizen science, or e-participation hold great potential to address societal challenges and need to be jointly researched to effectively realize democratic claims. This dissertation thus develops the overarching concept of *Digital Involvement* for their joint consideration. However, given that information is the very fundament of (digital) engagement, we must also ask ourselves in times of growing datafication: How can we design digital involvement in a way that it inclusively communicates complex data and empowers more people to access, understand, and use it? This dissertation focuses on addressing these complexities by exploring the design of digital involvement in a datafied society.

## 1.1. The Transformative Impact of Datafication on Society

The digital transformation and the pervasive phenomenon of datafication have fundamentally reshaped the foundations of modern society across social, economic, and political domains. The quantification of society through the increasing representation of its characteristics in measurements, numerical values, and data is not merely a technological transformation; data has evolved into a critical medium for interpreting, responding to, and making sense of different contexts. Today, data permeates every facet of life, influencing how individuals, organizations, and institutions perceive, interpret, and interact with the world. (Alaimo and Kallinikos, 2024; Hintz et al., 2018; Splichal, 2022)

Economically, datafication has disrupted traditional institutional boundaries and created new business models (Alaimo and Kallinikos, 2024). Where previously institutions operated as self-contained entities, the integration of external data sources and global networks has increasingly intertwined internal operations with external environments (Alaimo and Kallinikos, 2024). With declines in manufacturing profitability, data-driven business models have opened up new avenues for growth and profits. They ultimately culminate in the advent of platform business models, optimizing the extraction and control of data (Srnicek, 2016).

Politically, the rise of data-driven government and the blending of public and private spheres introduced new forms of power and control, raising concerns about privacy, surveillance, and the erosion of democratic processes (Alaimo and Kallinikos, 2024; Hintz et al., 2018; Splichal, 2022; Weinhardt et al., 2024). With access to vast amounts of granular data on citizens, from their movements to their social interactions, governments today obtain behavioral insights that are unprecedented in history, reconfiguring state-citizen relationships (Bigo et al., 2019). This is accompanied by an ideological component, the belief in neutrality and objectivity of data, and the self-evident relationship between aggregated data and the individual. However, this view inherently overlooks the influences of the systems and purposes that shape its collection (Hintz et al., 2018; Van Dijck, 2014). It is because of this that Bigo et al. (2019) argues for "a shift in focus from the politics of or in data to data as a force that is generative of politics" (p.4).

Socially, the shift towards the digital sphere has redefined human interactions. Relationships, communication, and even identities are now mediated by data (Splichal, 2022). The digital sphere offers numerous ways of keeping updated, connected and participate in society. Yet consisting of various forms of "holding for true", it introduces new obligations of finding, distinguishing and using relevant data, gatekeeping one's own flood of information (Splichal, 2022). Overall, these developments highlight the transformative force of datafication that is reshaping the very fabric of contemporary life.

## 1.2. Digital Platforms and Participation in a Datafied Society

The digital connectivity and collaboration, in conjunction with the increased complexities and new business opportunities in a datafied society, have additionally fueled opening developments that have been ongoing for years. Peer- or user-generated production forms and crowdsourcing emerged as participatory economic processes; Internet governance and digital activism shape political interactions and citizen journalism or science change traditional hierarchies of knowledge creation and distribution (Hintz et al., 2018; Weinhardt et al., 2020). Digital platforms have become a key mediator of these opening processes, blurring the boundaries between private and public life and between social and economic spheres (Alaimo and Kallinikos, 2024; Splichal, 2022). While bringing great opportunity to rethink traditional power structures and strengthen democratic and bottom-up practices, these developments do not unambiguously imply a transition towards a more decentralized, social restructuring (Alaimo and Kallinikos, 2024; Hintz et al., 2018). The digital public sphere is dominated by political and economic considerations, shaping the dynamics of participation. Exploitation and pseudo-participation can hide under the idea of opening processes (Altenried, 2020; Haklay, 2013b; Kittur et al., 2013; Palacin et al., 2020). Data-driven business models fuel the erosion of privacy, algorithms deepen inequalities in society, and harmful actors misuse the public sphere to spread disinformation (Splichal, 2022; Srnicek, 2016). The design of digital involvement as such – be it in economic, scientific, or political contexts – extends beyond mere participation; it is a critical component of maintaining and enhancing democratic processes.

In a society where data increasingly shapes social, economic, and political decisions and mediates agency and power, it is crucial to consider the implications of datafication when designing and implementing participatory processes, which are intrinsically about redistributing power and democratizing decision-making. With information being the very fundament of participation (International Association of Public Participation, 2007), the inclusive communication of relevant data becomes the building block of designing inclusive participation. To counteract deliberative deficits, education and skill-enabling remain important measures, and as citizens become more proficient in accessing, understanding, and using data, they are better equipped to participate meaningfully in democratic processes, challenge power structures, and advocate for their rights (Alaimo and Kallinikos, 2024; Splichal, 2022).

## 1.3. Information Systems Perspectives on Designing Digital Involvement

From an information systems (IS) perspective, an anchor to contributing towards an inclusive and democratic design of digital involvement evidently appears in the research on market and platform design (Weinhardt et al., 2024). Research fields such as peer-production, crowdsourcing, e-participation, or even citizen science have been established in IS for years and continue to develop (Weinhardt et al., 2020). They have contributed important insights on the design of computer-mediated work and collaboration, user-centric platform design, incentive schemes, and the design of participation within their respective domain. Multiple classifications and design frameworks of individual participatory practices have been developed (Estellés-Arolas et al., 2015; Haklay, 2013a; International Association of Public Participation, 2007; Shirk et al., 2012; Van Dijk, 2012). However, accounting for the increasing fusion of the public and private, economic and political spheres, it is beyond time to rethink these research streams toward a joint consideration of digital involvement.

This dissertation argues that facilitating mutual exchange and learning between different initiatives becomes imperative to advance and harness the diversity of approaches and adequately guide practitioners in the democratic design of digital involvement. Furthermore, it argues that it is essential to better integrate the implications of datafication into the very design processes shaping the formats and platforms of digital involvement. Adopting Feenberg's critical theory of technology, the design of platforms as humanly created internalizes values and politics (Feenberg, 1999; Haklay, 2013b). They can thus be equally used to reinforce traditional power structures and inequalities when embodying the interests of a small elitist group or empower democratic development when accounting for macro-societal needs. In the context of datafication and citizen participation, it is, thus, evident that platforms for digital involvement need to inherently account for the needs of citizens in a datafied society. Returning to the fundamental role of informed participation, accounting for datafication includes understanding the individual relevance of data for the participation process and aiming at providing support structures so that inclusive communication, understanding, and usage of this data is possible. Thinking of data as a foreign language, there are multiple pathways for system designers to go about this. We can either build tools that simply provide the necessary texts in the users' language, provide translation tools that skill-enable the user to interact with a foreign language or build tools that ultimately empower the user to learn the foreign language. For digital involvement, this translates to participation tools and formats that render relevant data as simple as possible, enable participants in the handling of complex data, or ultimately see digital involvement projects (DIP) as an opportunity for learning, empowering citizens in their self-determined participation in a datafied society beyond the individual projects' boundaries. Being associated with different levels of difficulty, effort, and self-determination, all pathways can have a 'raison d'être', depending on the involvement context.

Therefore, this dissertation takes upon the identified challenges exploring the design of digital involvement formats and tools for a datafied society aiming at uniting and advancing the different IS participation research fields of crowdsourcing, citizen science and e-participation. In three consecutive research parts it explores how the diverse field of digital involvement can be navigated by practitioners (Part II), the role of data in DIP and the support systems that simplify participation (Part III) and the design of support systems for skill-enabling and learning (Part VI). For this endeavor, this dissertation makes use of a mixed-method research approach including a variety of different qualitative and quantitative IS and social science research methods, like design science research (DSR) (Peffers et al., 2007), experimental research (Charness et al., 2012), structured literature and artifact reviews (Gnewuch and Mädche, 2022; Webster and Watson, 2002), and qualitative interview and focus group studies (Kaiser, 2014; Krueger and Casey, 2014).

## 1.4. Structure of the Dissertation

The overall structure of this dissertation consists of five parts, grouping ten chapters and is visualized in Figure 1.1.

**Part I: Foundation**: Besides this introduction (Chapter 1), the first part of the dissertation aims at defining the term digital involvement, thereby laying the foundation for the joint research of established participatory practices such as crowdsourcing, citizen science and e-participation (Chapter 2). The contents of this second chapter emerge from a published conference paper called "Same Same but Different - Towards a Taxonomy for Digital Involvement Projects" ultimately answering the research question:

Research Question 1: *What are the key characteristics to describe and distinguish DIP?*

**Part II: Navigating the Design of Digital Involvement**: Upon the definition of digital involvement, the second part of this dissertation focuses on the provision and application of tools helping practitioners to navigate the diverse design of digital involvement. Chapter 3 entails the content of a published conference paper called "From (Design) Theory to (Participation) Practice: Leveraging a Taxonomy for Digital Involvement Projects". The chapter creates two design artifacts based on the DIP taxonomy, answering the research question:

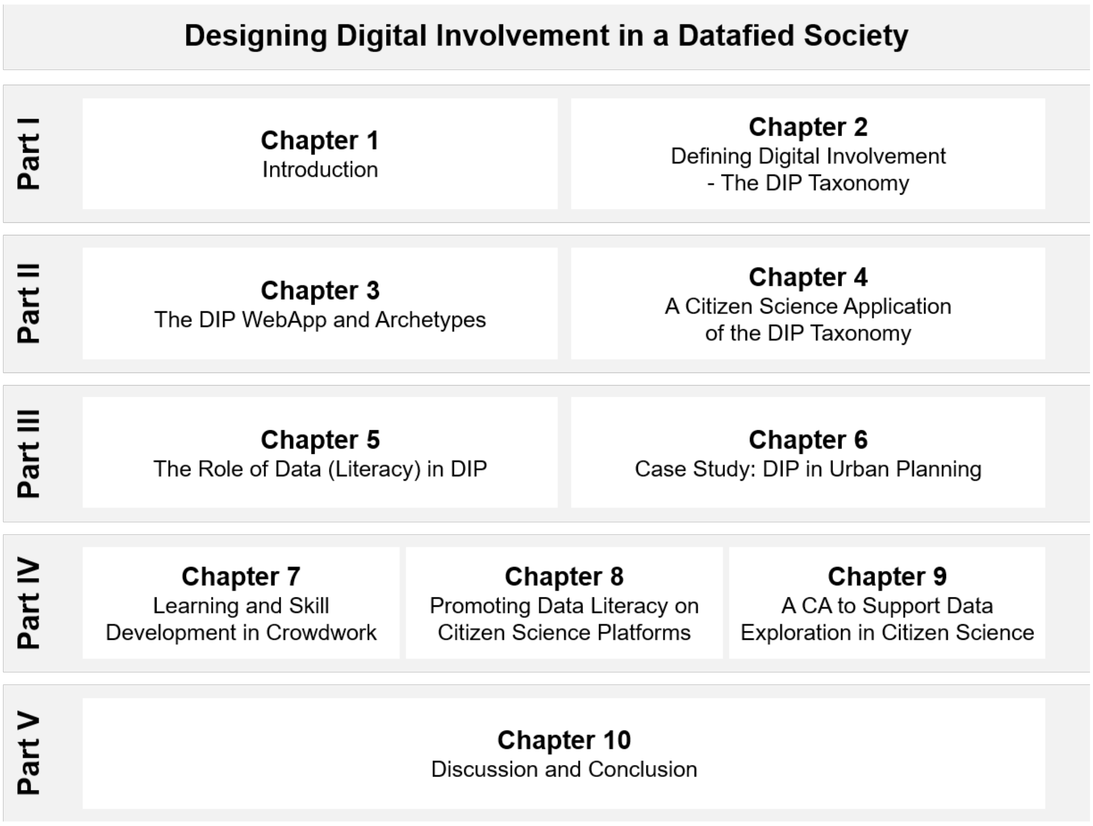


Figure 1.1.: Structure of the dissertation.

Research Question 2: *How can we use the DIP taxonomy to capture design knowledge and make it available to domain practitioners?*

Consecutively, the developed tools are applied at the example of citizen science in Germany. The contents of this chapter have been presented at a citizen science conference and outlined in an extended abstract called "Designing (for) Change - A Taxonomy-based Approach to Project Design." The full publication as a journal article is currently under review aiming to answer the following research questions:

Research Questions 3:

*3.1 What design characteristics do citizen science projects exhibit in an overarching taxonomy for participatory concepts and how are their preferences distributed?*

*3.2 Given the diversity of citizen science practices, to what extent can distinct design clusters be identified?*

*3.3 Regarding the multidisciplinary nature of citizen science, is there an association between a project's disciplinary focus and its assignment to specific design clusters?*

**Part III: Positioning and Simplifying Data in Digital Involvement**: In this part of the dissertation, the role of data and data literacy as fundament of digital involvement will be discussed. Chapter 5 focuses on the theoretical foundations of the data literacy concept and discusses its positioning as societal competence, including unpublished survey results on citizens' perspectives towards the relevance of data and data literacy. Chapter 6 then proceeds with two experimental studies, exploring the effects of different data representations and tools that simplify participation in the context of urban planning scenarios. The content for this chapter emerges from two published conference papers called "Bridging Realities: Exploring Enablement Factors for XR Participatory Urban Planning" and "The Devil is in the Details? Investigating 3D Visualization Types for E-Participation in Urban Planning" and discusses the research questions:

> Research Question 4: *How does the usage of XReality (XR) affect participants' literacy in the context of participatory urban planning and what specific factors are crucial for successful XR participation formats?*

> Research Question 5: *How do properties of 3D visualizations affect their suitability for processes of e-participation in the context of urban planning?*

**Part IV. Designing Enablement and Learning in Digital Involvement**: Upon discussing the role of data and literacy in the context of digital involvement, the fourth part of the dissertation focuses on the exploration of tools to enable and educate citizens within the process of their digital involvement. Chapter 7 structurally captures the current state of the literature on teaching and learning in digital involvement and investigates learning mechanisms on an operating participation platform. The contents emerge from the published conference paper "Learning while Earning? A Literature Review and Case Study on Learning Opportunities in Crowdwork" and answer the research question:

> Research Question 6: *How can crowdsourcing platforms support the learning and skill development of crowdworkers?*

Chapter 8 specifically focuses on the realm of citizen science, structurally capturing existing citizen science platforms and analyzing potentials and challenges for promoting citizens' data literacy. The contents emerge from a published journal article called "Digital Participation for Data Literate Citizens – A Qualitative Analysis of the Design of Multi-Project Citizen Science Platforms" and answer the research questions:

> Research Questions 7: *What multi-project citizen science platforms exist and how do they support the conduction of citizen science projects? What are the potential and challenges for promoting citizens' data literacy through digital citizen science?*

Finally, Chapter 9 delves into the design of an assistant tool for citizen science projects in the form of a conversational agent (CA). The contents for this chapter emerge from a published journal article called "Designing a Conversational Agent for Supporting Data Exploration in Citizen Science" and answer the research question:

Research Question 8: *How should a CA be designed to support data exploration in citizen science applications?*

**Part V: Finale**: The fifth part closes this dissertation with a discussion and conclusion chapter. Bringing together the results of the various research parts and discussing the findings in their broader context, the contributions, limitations, and needs for further research uncovered by this dissertation will be positioned in a research framework for designing digital involvement in a datafied society.

# 2. Defining Digital Involvement - The DIP Taxonomy

## 2.1. Introduction

The participation of citizens via digital platforms has become increasingly common in the public sector (Santamaría-Philco et al., 2019; Steinbach et al., 2020). Amidst the Covid-19 pandemic, digital means have become crucial for the state to maintain interaction with its citizens (United Nations, 2022). As e-participation initiatives are predominantly organized through municipalities and cities on a local basis, a variety of participation formats have been established that are difficult to compare and to evaluate (Pirannejad et al., 2019). The practice is lacking appropriate formats for experience and knowledge sharing and thereby enabling comparisons. This is especially concerning, as not all participation efforts result in successful projects: Depending on design and implementation, approaches can become pseudo-participative processes with negative effects on decision-making and co-design (Davies and Procter, 2020; Palacin et al., 2020). The appropriate design and usage of IS for participation can be challenging and costly, especially as their deployment depends on the local context (Andersen et al., 2007; Hofmann et al., 2020). While digital adoption of new technology in large enterprises, is well advanced across EU countries, the leveraging of innovative technologies is yet a key challenge for the public sector, requiring significant investment (European Comission, 2022). The relevance of digital participation, however, is not limited to the political sphere. On the contrary, these approaches are gradually finding their way into business and science. For economic institutions, it became clear that involving outsiders in processes is a fruitful endeavor (Chesbrough, 2006), resulting in the emergence of crowd approaches, like crowdsourcing, -funding, -work, etc. summarized here under the term crowd-X. In the scientific world, it has been acknowledged that the increasing complexity of problems accompanied by a constant skepticism of citizens towards scientific findings must be countered with new ways of communicating and collaborating (Head and Alford, 2013; Wormer, 2020). As such, citizen science initiatives enabling the involvement of non-professionals in scientific research activities, are increasingly popular throughout research domains (Shirk et al., 2012; Spasiano et al., 2021). In the digitalized world, these participation approaches from e-participation, crowd-X or citizen science consequently take place online simultaneously and encounter similar challenges in design and implementation. Therefore, it seems necessary to have a domain-independent framework that creates a common knowledge base while taking individual differences into account. By understanding the particularities of the disciplines and adapting the digital participation processes accordingly, it becomes possible to learn from other involvement projects and transfer design knowledge (Koch et al., 2011). Thus, a robust understanding and an awareness of individual characteristics would be necessary to support practitioners and policymakers to harness the benefits of the diverse participatory approaches (Andersen et al., 2007; Haklay et al., 2021). While individual domains have produced numerous definitions and classifications of their field (Haklay, 2013a; Haklay et al., 2021; Koch et al., 2011; Royo and Yetano, 2015; Spasiano et al., 2021; Van Dijk, 2012), to achieve this objective, it needs a larger framework to discuss characteristics of involvement projects, outside the conceptual boundaries of participation in the public, economic or scientific sector. As such, we introduce the term DIP as an umbrella term for e-participation, crowd-X, and citizen science, describing projects utilizing digital platforms to involve multiple external individuals in a defined participation process. Our research aims at developing a taxonomy for DIP, exploring similarities and differences, and forming a common ground for intersectoral exchange. Therefore, our research is guided by the following research question: *What are the key characteristics to describe and distinguish DIP?*

Within a tricyclic DSR project, we develop a taxonomy for DIP to profoundly answer the research question at hand. In the first DSR cycle that is presented in this chapter, a preliminary taxonomy is developed following the iterative process based on Nickerson et al. (2013). The taxonomy is thereby conceptually grounded in e-participation, crowd-X, and citizen science theory and combined with an empirical project example analysis. Through using a focus group for the taxonomy evaluation, we find strong support for the necessity of providing guidance within DIP. After refining our taxonomy based on the practitioners' feedback, we thus propose a preliminary taxonomy of 19 dimensions that is suitable to describe DIP according to their participation process, the involved individuals, and the necessary digital infrastructure.

## 2.2. Theoretical Background

The following chapter describes the theoretical background and lays the theoretical foundations for the first DSR cycle by reviewing related literature in the domains of crowd-X, e-participation, and citizen science.

Crowd-X is usually defined as a model targeting to solve problems and complete tasks by involving a(n undefined) digital crowd (Estellés-Arolas et al., 2015; Howe, 2006a,b). In crowd labor markets the crowd processes tasks ranging from simple repetitive email tagging to complex jobs (Kittur et al., 2013). Crowdfunding helps startups to get their capital by using the crowd as investors (Feldmann and Gimpel, 2016; Hammon and Hippner, 2012). Open innovation contests outsource new product developments to crowds (Chesbrough, 2006; Terwiesch and Xu, 2008). Using prediction markets the crowd predicts future events by trading stocks (Kloker et al., 2017; Kranz et al., 2014; Spann and Skiera, 2003; Teschner et al., 2011). While differences exist, all approaches try to harness the human capital available online (Howe, 2006a,b). Although several authors worked on definitions, typologies, and taxonomies of crowdsourcing (e.g. Estellés-Arolas et al., 2015; Prpić et al., 2015a,b), there still does not exist an overall broadly accepted taxonomy. To produce a common ground Estellés-Arolas et al. (2015) proposed eight key elements for an integrated definition of crowdsourcing and identified five main crowdsourcing types.

Prpić et al. (2015b) defined relevant characteristics of crowdsourcing, e.g., cost, anonymity, scale, IT, time, task magnitude, and reliability of the crowd. Furthermore, Prpić et al. (2015a) concluded that 4 crowdsourcing alternatives exist, diverging from Estellés-Arolas et al. (2015) due to a narrower definition of crowdsourcing. Straub et al. (2015), on the other hand, included more concepts in his definition of crowdsourcing, classifying approaches based on the payment schemes in non-payment and payment: piece rate and rank order tournaments. While thereby including all relevant subforms of crowd-X, it only suits a broad classification. As such, crowd-X could be seen as an overarching framework for other phenomena such as e-participation or citizen science, also involving individuals in public policy or in research respectively. However, these domains have created their individual typologies, setting various demarcations to crowd-X.

E-participation itself is a subform of digital government (Macintosh, 2004) and describes the use of different digital means for consultation and democratic decision-making (Sanford and Rose, 2007). Its focus lies on improving communication and cooperation between the stakeholders involved in policy-making (e.g., citizens, politicians, and civil servants). With the introduction of more sophisticated tools, IS research has a growing interest in investigating the underlying forces that impact the quality of democratic processes (Macintosh et al., 2009). The goal of implementing participation is to seek true collaboration between citizens and governments (Pristl and Billert, 2022). Moreover, participation is embedded in a socio-political environment with a certain degree of complexity. Participation can be divided into three different levels of involvement (e-enabling, e-engaging, and e-empowering (Macintosh, 2004)). Other classifications distinguish between forms initiated by citizens and those which address citizens but are run by governmental agencies (Van Dijk, 2012). Some participation models emphasize the ideal of direct democracy with citizens having an actual say in decision-making processes, corresponding with higher levels of participation in some frameworks (Grönlund, 2009; Macintosh, 2004; Pristl and Billert, 2022). Nonetheless, e-participation does not necessarily compete with representative decision-making but rather can be seen as a complement to enhance civic engagement (Aichholzer and Allhutter, 2009). In terms of demarcations to crowd-X, some authors acknowledge crowd-X as a tool that can be incorporated into e-participation projects (Royo and Yetano, 2015). However, e-participation has recently made efforts to self-develop digital infrastructure rather than using established crowdsourcing platforms (Fegert, 2022). As such, concepts like digital citizen participation are arguing for using interdisciplinary research to guarantee inclusivity, interoperability, and democracy in the platform design. Similar efforts in the design of platforms can be observed in the domain of citizen science (Liu et al., 2021).

Citizen science generally describes the involvement of citizens in scientific research activities (Socientize Consortium, 2014). It is a useful means in many research fields (Pettibone et al., 2017; Spasiano et al., 2021) that nowadays frequently utilize Information and Communication Technology (ICT) for organizing participation under the name of virtual, online, or digital citizen science (Reed et al., 2012; Weinhardt et al., 2020). Through the engagement of non-academics, more insights can be generated, and workforce can be contributed (Shirk and Bonney, 2018). As such citizen science describes a broad range of projects, which led to a pluralization of definitions and understandings (Haklay et al., 2021). While some definitions emerged from a political point of view, others approach citizen science from a societal or scientific perspective (Haklay et al., 2021). At the heart of all classifications, however, is the idea to define the extent and implementation of civic involvement in science (Haklay et al., 2021). In the demarcation of approaches in e-participation, such as, e.g., the "ladder of participation," some citizen science authors clearly differentiate that participation levels within citizen science are assessment-free (Haklay, 2013a). Other authors, however, draw a connection between contributory citizen science and consultative e-participation approaches when involving public institutions in citizen science projects (Spasiano et al., 2021). Common classifications for citizen science differentiate between four to six different types of citizen science projects, starting with crowdsourcing-like initiatives or commissioned research to extreme citizen science approaches or collegial work (Haklay, 2013a; Shirk et al., 2012). As such, some authors argue that crowd-X (especially crowdsourcing) can be seen in some cases as a low-involvement subtype of citizen science (Haklay, 2013a). Others argue that crowd-X is a tool used within the digitalization of citizen science (Spasiano et al., 2021). Although demarcations within the domain are not clear, the field is rapidly growing and evolving, and therefore authors call for embracing and exploring plurality rather than creating hypothetical definitions (Haklay et al., 2021).

In conclusion, the review of literature in crowd-X, e-participation, and citizen science reveals that categorization frameworks structuring the domains and their demarcations have yet failed to capture the complexity and diversity of participatory approaches. Similar approaches and challenges in classifying subtypes emphasize the common context of the three fields and support the idea that rigid definitions in the domains are unsuitable for establishing a common classification and, thereby, a basis for conversation about DIP.

### 2.3. Methodology

The general methodological framework for the taxonomy development within a DSR approach is presented hereafter. This introduction is meant to lay the methodological basis for the presentation of the activities and results of the first cycle in the following. DSR describes a problem-solving paradigm with the objective to deliver innovation and prescriptive knowledge through the creation of artifacts (Vom Brocke et al., 2020). As such, embedding taxonomy development within the DSR paradigm provides guidance and transparency to taxonomy creation and evaluation (Kundisch et al., 2021). Additionally, it enables the iterative evolution of taxonomies, allowing to communicate an initial version of a taxonomy that can be used by the research community for further work, evaluating and eventually updating initial versions (Kundisch et al., 2021). This is especially useful, as the field of DIP is constantly evolving in individual areas, thus rigid definitions and demarcations are impractical. We, therefore, embedded our overall research design in a tricyclic DSR approach (see Figure 2.1) following Peffers et al. (2007) six-step pro-

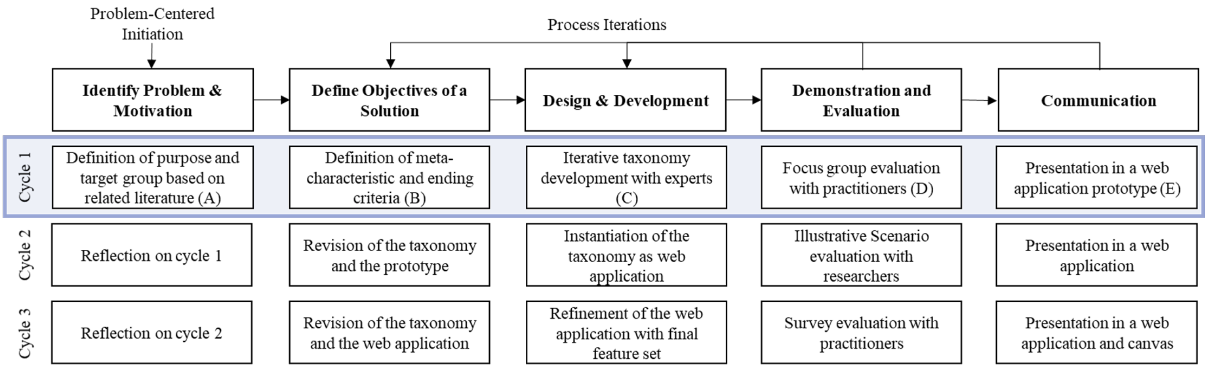


Figure 2.1.: Presented research activities (blue) within the tricyclic DSR approach based on Peffers et al. (2007).

cedure in every cycle. The first cycle served the purpose to develop a preliminary taxonomy. The second and third cycles will further refine the taxonomy, improving its practical usability by instantiating the taxonomy in a web application and a canvas format. By combining conceptual and empirical iterations and qualitative and quantitative methods for taxonomy assessment, we will provide evaluative rigor while elaborating on the artifact's practical relevance (Hevner, 2007; Kundisch et al., 2021; Szopinski et al., 2019).

The six DSR activities of the first DSR cycle targeted the development of a preliminary taxonomy (see Figure 2.1, Cycle 1). For this cycle we methodologically followed Kundisch et al. (2021) who present an approach to embed iterative taxonomy design in DSR: The first step in this method is the definition of the taxonomy's purpose and target group to specify the problem that the taxonomy tries to solve (Kundisch et al., 2021). By reviewing literature in the three domains crowd-X, citizen science, and e-participation, we specified their relation to DIP, partially shown in the related literature section (A). Second, to specify objectives for the taxonomy design, meta-characteristics, to narrow down the scope of the taxonomy's characteristics, and ending criteria should be defined (Kundisch et al., 2021; Nickerson et al., 2013). We thus derived meta-characteristics from the taxonomy's purpose and objective ending criteria to define when our iterative design process should be concluded (B). Third, to design and develop a taxonomy, researchers should follow an iterative process, combining conceptual or empirical iterations to identify dimensions and characteristics of a taxonomy (Kundisch et al., 2021). While conceptual iterations elaborate characteristics based on domain literature, empirical iterations evaluate practical domain artifacts to conceptualize knowledge (Nickerson et al., 2013). We, therefore, started the iterative process with conceptual iterations (C1) utilizing the literature identified in (A) to identify a set of characteristics. After that, we extracted real-life project examples of DIP to evaluate and enrich the set of characteristics based on their applicability in practice in empirical iterations (C2). This was undertaken in an expert group of three researchers, two of them senior researchers, chosen based on their expertise either in e-participation, crowd-X, or citizen science. We concluded the empirical iterations based on the ending criteria defined in (B) with a preliminary taxonomy. For the artifact's demonstration and evaluation, the researcher should evaluate the degree of fulfillment of subjective criteria for a taxonomy (Kundisch et al., 2021). Therefore, we conducted a focus group evaluation with seven industry practitioners (D), which is an established method for taxonomy evaluation (Szopinski et al., 2019). Focus groups generally produce qualitative data by contrasting different groups and benefiting from their interaction (Krueger and Casey, 2014). As such, they can guide design and development processes by improving the experts' understanding of issues related to the artifact, do pilot-testing of prototypes to help with fine-tuning ideas and concepts and to evaluate an artifact (Krueger and Casey, 2014). In this first DSR cycle, a qualitative evaluation method was promising for testing the taxonomy, receiving feedback on the quality of the artifact, and understanding areas for improvement from the perspective of industry practitioners across different DIP domains. Therefore, our focus group was structured as follows: First, practitioners jointly classified new objects by applying the taxonomy (Szopinski et al., 2019). Second, practitioners revealed their individual assessment in a questionnaire including two questions for each subjective ending criterion. Third, the actual focus group discussion was conducted, based on the impressions from the joint classification and the assessments of meeting the subjective ending criteria based on the questionnaire. The on-site evaluation was structured by an interview guideline, documented by an uninvolved professional, who protocoled the session and used an audio recording to enrich the transcript after the session (Krueger and Casey, 2014). As a final activity in the DSR cycle, the researcher should communicate the taxonomy in a way that is appropriate to the target group (Kundisch et al., 2021). We thus communicated the revised taxonomy in a first prototype of a web application (E), which enables practitioners as a target group to classify current approaches of DIP based on their characteristics.

## 2.4. Designing a Taxonomy for DIP

This chapter presents the results of our first DSR cycle, depicting all conducted activities (A)-(E) for the design of a preliminary taxonomy: We start with the problem awareness stage (A) as we foster a problem-centered initiation. Next, we define our solution objectives (B) before iteratively developing a set of dimensions and characteristics as a taxonomy (C1, C2). Finally, we present our evaluation stage (D), where a focus group of practitioners empirically assessed the usability of our taxonomy, and depict our communication efforts in the form of a web application prototype (E).

### 2.4.1. Problem Awareness

The problem awareness stage (A) should create a thorough evaluation of the initial conditions and should question the necessity of creating a new artifact. We elaborate on the need to create a DIP taxonomy by outlining the goal and purpose of the taxonomy and its possible target audiences. The input to this phase is the inspection of related literature in the domains of e-participation, crowd-X, and citizen science, that have been presented as theoretical background. All three phenomena have experienced an increase in application that has resulted not least in a pluralization of concepts, projects, and digital infrastructure

(Estellés-Arolas et al., 2015; Liu et al., 2021; Shirk et al., 2012; Van Dijk, 2012). Thus, efforts were made in all areas to define and subdivide the domains more precisely and to distinguish them from another (Estellés-Arolas et al., 2015; Haklay, 2013a; Macintosh, 2004; Shirk et al., 2012; Van Dijk, 2012). However, these demarcations are neither uniform nor clear. Disagreements in the literature show that none of the conceptualizations is suitable to function as an umbrella term. This prevents the domains from sharing knowledge and tools. We thus identify as a first result, the need for a more general concept.

***DIP:*** *In the absence of an overarching umbrella term, we define the term DIP to describe projects that utilize digital platforms for the involvement of multiple external individuals in a defined participation process. We seek to reconcile with this term the three domains e-participation, crowd-X, and citizen science, although DIP is not limited to these, as other concepts could meet the definition. The proposed definition is used to contrast DIP from use cases of pure computer-mediated work or computer-mediated collaboration (Rogers, 1992; Turner, 1984). On the other hand, we intend to differentiate DIP from Big Tech's social networks that allow registered users to connect with other users and advertisers to use those network effect for their corporate interests (Boyd and Ellison, 2007). However, both of these areas can play a role in conducting DIP (Spasiano et al., 2021; Van Dijk, 2012).*

By defining the concept of DIP and developing a set of differentiating project characteristics for it, we create a framework that allows users to discuss and compare involvement projects across domains. A taxonomy helps to identify similarities and differences between its objects, on the other hand, it provides guidance by structuring past, present, and future developments within a domain (Szopinski et al., 2019). Thus, the target group for our taxonomy is users that are either planning a DIP, looking for guidance without limiting themselves to one domain, or practitioners that want to classify and further develop their participation concept. The taxonomy can also be used by researchers to further refine definitions of subparts of DIP, based on project characteristics from the field.

### 2.4.2. Solution Objectives

To formulate objectives for a solution within our DSR framework (B), we define meta-characteristics, ending conditions for the iterative development process, and evaluation goals for our initial taxonomy (Kundisch et al., 2021). In order to ensure the classification of DIP, we derive our three meta-characteristics from the definition of DIP, namely its cornerstones, the participation process, the projects' individuals, and the digital platform projects. In addition, we utilize the objective ending criteria, suggested by Nickerson et al. (2013) as a condition to conclude our iterative design process. The subjective ending criteria, namely that a taxonomy should be concise, robust, comprehensive, extendible, and explanatory, are key proxies for the taxonomy's usefulness, which can ultimately be judged by observing the taxonomy's application by others (Nickerson et al., 2013). Therefore, we set the taxonomy's usefulness and applicability as evaluation goals and will assess them with individuals from the target group in a focus group.

### 2.4.3. Design and Development: Taxonomy

The design and development of the preliminary taxonomy (C) was structured by an iterative process to elaborate dimensions and characteristics of DIP, utilizing conceptual and empirical iterations. In the following, we will first present the conduction of the conceptual iteration and its initial set of dimensions and characteristics. Then, the results of the empirical iterations will be depicted. Finally, we employed a second-level grouping of dimensions (Kundisch et al., 2021) to organize the results of all iterations.

**Conceptual Iterations (C1):** First, we used the identified literature from (A) to narrow our conceptual base to five theoretical contributions as a starting point for DIP. The contributions were selected as they represent frequently cited frameworks for classifications or groupings of crowd-X (Estellés-Arolas et al., 2015), citizen science (Haklay, 2013a; Shirk et al., 2012), and e-participation (International Association of Public Participation, 2007; Van Dijk, 2012). Second, the three meta-characteristics developed in (B) were utilized to guide the identification of characteristics from the literature. Thus, characteristics either refer to the purpose and design of the participation process, the individuals with their characteristics, roles, and relations, or the digital infrastructure. Through the conceptual iterations, we could identify ten dimensions grouping between two and five characteristics, which will be explained in the following:

*1. Participation Degree:* The scope and intensity of civic participation is a common differentiator for projects within subcategories of DIP such as citizen science (Haklay, 2013a; Shirk et al., 2012) or e-participation (International Association of Public Participation, 2007). Some types of projects require a high involvement of citizens such that a large share of conducted activities are participatory e.g. co-created projects in citizen science (Shirk et al., 2012) or collaboration and empowerment in e-participation (International Association of Public Participation, 2007). Other projects offer very limited possibilities to participate, such as crowdsourcing or distributed intelligence projects (Haklay, 2013a) or information and consultation approaches (International Association of Public Participation, 2007).

*2. Type of Participation:* In addition to the scope of participatory activities, the type of participatory work differs within subcategories of DIP such as e-participation or crowd-X. Focusing on the spectrum of public participation, we note that the first level of informing citizens includes no active participation from the side of the participants. On the other hand, all other levels include feedback or collaborative work, representing an effort to be provided. For crowd-X, usually, all forms imply active participation of participants, however, some are based on providing a resource (e.g. money in crowdfunding), while others comprise mental effort (e.g. ideas in crowdstorming) (Estellés-Arolas et al., 2015).

*3. Structure of Participation:* Especially in crowd-X and citizen science, different concepts require participants to follow a different work structure. From the definition of crowdsourcing, we draw that there can be teamwork or work undertaken individually (Howe, 2006b). Similar participation structures can be observed in citizen science, where experts differentiate between crowdsourcing or contributory citizen science, typically involving individual contributions, and co-created or collegial citizen science requiring teamwork (Shirk et al., 2012).

*4.-5. Outcomes and Publicity:* Another conceptual difference within as well as across subdimensions of DIP comprises the intended outcomes of the participation processes. In e-participation, we note that the first level in the spectrum of public participation comprises knowledge sharing while the other levels additionally work towards making or refining a decision (International Association of Public Participation, 2007). These outcomes are usually publicly available, for instance, decisions are announced on a municipal level. For citizen science, traditionally, the intention is producing knowledge, by contributing to a research project (Spasiano et al., 2021). While it is desirable to share project outcomes publicly, in reality, the public often has to wait for a publication of the scientific work, that in turn may only be available for a restricted group (European Citizen Science Association, 2015). In comparison to that, crowd-X is an umbrella term for a variety of concepts that target different outcomes. For instance, crowdfunding can result in the production of a certain product, while crowdopinion can target decision-making (Estellés-Arolas et al., 2015). On the other hand, crowdcontent processes such as crowd-searching or -analyzing produce knowledge contributions (Estellés-Arolas et al., 2015).

*6.-7. Project Driver and Owner:* Looking at the different stakeholders involved, we identify four important roles within citizen science, e-participation, and crowd-X. For all subgroups we identify citizens and experts that can either act as a group (crowd/ organization) or individually (non-expert individual/ expert). Depending on who fulfills the role of initiating a project and driving it during the project execution, participation projects can be classified: An example from citizen science comprises contractual projects that are crowd initiatives that are however driven by one or multiple experts, while co-created projects are initiated by experts but driven in equal partnership (Shirk et al., 2012). E-participation tools can be differentiated into citizen initiatives and government initiatives (Van Dijk, 2012): In e-petitioning citizens can individually or collectively initiate projects opening up a petition, while for example e-voting procedures are usually initiated by a public organization (Van Dijk, 2012).

*8. Reasons for the Participatory Design:* Comparing concepts across subdomains in DIP, there are conceptual differences in the reasons why projects are designed in a participatory manner. In e-participation, the spectrum of public participation is grounded in the belief that participation in a decision is a fundamental right of those who are affected by it – mostly the citizens (International Association of Public Participation, 2007). Additionally, e-participation approaches are used to shape relationships between citizens and administration in a more democratic and trustful way (Van Dijk, 2012), implying acceptance and legitimation as motivating factors. For citizen science, traditionally, three perspectives of different research fields exist: the sociological, the natural science, and the policy perspective (Levy and Germonprez, 2017). While the sociological perspective indicates that there are desirable effects on society, which can motivate the project initiation, the natural science perspective adds a more economic reasoning for creating scientific endeavors (Levy and Germonprez, 2017). Crowd-X became popular, first and foremost because it enables companies or individuals to profit from the crowd's input (Estellés-Arolas et al., 2015). This can be accomplished through outsourcing work in any form, but also physically by receiving funding (Estellés-Arolas et al., 2015).

*9. Direction of Communication:* Analyzing differences within subcategories of DIP, we find that participation projects differ regarding the direction of communication flows between stakeholders. There are one-sided formats, where mainly one actor communicates information towards another, such as information (International Association of Public Participation, 2007), citizen science crowdsourcing (Haklay, 2013a) or crowdfunding (Estellés-Arolas et al., 2015). Other formats are designed to generate a dialogue between two parties, such as consultations (International Association of Public Participation, 2007) or crowdopinion (Estellés-Arolas et al., 2015), while a third category of projects relies on a communication also between and within parties, which we categorize as multi-sided. This is the case in citizen science, for example, in collegial projects (Shirk et al., 2012), or for crowd-X in crowdstorming (Estellés-Arolas et al., 2015) approaches.

*10. Incentives:* Depending on the possibilities of the initiators and the contents of the participation, different incentive systems exist across and within subdomains of DIP. In crowd-X, frameworks differentiate between tangible (e.g., money) and intangible (e.g., recognition) rewards. Furthermore, in citizen science, it is acknowledged that depending on the project design, different incentive schemes are needed (Haklay, 2013a). Participants are often intrinsically motivated through, e.g., curiosity or an interest in the topic or science (Haklay, 2013a; Jennett et al., 2016). Another reason for participation is the assumption that it is important to contribute to science, focusing on the created impact (Jennett et al., 2016). Both self-related intrinsic and impact-related motivations can be observed in e-participation: Through e-participation, participants can articulate their ideas, issues or complaints to administrative institutions, hoping to influence an agenda, policy or decision (Van Dijk, 2012).

**Empirical Iterations (C2):** Building upon the conceptual iterations, we evaluate and enrich the initial set of dimensions and characteristics by assessing real-life project examples of DIP in a group of expertise consisting of three experts. For the conduction of the empirical iterations we thus utilize three inputs. As for prior iterations, we use the meta-characteristics for guidance. Additionally, we make use of a sample of nine easily accessible projects from participation platforms (see Table 2.1) that have been mentioned in the literature or the public discourse (Nickerson et al., 2013). Finally, we employ the objective ending criteria defined in (B) to determine the conclusion of the process. Through this approach we find support for our initial set of dimensions, additionally adding seven dimensions grouping between two to three characteristics. The new dimensions will be explained in the following.

| Domain | Platforms |
|---|---|
| citizen science | https://eu-citizen.science; https://www.spotteron.net; https://www.zooniverse.org |
| crowd-X | https://www.mturk.com; https://www.kickstarter.com; https://ideas.lego.com |
| e-participation | https://www.change.org; https://www.democratsabroad.org; https://liqd.net |

Table 2.1.: Platforms utilized for project extraction in empirical iterations.

*1. Implementation:* Notably, digital projects used different infrastructure designs to implement mechanisms of participation. While some projects (2) mainly relied on synchronous digital formats (e.g. digital meetings), others were based on asynchronous formats. For this category we identified projects facilitated via a web-based platform (7) and projects that utilize a mobile application (3).
*2. Enabling Effort:* Analyzing the necessary infrastructure we identify that some projects (3) have prerequisites such as specialized equipment or skills that are necessary for participation. This results in a high enabling effort in order to involve people (e.g. training sessions or the distribution of physical resources). Other projects involve only equipment, knowledge, and skills that can be assumed given for the majority of people (e.g., mobile phone camera), implying little to no effort to enable participation.
*3. Suggestions feedbacked:* Depending on the design of the platform, project structures differed regarding the handling of citizens' contributions. While for some initiatives (5), feedback in any form is not necessary, many projects (6) generally foresee feedback on the citizens' contributions. Some project structures (2) generate this feedback through experts while others rely on crowd feedback (4).
*4. Participation Offer:* Reviewing different participation projects from the field, we identified that they differ in terms of the participatory tasks variety that is offered to the citizens. While some projects (4) propose only one distinct activity (e.g. signing a petition), other projects offer multiple tasks. For some projects (3), participants can self-select which tasks they want to do (e.g., tasks are independent); for others, a successful participation includes doing all the tasks (e.g., a series of sequential activities).
*5. Time Requirements:* Depending on the design of the participation project and infrastructure, initiatives differed in their time requirements for a successful participation. Some projects (5) clearly focus on brief participation processes (e.g. taking a picture with a mobile application or giving a signature digitally). One other project offered tasks of different time intensity; therefore, citizens can self-select how much time they want to spend on the project. A final category of projects (3) offers one or more tasks that are time-consuming (e.g. data collection over a longer time period or innovating a product).
*6. Target Group:* A large share of projects (6) is open to the general public, meaning that everyone who is interested can participate in the project (e.g. everyone can create an account, visit the platform or download the application). For other projects, participation is only under certain conditions possible: We identify restricted projects (2) that allow participation for everyone who fulfills certain criteria (e.g. project initiators can set filter criteria for the participant selection). Closed projects (1), however, describe an involvement project executed with a specified group of volunteers (e.g. a project executed with a certain school class).
*7. Community Building:* A final difference regarding the different handling of the projects' individuals comprised features to support a community building within a project. While some projects (5) clearly intend to build a community between interested members (e.g., through the option to follow or support each other), other projects (4) do not include such structures.

**Organizing the taxonomy:** Conceptual and empirical iterations identified 17 dimensions, each with two to five characteristics. We introduce a second-level ordering of the dimensions (Kundisch et al., 2021) to maintain practical overview and thematically group the 17 (sub)-dimensions into six (superior) dimensions. For 10 of the 17 sub-dimensions, the quality criterion of mutual exclusivity could be employed. For the remaining dimensions, it was more reasonable to model non-exclusive characteristics, due to the complexity of project designs and their activities even in subcategories of DIP (Estellés-Arolas et al., 2015; Spasiano et al., 2021). An overview of the results can be seen in Figure 2.2

| Dim-ension | Sub-Dimension | Characteristics | | | | |
|---|---|---|---|---|---|---|
| Degree of Participation | D1 Participation degree | High (4) | Low (5) | | | |
| | D2 Participation offer | Single task (4) | Multiple tasks optional (3) | Multiple tasks mandatory (2) | | |
| | D3 Type of participation | Active-effort (7) | Active-ressources (3) | Passive (1) | | |
| Implementation of Participation | D4 Implementation | Asynchronous-web-based platform (7) | Asynchronous-mobile application (3) | Synchronous (2) | | |
| | D5 Structure of participation | Team work/ participation (4) | Individual work/ participation (9) | | | |
| | D6 Time requirements | High (3) | Low (5) | Self-selected (1) | | |
| | D7 Enabling effort | High (3) | Low (6) | | | |
| Incentives | D8 Incentives for participation | Self-related extrinsic (3) | Self-related intrinsic (5) | Impact-related (7) | | |
| | D9 Reasons for the participatory design | Acceptance & legitimation(2) | Funding (1) | Access and ressources (4) | Value-based (3) | Profit maximisation (2) |
| Communication | D10 Direction of communication | Top-down (2) | Two-sided (2) | Multi-sided (5) | | |
| | D11 Suggestions feedbacked | Expert feedback (2) | Crowd feedback (4) | No feedback (5) | | |
| | D12 Community building | Yes (5) | No (4) | | | |
| Project Stakeholder | D13 Project driver | Crowd (2) | Non-expert individual (1) | Organziation/ expert (5) | Equal partnership (1) | |
| | D14 Project owner | Crowd (1) | Non-expert individual (1) | Organization/ expert (7) | | |
| | D15 Target group | Open (6) | Restricted (2) | Closed (1) | | |
| Gains and Outcomes | D16 Project outcome | Product (4) | Knowledge (4) | Decision (5) | Sharing things (3) | |
| | D17 Publicity of the outcome | Public (5) | Accessible for the participants (1) | Non-public (3) | | |

Mutually exclusive dimension   Non-mutually exclusive dimension

Figure 2.2.: Initial taxonomy applied to nine projects from the field.

### 2.4.4. Demonstration and Evaluation

To demonstrate and evaluate the taxonomy for DIP (D), a focus group with seven industry practitioners was held, to assess the taxonomy's applicability and usefulness. The practitioners came from for-profit companies and nonprofit organizations where they fulfilled various roles e.g., project or program management, fundraising, or platform operation. In general, they stated that they have little prior knowledge of taxonomies and have not come across a taxonomy for DIP. Some reported typologies used for specific subsets of participation projects, however, participants indicated that the development of a superior taxonomy could be useful for several reasons: On the one hand, a taxonomy could unify the understanding of participation from different stakeholders in the participation formats. By standardizing terminologies across domains and across different language areas, collaboration and communication could be eased. One participant opinion should be presented to illustrate this point: *"[...] with the topic of participation, you notice in an exceptional way when you talk to different actors that everyone understands something different. [...] I would say that it definitely helps if you just use the same vocabulary – you are referring to the same things"*. On the other hand, a taxonomy could be helpful to classify oneself as a digital platform provider regarding one's participation offers. In this sense a taxonomy could be used for expanding and refining existing participation formats. Above that one participant added about the utility of a taxonomy for setting up a project that it would be *"[...] a bit like checking, have I taken everything into account, or have I now forgotten something important, or could I even add an innovation"*. When presenting and jointly applying our taxonomy for DIP, a first impression of the practitioner's individual assessment was collected through a questionnaire[1]. Overall, the taxonomy could score positive results in all ending criteria, which means, on average, the seven practitioners agreed that the taxonomy is concise, robust, comprehensive, extendible, and explanatory ($Mean > 4$). While the approval was the strongest for the extendibility of the taxonomy ($Mean = 5.43$), it was the weakest for its comprehensiveness ($Mean = 4.21$) and scored the most diverse results regarding its robustness ($Mean = 4.29$, Standard Deviation $(SD) = 1.33$). The questionnaire results were displayed in the focus group as a discussion basis, which made it possible to gather more detailed insights into the practitioners' assessments, as well as concrete proposals for improvement of the taxonomy. As for the questionnaire, the discussion was structured by the five dimensions of subjective ending criteria:

*Conciseness:* Participants stated that the taxonomy enables a suitable overview and the framework gives room to further add aspects. Its compactness seems to be its strength and thereby creating an environment in which relatively little time is needed to understand and apply the taxonomy. However, a discussion evolved on whether the taxonomy might not be concrete enough for its application in certain subdomains

[1] Two items per ending criterion, self-developed based on Nickerson et al. (2013) and measured on a 7-point Likert scale: The taxonomy… 1a) is concise. 1b) is too extensive. 2a) includes enough dimensions to sufficiently describe all objects of DIP. 2b) includes too little dimensions to adequately differentiate between the objects of interest. 3a) includes all dimensions of interest to classify objects of DIP. 3b) is not comprehensive for DIP. 4a) is extendible in case of new developments in the category of DIP. 4b) can not be adapted to fit future developments in the category of DIP. 5a) provides useful explanations of the nature of the objects under study. 5b) is too detailed.

of DIP. Some dimensions might need a stronger differentiation. However, adding dimensions that are very unique to certain domains would increase the overall complexity. A consideration in which a prioritization was made in favor of creating a concise taxonomy.

*Robustness:* The discussion about insufficient differentiation of the taxonomy was taken up again when it came to its robustness. However, the practitioners also emphasized the importance of the taxonomy's dimensions keeping the big picture in mind and not getting lost in individual aspects of specific areas. In this sense, they expressed that looking at characteristics of well-known dimensions that are unusual in their domain, can be very valuable for rethinking their own practices. However, they pointed out that certain dimensions could not be classified as value-free for some participation domains. One citizen science practitioner formulated, for example, regarding the subdimension participation degree: *"I suspect that many would mark "high" because it's almost desired to mark high there. [. . . ] High and low are almost judgmental [categories] when it comes to the participation degree."*

*Comprehensiveness:* Participants made several proposals on how to further enrich or improve the taxonomy. In terms of improving existing dimensions and characteristics, the sub-dimension degree of participation was perceived ambiguously: Participants pointed out that the classification into high and low might not sufficiently depict the reality, but rather, a sliding scale would be required. Otherwise, the sub-dimension should be further divided into sub-categories or redefined. Another improvement targeted the sub-dimension enabling effort, where participants added that it would be necessary to distinguish between knowledge and technical resource prerequisites. Regarding the platform communication, participants suggested a renaming of the characteristic top-down to one-sided and pointed out that an important enrichment for the taxonomy would be to discuss next to the communication also the direction of power. Additionally, participants discussed whether it would be of interest to add details about aspects such as moderation, communication features, the type of desired input, or the duration in which the communication is stored. Furthermore, regarding the ownership of the project, the practitioners raised the question of whether the relevant question would be who the project owner is in terms of infrastructure or in terms of power. Additionally, one participant reported, that funding might be a reason for the participatory design in the public sector. For some practitioners, including participative offers could lead to advantages when it comes to applying for public funding, in comparison to crowdfunding, which was meant by this characteristic. A final improvement targeted the sub-dimension implementation, where participants stressed that it would be important to capture whether there is an analogous or hybrid component to the digital project. Overall, participants agreed that most of the proposed dimensions are already very useful, but would need additional explanations and examples to exclude misunderstandings and ambiguity. As such one participant formulated: *"We had talked a bit in our group [. . . ] and asked if there could be explanations to the subdimensions [. . . ]. If the terms should stand for themselves then the terms should not be so ambiguous as they are right now."*

*Extendibility:* The practitioners consented that the frame the taxonomy provides would allow for an easy adaption in case of further developments either through adding dimensions or characteristics or splitting up existing aspects.

*Explanatory capabilities:* In this category the participants recapped that the added value for them as practitioners would be especially to be able to compare their own participation activities to others across domains. One participant highlighted the value as follows: *"in terms of such a broad participation landscape, just understanding whether I am in a specific bubble of participation processes that are all low-threshold citizen processes, or what other projects might be in this landscape; where do others position themselves?; How can we mutually benefit from each other?"* Additionally, the taxonomy would be very helpful to evaluate one's motives, views, and assessments, which could lead to more practical delineations. One participant used the following words for the description: *"There are simply so many different participatory formats [...] that have somehow found their own term and are then very eager to distinguish themselves from other participatory formats. [...] if you take these different types that are circulating and say now please fill out the form [taxonomy, . . . ] perhaps this also becomes tangible in such a self-assessment, and no longer runs that much on an anecdotal basis."*

In conclusion, we extract from the focus group evaluation a strong support for the general need for a taxonomy for DIP and a satisfying result for its usefulness. To further improve utility and applicability, we collected some ideas for optimization and the demand to find a communication format that accounts for more explanations and examples.

### 2.4.5. Communication

For the communication of our first DSR cycle's results (E), we first implemented the points for improvement collected from the practitioners and then instantiated a version of the taxonomy in a navigable web application prototype. The results of the improvement cycle can be seen in Figure 2.3. Based on the practitioners' feedback, we add two necessary dimensions, being the participation format and the moderation. The prior accounts for differences in completely digital versus mixed formats that either have parallel or sequential analogous and digital formats. The latter describes whether or not dialogue is moderated and by whom. Additionally, we revise the subdimension degree of participation based on a more detailed framework (Somech, 2002) and rename characteristics in the direction of communication, project owner and driver, and the subdimension enabling effort. As such, the final taxonomy of our first DSR cycle includes 19 dimensions, each grouping between two to five characteristics. To be able to better explain the dimensions and characteristics for the usage by practitioners, we design a new format for communication, which can be seen in Figure 2.4. The navigable web application prototype guides the user through the respective dimensions and sub-dimensions of the taxonomy, allowing them to choose characteristics based on detailed explanations (see Figure 2.4, top). Upon this successive classification, the user receives an overview of their selected characteristics in an overview page (see Figure 2.4, bottom). With this communication format, we intend to better reach DIP practitioners and circumvent the problem that individual terms can have different connotations in certain domains or roles, which can lead

| Dimension | Sub-Dimension | Characteristics | | | | |
|---|---|---|---|---|---|---|
| Degree of Participation | D1 Extent of participation | Information sharing | Consultative | Democratic | | |
| | D2 Participation offer | Single task | Multiple tasks optional | Multiple tasks mandatory | | |
| | D3 Type of participation | Active-effort | Active-ressources | Passive | | |
| Implementation of Participation | D4 Format | Digital | Analog/ digital (parallel) | Analog/ digital (sequential) | | |
| | D5 Implementation | Asynchronous-web-based platform | Asynchronous-mobile application | Synchronous | | |
| | D6 Structure of participation | Team work/ participation | Individual work/ participation | | | |
| | D7 Time requirements | High | Low | Self-selected | | |
| | D8 Prerequesites | Domain knowledge | Domain-specific equipment | Assumes preconditions | | |
| Incentives | D9 Incentives for participation | Self-related extrinsic | Self-related intrinsic | Impact-related | | |
| | D10 Reasons for the participatory design | Acceptance & legitimation | Funding | Access and ressources | Value-based | Profit maximisation |
| Communication | D11 Direction of communication | One-sided | Two-sided | Multi-sided | | |
| | D12 Suggestions feedbacked | Expert feedback | Crowd feedback | No feedback | | |
| | D13 Community building | Yes | No | | | |
| | D14 Moderation | Crowd | Individual from crowd | Organization/ expert (intern) | Organization/ expert (extern) | None |
| Project Stakeholder | D15 Project driver | Crowd | Individual from crowd | Organziation/ expert | Equal partnership | |
| | D16 Project owner | Crowd | Individual from crowd | Organization/ expert | | |
| | D17 Target group | Open | Restricted | Closed | | |
| Gains and Outcomes | D18 Project outcome | Product | Knowledge | Decision | Sharing things | |
| | D19 Publicity of the outcome | Public | Accessible for the participants | Non-public | | |

Mutually exclusive dimension   Non-mutually exclusive dimension

Figure 2.3.: Revised taxonomy after the evaluation.

to misunderstandings. In the upcoming cycles, the prototype will be evaluated and further developed into a publicly available web application. This will enable us to easily describe and classify a large set of DIP in an illustrative scenario-based evaluation (Szopinski et al., 2019) while capturing and presenting the results directly in the application. In this way, future users can directly compare their participation initiatives with others after the classification. In the third DSR cycle, we will reflect on our taxonomy improvements together with an extended circle of industry practitioners from the focus group through a survey evaluation, thus, communicating the final web application directly to its target user group.


Figure 2.4.: Snapshots of the navigable web application prototype; Classification page with detailed explanations (top) and results page showing an overview of the chosen combinations of characteristics (bottom).

## 2.5. Discussion and Conclusion

Digital participation, as it appears in government processes in e-participation, in business initiatives in crowd-X, or in the scientific community in citizen science, is a continuously increasing development, creating more and more demand for adequate infrastructure (Estellés-Arolas et al., 2015; Liu et al., 2021; Shirk et al., 2012; Van Dijk, 2012). Instead of joining forces, domains undertake efforts to distinguish their domains from another and thereby failing to achieve uniform definitions which might strengthen the collaboration. In our work, we, therefore, create the umbrella term DIP that allows us to discuss digital involvement concepts across domains solely based on their characteristics. Following the iterative design process by Nickerson et al. (2013), we derive a taxonomy for DIP grounded in the domains of e-participation, crowd-X, and citizen science based on conceptual and empirical iterations. We ensure its practical relevance and applicability through an evaluation and refinement of the taxonomy with practitioners from the field. After the first DSR cycle, our taxonomy includes 19 dimensions to describe key characteristics around the participation process, its individuals, and the digital infrastructure. With this theoretical contribution, we are the first to contribute to a framework, enabling the structured analysis and comparison of DIP subgroups based on their design and characteristics instead of conceptual definitions. Dimensions and characteristics such as degree of participation or communication can make design choices transparent, that are influential for a project's outcome like effects on empowerment and learning (De Albuquerque and Almeida, 2020; Palacin et al., 2020). Upon the creation of an appropriate database, the taxonomic descriptions could help advance further research on these topics, sharpen existing conceptual boundaries of DIP subgroups, or identify existing and new relevant archetypes within DIP.

As for practical implications, with the taxonomy, and especially the navigable web application, we provide a tool, practitioners can deploy to analyze and characterize their participation offer or initiatives and strategically compare them to others. On the other hand, it provides the possibility to generate new ideas and develop new forms of participation projects by identifying new combinations of characteristics. Although our evaluation underlined the difficulty of creating a framework fitting and presenting the variety of DIP, it especially showed that the taxonomy is a solid basis for discussion that is urgently needed. By providing a common language and a structured framework, it enables practitioners to talk about DIP thus potentially increasing knowledge sharing and mutual learning. Especially upon the creation of a database, practitioners lacking expertise could more easily identify similar projects across domains and learn from them by comparison or even share used technological infrastructures if necessary, since, for instance, in the field of citizen science, many platforms can be used for free (Stein et al., 2023c).

When utilizing the preliminary taxonomy for DIP, some limitations in its development should be stated. In the attempt to find a starting point for the specification of DIP, we focus on the three domains of e-participation, crowd-X, and citizen science, which shape our taxonomy through its design and evaluation. Thereby we disregard other subparts of DIP, that might add interesting new characteristics or domains to the taxonomy. Examples for other domains that might be valuable to analyze for DIP are movements such as open science, or open innovation. Moreover, in the definition of DIP we solely focus on digital involvement *projects*, which implies a time-bounded feature to the participation process at hand. This, however, may unnecessarily restrict the application of our taxonomy, e.g., to continuous initiatives, and as such, may argue for a reframing of the wording. Additionally, the sample of projects utilized in the taxonomy development process did not strive for completeness nor representativity for DIP. Rather we focused on a limited sample and critically evaluated results with practitioners from the field. Thereby, we ensure practical relevance, however, very rare characteristics or characteristics typical of a particular participation area might have been overlooked in our taxonomy so far. Thus, future work should question what other phenomena match the definition of DIP and evaluate whether the taxonomy enables their classification. A framework for structuring a discussion of DIP, however, provides only limited value to practitioners. The benefits are only fully realizable when utilizing the taxonomy on a large set of projects, thereby identifying archetypes and providing a baseline for comparison and learning. While we encourage both practitioners and theorists to use and further develop the initial taxonomy, we ourselves want to address some of the identified limitations of our taxonomy in the mentioned second and third design cycle.

In this chapter, we presented the results of the first design cycle in a tricyclic DSR project to build a taxonomy for DIP. Our initial taxonomy of 19 dimensions was communicated after ensuring practical relevance and comprehensibility with practitioners from the field. While an empirical evaluation will only follow in cycles two and three of the DSR approach, the methodological approach of the first cycle guarantees the usefulness of this initial taxonomy for the target group, namely project initiators and practitioners, that seek guidance and inspiration. By uniting the fields of e-participation, citizen science, and crowd-X and proposing a framework to discuss projects based on their main characteristics, we provide a first theoretical basis for the new concept of DIP. With this we aim to invite scientists and practitioners likewise to engage in the discussion on DIP, thereby contributing to building a transdisciplinary knowledge base that can be used for research and innovation, as well as knowledge and resource sharing.

Part II.

# Navigating the Design of Digital Involvement

## Abstract Part II

*Building on the concept of digital involvement, Part II, titled "Navigating the Design of Digital Involvement" explores the diverse design approaches within DIP and provides a structured framework for understanding and addressing this variety. With the growing spread of digital involvement across political, economic, and scientific domains in datafied societies, for practitioners, the complexity of selecting appropriate formats and tools continuously increases. At the same time, social polarization and democratic backsliding make it a priority to ensure the effectiveness of citizen involvement, demanding rigorous evaluation and exchange of best practices. Therefore, this part deals with practitioner-oriented research on navigating the design of digital involvement, leveraging the DIP taxonomy. It explores the instantiation of the taxonomy in two practical artifacts. A navigable web application facilitates access to the taxonomy for practitioners and DIP archetypes break down the variety of DIP designs. This research was carried out with Alicia Wittmer, Jonas Fegert, and Christof Weinhardt and presented at the 58th Hawaii International Conference on System Sciences (Stein et al., 2025b). Furthermore, in collaboration with the platform mit:forschen!, the DIP artifacts are applied to the realm of citizen science to shed light on the self-positioning of citizen science projects in the design landscape of digital involvement. As a joint effort together with Moritz Müller and Jonas Fegert, this research was presented at the Conference of the European Citizen Science Association 2024 (Stein et al., 2024a) and has been submitted for review in the current form to the journal Citizen Science Theory & Practice. In this part, titles, tables, and figures from the original publications have been renamed, reformatted, and re-referenced to align with the structure of the dissertation. Additionally, chapter and section numbering were adjusted, and formatting, abbreviations, and references were standardized for consistency.*

# 3. The DIP Web Application and Archetypes

## 3.1. Introduction

The increasing complexity of global societal challenges such as the Covid-19 pandemic or the climate crisis results in governments worldwide seeking new approaches to involving experts, citizens, and other stakeholders in government action (United Nations, 2022). With virtual communication and collaboration becoming the norm, digital participation formats take a central role, and the market for participation and deliberation technologies is rapidly expanding (García et al., 2023; United Nations, 2022). While e-participation, as part of digital government, has been available for two decades, the field experiences lately a diversification of participatory approaches. Concepts emerging from other domains, such as crowdsourcing or citizen science, are increasingly finding their way into the practices of governments and local administrations (Shanley et al., 2019). This holds promising benefits for policymakers, improving their possibilities of finding suitable approaches (Flores and Rezende, 2022; Fritz et al., 2019). However, the dynamic developments fuel the complexity of designing, conducting, and evaluating projects, especially as the design of participation is decisive in ensuring that projects have no adverse effects or become pseudo-participatory (Palacin et al., 2020; Pristl and Billert, 2022). In times of societal polarization and the threat of democratic backsliding (Weinhardt et al., 2024), guiding practitioners in the evolving field of participatory approaches becomes paramount. In addition to the scientific field of e-participation, the research areas of crowdsourcing and citizen science are increasingly relevant to finding orientation and best practices. To benefit from theoretical knowledge across domains Stein et al. (2023b) have developed a unified taxonomy, deriving key design characteristics to describe DIP. While this enables the joint consideration of participation projects, the taxonomy as a theoretical artifact remains somewhat unaccessible to practitioners such as community organizers, government officials, administrative staff, or industry practitioners, who want to leverage existing design knowledge. Thus, we turn to the research question: *How can we use the DIP taxonomy to capture design knowledge and make it available to domain practitioners?*

In a DSR approach, we take upon the work of Stein et al. (2023b), further developing the DIP taxonomy into an interactive web application that we demonstrate and evaluate in a real-life scenario with 60 practitioners. Furthermore, we apply the taxonomy in illustrative scenario analysis to 46 projects, deriving design clusters that we use to formulate preliminary DIP archetypes. Thereby, we aim to contribute to the knowledge base on the diversity of design approaches for digital involvement, making a theoretical taxonomy accessible for both research and practice.

## 3.2. Theoretical Background

Formats of laypeople's digital participation are being discussed in political, economic, and scientific contexts. As a branch of digital government, e-participation focuses on using information and communication technology for consultative and democratic political processes (Macintosh, 2004; Pristl and Billert, 2022; Sanford and Rose, 2007). Similarly, in the economic context, crowdsourcing describes a problem-solving model where digital means enable the involvement of an (often undefined) digital crowd in completing tasks (Estellés-Arolas et al., 2015; Howe, 2006b). Finally, scholars in the field of participatory science subsume research under the term citizen science (Haklay et al., 2021), with online citizen science focusing on digitally mediated forms (Aristeidou and Herodotou, 2020). Individually, frameworks to guide, discuss, and reflect involvement practices have been established in all three research fields (e.g., Estellés-Arolas et al., 2015; Haklay, 2013a; Van Dijk, 2012). However, as crowdsourcing and citizen science practices are gradually finding their way into theory and practice of digital government (e.g., Flores and Rezende, 2022; Halmos et al., 2019; Schmidthuber et al., 2022), traditional e-participation typologies can no longer adequately reflect the diversity of participation formats. With DIP, a concept has been proposed to jointly discuss participatory initiatives, describing "projects that utilize digital platforms for the involvement of multiple external individuals in a defined participation process" (Stein et al., 2023b, p.5). The DIP taxonomy proposes 19 design subdimensions based on five existing theoretical classification frameworks from e-participation (International Association of Public Participation, 2007; Van Dijk, 2012), citizen science (Haklay, 2013a; Shirk et al., 2012) and crowdsourcing (Estellés-Arolas et al., 2015). The taxonomy (see Figure 3.1) includes three subdimensions for the participation level, five subdimensions describing the project's implementation process, two subdimensions on participants and organizers' incentives, four subdimensions depicting communication structures, three subdimensions addressing project stakeholders, and two subdimensions focusing on the projects' outcomes. Eleven of the 19 dimensions employ the mutual exclusivity criterion, while eight allow for multiple selections (gray-shaded). The taxonomy extends classical e-participation frameworks, which differentiate projects, e.g., based on the level of involvement (e.g., Macintosh, 2004) or their driving party (e.g., Van Dijk, 2012), by connecting multiple design dimensions and, importantly, including dimensions that are applicable across contexts, which helps to describe, compare, and potentially include emerging involvement practices in the digital government sector.

## 3.3. Methodological Approach

DSR is a problem-solving paradigm that creates artifacts, constructs, models, methods, and instantiations, providing descriptive or prescriptive knowledge and innovative solutions (March and Smith, 1995; Vom Brocke et al., 2020). It is a suitable method for taxonomy development, as it supports an iterative approach and enables revision based on evaluation results (Kundisch et al., 2021). Building on the initial DSR cycle of Stein et al. (2023b), we aim to further operationalize the taxonomy as an extant

| Dimension | | Characteristics | | | | |
|---|---|---|---|---|---|---|
| Degree of participation | 1.1 Participation extent | Information sharing (52%) | Consultative (30%) | Democratic (17%) | | |
| | 1.2 Participation offer | Single task (65%) | Multiple tasks optional (35%) | Multiple tasks mandatory (0%) | | |
| | 1.3 Type of participation | Active-effort (85%) | Active-resources (17%) | Passive (2%) | | |
| Implementation of participation | 2.1 Format | Digital (96%) | Analog/ digital (parallel) (0%) | Analog/ digital (sequential) (4%) | | |
| | 2.2 Implementation | Asynchronous web-based platform (91%) | Asynchronous mobile app (11%) | Synchronous (0%) | | |
| | 2.3 Structure of participation | Teamwork/ participation (13%) | Individual work/ participation (98%) | | | |
| | 2.4 Time requirements | High (4%) | Low (70%) | Self-selected (26%) | | |
| | 2.5 Enabling effort | Domain knowledge (13%) | Domain-specific equipment (7%) | Assumed preconditions (80%) | | |
| Incentives | 3.1 Incentives for participation | Extrinsic self-related (17%) | Intrinsic self-related (39%) | Impact-related (78%) | | |
| | 3.2 Reasons for the participatory design | Acceptance/ legitimation (28%) | Funding (13%) | Access/ resources (83%) | Values (2%) | Profits (7% ) |
| Communication | 4.1 Direction of communication | One-sided (41%) | Two-sided (22%) | Multi-sided (37%) | | |
| | 4.2 Feedback | Expert feedback (22%) | Crowd feedback (43%) | No feedback (46%) | | |
| | 4.3 Community building | Yes (28%) | No (72%) | | | |
| | 4.4 Moderation | Crowd (0%) | Individual from crowd (4%) | Internal organization/ expert (2%) | External organization/ expert (2%) | None (91%) |
| Project Stakeholder | 5.1 Project driver | Crowd (11%) | Individual from crowd (11%) | Organization/ expert (78%) | Equal partnership (0%) | |
| | 5.2 Project owner | Crowd (4%) | Individual from crowd (7%) | Organization/ expert (89%) | | |
| | 5.3 Target group | Open (80%) | Restricted (20%) | Closed (0%) | | |
| Gains & Outcomes | 5.4 Project outcome | Product (30%) | Knowledge (52%) | Decision (24%) | Sharing (24%) | |
| | 5.5 Publicity of the outcome | Public (67%) | Accessible for the participants (13%) | Non-public (20%) | | |

Figure 3.1.: DIP taxonomy based on Stein et al. (2023b); In parentheses, distribution of 46 sample projects.

model in the forms of additional design artifacts, i.e., instantiations and models (Gregor and Hevner, 2013; March and Smith, 1995; Vom Brocke et al., 2020). Methodologically, the first cycle integrated taxonomy development within Peffers et al. (2007) DSR approach (Kundisch et al., 2021; Nickerson et al., 2013) (see Figure 3.2, Cycle 1). Consequently, we also structure our research activities within the established procedure of problem identification, definition of objectives, design, and development, and demonstration and evaluation (Peffers et al., 2007) (see Figure 3.2, Cycle 2). We started our DSR cycle by reflecting on the DIP taxonomy and its development, motivating further artifact development, and deriving objectives for improved solutions. Identifying two independent areas for improvement, we proceeded in two individual subcycles (a) and (b), before combining the endeavors for their communication. In subcycle (a), we improved the usability of the taxonomy as a theoretical framework by contributing a physical implementation in the form of an interactive web application. While the initial DSR cycle emphasized qualitative evaluation, we set a quantitative focus when assessing the taxonomy as a theoretical model, and the web application as its situated instantiation (Gregor and Hevner, 2013; March and Smith, 1995). We demonstrated the artifact in a field test to domain practitioners, asking them to use the web application themselves to classify their individual DIP. Consecutively they were asked to share

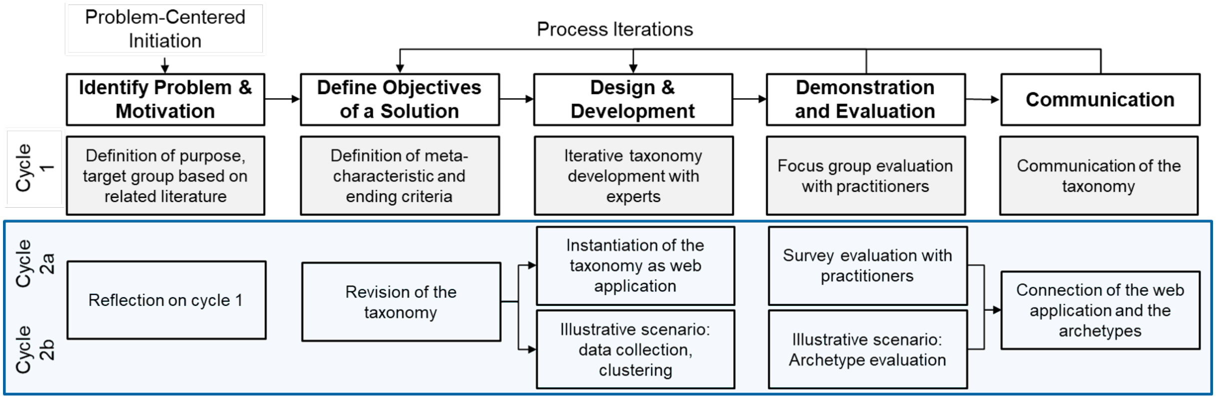


Figure 3.2.: DSR cycles based on Peffers et al. (2007) with the current research activities (in blue) and the previous research activities (in grey).

their experiences through a survey, a standard method for both taxonomy and software testing (Gediga et al., 2002; Szopinski et al., 2019). The survey evaluated qualitative criteria for the taxonomy's quality (Nickerson et al., 2013; Stein et al., 2023b) (one positive and one negative item per criterion arranged on a scale of -3 to 3), and the net promoter score (Sasmito et al., 2019), as a means of determining overall satisfaction with the web application. To reach a large group of active DIP practitioners, we collaborated with the largest German citizen science platform, "mit:forschen!" (e.g., Moczek et al., 2021), contacting 240 project initiators between June and August 2023. 60 of them utilized the web application to classify their project and 24 proceeded to take part in the survey evaluation. No monetary incentive for application utilization and evaluation was set. Second, subcycle (b) focused on using the DIP taxonomy to capture current design practices in the form of operationalized models, the DIP archetypes. We followed an illustrative scenario approach, applying the taxonomy to real-world projects (Szopinski et al., 2019). To sample a diverse set of projects, we utilized both scientific and public sources. From the literature, overviews for e-participation (Fegert, 2022) and citizen science platforms (Stein et al., 2023c) were used to randomly extract projects from all presented platforms. Additionally, a random sample of projects was selected from public project overview websites on crowdsourcing projects and internet-based activism (Wikipedia, 2020, 2024), resulting in a total of 46 DIP projects. Two independent reviewers coded the project examples based on their websites and publicly available information, achieving a Cohen's Kappa of 0.767, indicating substantial agreement (Landis and Koch, 1977). To identify common design patterns and relationships among the DIP projects, we used hierarchical, agglomerative clustering following guidance by Kassambara (2017). This technique enables the evaluation of similarities of objects, sorting them into a tree-based representation (Kassambara, 2017; Romesburg, 2004). Working with a binary data structure, we measured object distance based on the Yule similarity (Choi et al., 2010), suitable for clustering observations based on overall profiles rather than magnitudes (Kassambara, 2017). The clustering results were assessed with the silhouette coefficient (Range -1 poor to 1 optimal) (Kassambara, 2017; Rousseeuw, 1987) and qualitatively evaluated by two domain experts to derive preliminary DIP archetypes.

## 3.4. Results of the DSR Activities

### 3.4.1. Reflection and Definition of Objectives

The DSR cycle of Stein et al. (2023b) produced an initial DIP taxonomy, developed through a combination of literature review and practical examples, and evaluated by an expert focus group. The evaluation confirmed the need for a DIP taxonomy but also highlighted areas for improvement and further development. This included content adjustments for dimensions 1.1, 2.1, 2.4, 4.1, and 4.4, which have already been addressed but not yet evaluated. Additionally, opportunities to enhance usability and explanatory clarity remained unexplored: Currently, the taxonomy, as a model artifact, is challenging for its primary users being practitioners, as focus group participants noted difficulties with terminology and the need for more explanations. This motivates considerations of instantiating the taxonomy in a more flexible and interactive form (Stein et al., 2023b). We embrace this notion, directing our focus toward designing and developing a publicly accessible web application (subcycle a). The artifact should ease access to the theoretical construct while providing necessary guidance in the application process. Thereby, practitioners could independently access the design knowledge while contributing to it when capturing their project designs. Furthermore, although the DIP taxonomy is rooted in theoretical and practical knowledge, its application in the creation process was limited to nine project examples. Therefore, a comprehensive knowledge base for practitioners on typical design configurations and their differences is still lacking. Given that this was seen as a major potential of the DIP taxonomy (Stein et al., 2023b), a more thorough analysis of illustrative scenarios is essential. By expanding the application of the taxonomy to a broader spectrum of project examples, we aim to uncover existing design patterns and ultimately delineate them as archetypical projects. Archetypes help to investigate configurations and dynamics of IS phenomena (Schilling et al., 2017) and can thus contribute to capturing additional design knowledge in the DIP research field.

### 3.4.2. DIP Web Application

#### 3.4.2.1. Design and Development

The design of the web application aims to enhance the understandability of the taxonomy. As such, we enrich the taxonomy by developing brief explanations for its contents, e.g., subdimensions and characteristics, to provide them for more inclusive reach in German and English. Furthermore, the objective is to ease access and empower practitioners to contribute to the knowledge base. We thus design an interactive experience, leading through the classification process of a DIP project (see Figure 3.3, top-left). For every dimension, a page is developed with the corresponding subdimensions and their characteristics and provides interactivity by selecting fitting project characteristics. To gain further insights into the user's experiences, for every subdimension, users can indicate difficulties with the selection if it occurred. Finally, the user receives an overview of the chosen design configuration and can decide to share this classification, providing additional details on the project (see Figure 3.3, bottom-left). We develop the web application utilizing the vue.js framework for frontend development and integration to MongoDB for the backend.

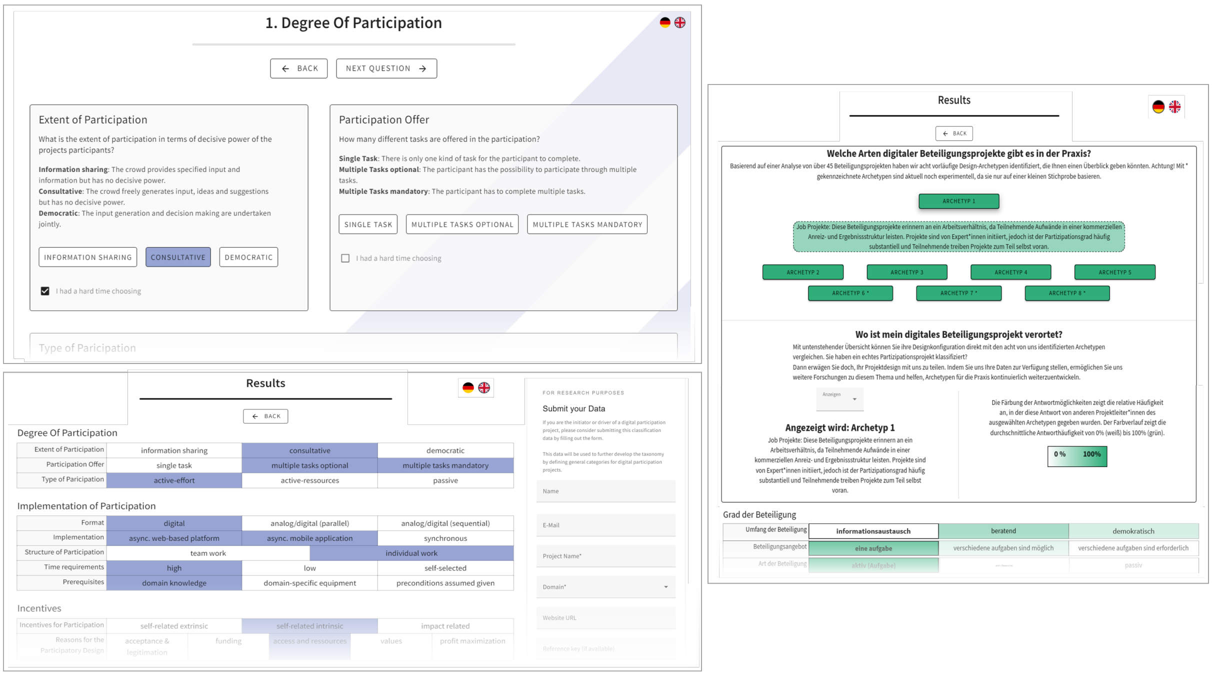


Figure 3.3.: DIP web application publicly available under hop.fzi.de/taxonomy (left) and prototypical extended results page (right).

#### 3.4.2.2. mit:forschen! Demonstration and Evaluation

We received 60 project descriptions from coordinators via our web application. In evaluating the difficulty of dimension selection, only four out of the 19 subdimensions were flagged by more than 10% of the users ($n = 60$). These included the extent of participation (0.30), reasons for the participatory design (0.25), community building (0.13), and time requirements (0.12). Feedback from an optional open-text entry provides insights into associated challenges. One participant stated: *"The project offers many different tasks to participate in. Some of the questions, therefore, do not apply equally to all offers."*; while another added: *"Sometimes several answers apply or something in between[..]"*. For objective quality criteria, mean values for all five criteria exceeded the neutral level of 0 (Scale -3 to 3, $n = 24$). The highest approval was given for explanatory capabilities ($Mean = 1.67$, $SD = 0.87$) and extendibility ($Mean = 1.38$, $SD = 0.97$), followed by conciseness and robustness ($Mean = 0.88$, $SD = 1.33$ and $SD = 1.57$). Comprehensiveness received the lowest approval ($Mean = 0.83$, $SD = 1.43$). Evaluating participants' willingness to recommend (Scale 0-10), we found, on average, a positive likelihood that participants would recommend the DIP web application to a colleague ($Mean = 6.29$, $SD = 2.46$).

### 3.4.3. Illustrative Scenario and DIP Archetypes

#### 3.4.3.1. Illustrative Scenario Analysis

Applying the taxonomy to a sample of 46 projects revealed a rich variety of design configurations. Figure 3.1 visualizes the distribution of characteristics in the sample. While for most dimensions, the samples' features are diverse, in some, we find particularly little variation, and some characteristics are not present. Specifically, the format is predominantly digital, with few analogous components, and moderation is rarely present. Additionally, the sample lacks projects offering multiple obligatory tasks, synchronous participation, closed target groups, and projects driven by equal partnerships.

Using hierarchical clustering on the dataset, we identify an optimal number of three clusters, with an average silhouette coefficient of 0.49. Beyond this, the average silhouette coefficient decreases until a sharp increase is observed at eight clusters (0.47). The dendrogram indicates that the three-cluster solution produces two smaller clusters (see Archetypes 1, 6) and one larger cluster (see Archetypes 2-5, 7-8). When expanding to eight clusters, the larger cluster splits into six, resulting in more balanced partitions.

#### 3.4.3.2. DIP Archetypes

Evaluating the eight-cluster solution, we identified five robust (five to twelve project examples) and three experimental ($<$ five project examples) DIP archetypes, which we describe in the following and have summarized visually in Figure 3.4.

**Archetype 1 "Job Projects":** This cluster contains six projects that resemble employment relationships, characterized by extrinsic participation incentives (100%) and non-public product results (100%) (Figure 3.4, dark-blue graph). The projects have a medium to high degree of participation, i.e., a consultative (83%) or democratic character (17%) and one (67%) or more (33%) effort-based participation tasks. Although participation is exclusively digital, specialist knowledge is often required (67%), and time expenditure can vary greatly, resulting in a medium to high participation threshold. The incentive and result structure is commercial: extrinsic incentives, non-public product results, access to knowledge/ resources (100%), and partially profits (50%). The communication is limited, with predominantly two-sided communication (83%), often without feedback (83%), no moderation (100%), and limited community building (67%). Projects are initiated by an expert/ organization (100%) but have different drivers and are open for everyone to join.

**Archetype 2 "Research Projects with Crowdsourced Tasks":** The cluster includes twelve projects, characterized as research projects with crowdsourced tasks, primarily due to their knowledge outcomes (see Figure 3.4, yellow graph). They exhibit a low to medium degree of participation, either informa-

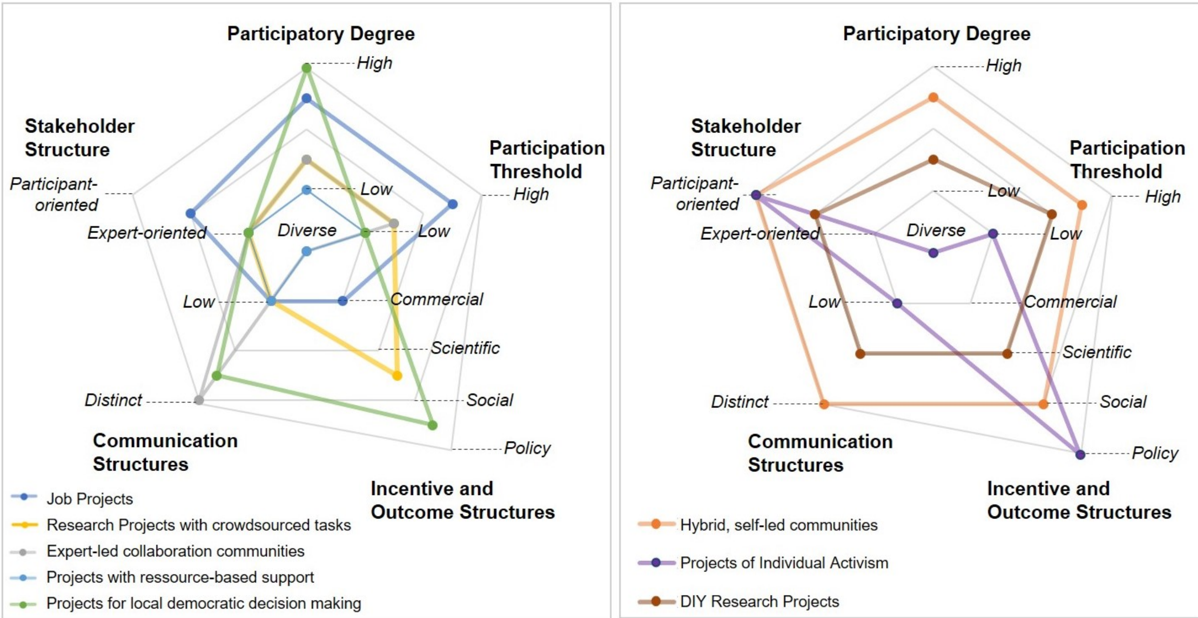


Figure 3.4.: DIP archetypes with robust archetypes (left) and experimental archetypes (right).

tive (58%) or consultative (42%), with typically one effort-based task (83%). Projects have a low to medium participation threshold, being web-based (75%) with low time requirements (83%) and individual work (100%), though some require specialized equipment (17%) or a mobile app (25%). Incentives and outcomes are socially and scientifically oriented: impact-related incentives (100%) and access-based motivators on the initiators' side (100%) to generate knowledge outcomes with varying publicity. Communication is limited, typically one-sided (83%), without feedback (75%), community building (83%), or moderation (100%). Initiated and driven by experts or organizations (100%), these projects generally invite everyone (83%), though some restrict participation (16%).

**Archetype 3 "Expert-led Collaboration Communities":** This cluster contains eleven projects described as expert-led collaboration communities (Figure 3.4, grey graph). They are community-oriented but lack decisive power for the community and are initiated and driven by experts. The participatory degree is low to medium, with low decisive power (82%), but one (55%) or multiple (45%) effort-based tasks. Participation thresholds are low to medium, being web-based (91%), with low (72%) or self-selected time requirements (27%), individual work (100%), and no prerequisites (91%). Communication means are distinct, including multi-sided communication (82%), feedback from participants (100%) and experts (27%), and means to promote community building (55%). The incentive and outcome structure is more diverse than for cluster 2: Although the need for access and resources remains a motivator for initiators (100%), the incentives for participants vary, and different public outcomes are achieved.

**Archetype 4 "Projects with Resource-Based Support":** This cluster includes five projects where participation is limited to resource-based support (Figure 3.4, light-blue graph). They have a low participatory degree, being informative (100%) with a single participatory task (100%), either resource-based

(80%) or passive (20%). Participation thresholds are low, with no prerequisites (100%), low time requirements (100%), and individual work (100%). Communication is limited, with one-sided communication (80%), no feedback (80%), moderation (80%), or community building (80%). The incentive and outcome structure is very diverse, with intrinsic (40%) and impact-related incentives (80%), financing (80%) or acceptance-related (20%) motivators, and public or non-public outcomes, including products (60%), sharing of things (20%) or knowledge (20%). Projects are initiated and driven by experts and open to join for everyone (100%).

**Archetype 5 "Projects for Local, Democratic Decision Making":** This cluster contains five projects focused on democratic decision-making, i.e., they are all democratic and produce a decision (Figure 3.4, green graph). Thus, their participatory degree is high with participants engaging in one (40%) or multiple effort-based tasks (60%). Projects have low thresholds for participation, i.e., web-based with no prerequisites (100%), individual work (100%), and low time requirements (80%). Multiple characteristics suggest that projects have a local focus, including a restricted target group (100%), intrinsic incentives of participants (100%), and the absence of community-building measures (100%). Simultaneously, they have a socially and policy-oriented incentive and outcome structure with impact-related incentives (100%) and access and acceptance motivators (100%) while producing decisions (100%), and additionally sharing of things (60%) as either public (80%) or participant-restricted outcomes (20%). Although projects typically do not promote community building nor provide moderation (80%), they have medium to distinct communication structures with multi-sided communication (80%), expert (80%), and crowd (80%) feedback. They are initiated and driven by experts (100%).

**Experimental Archetype 6 "Hybrid, Self-led Communities":** This cluster contains two projects, which need to be considered outliers in the dataset (Figure 3.4, orange graph). What distinct them from other projects are their analogous components and their initiation and control by the entire participant group. This occurs with other participant-centric characteristics such as particular distinct communication structures, team-based participation, and the sharing of endeavors as a primary project outcome. We thus describe these outliers as hybrid, self-led communities.

**Experimental Archetype 7 "Projects of Individual Activism":** This cluster contains three outlier projects (Figure 3.4, purple graph). What sets them apart from others is the importance of individual participants as initiators and drivers of the project and the exclusive focus on producing legitimized decisions. Besides this, in their structure, they resemble archetype 4, including low-threshold, resource-based participation without a strong community focus. We thus characterize them as projects of individual activism.

**Experimental Archetype 8 "DIY Research Projects":** This cluster groups two outlier projects that are situated between archetypes 2 and 3 with a strong research focus and community orientation, i.e., distinct communication structures and outcomes that are available for the participants (Figure 3.4, brown graph). Thresholds to participation are higher than in cluster 2, indicating more personal responsibility for participants and, thus, a do-it-yourself research character.

### 3.4.4. Connecting the DIP Web Application and Archetypes

To communicate results to DIP practitioners, we connect the DIP web application and the DIP archetypes. While the web application improves the accessibility and usability of the initial taxonomy, the archetypes extend its informative value. By connecting both developments, we aim to bundle these potentials, making the DIP archetypes more accessible while increasing the attractiveness of the web application. Based on the web application's results page, we thus design a prototypical extended results page, which enables the user to be informed about the DIP archetypes. In addition to viewing their own classification, users receive short summaries on the different archetypes and can view the archetypes' characteristics next to their classification in the taxonomy's display (see Figure 3.3, right).

## 3.5. Discussion and Conclusion

In this chapter, we presented our research on making a theoretical taxonomy for DIP available to domain practitioners using the iterative process of DSR. Reflecting on the initial taxonomy, the necessity to further address practitioners' requirements regarding application and usability became apparent. Our instantiation of the taxonomy as an interactive web application yielded promising results in improving usability, empowering practitioners to independently use the taxonomy, with most subdimensions proving accessible. Furthermore, the positive feedback on the evaluated quality criteria underlined the value of the DIP taxonomy as a theoretical construct. As a challenge, we remark that the complexity of some participatory projects (i.e., multi-staged projects or multi-faceted activities) is not inherently considered in the taxonomy. For a more precise classification, those projects currently would have to be considered individually in their phases. Furthermore, while intentions to recommend the web application were, on average, positive, the evaluation of the classical net promoter score yields negative results, indicating that the web application does not operate as an individual service or product. If this is desired, further enhancements might be necessary, with the extended results page featuring the DIP archetypes or potentially an automatic self-assignment to an archetype being one promising avenue.

With our illustrative scenario analysis of 46 DIP projects, we advanced the application of the taxonomy, uncovering both the rich diversity of design variations in certain domains as well as underrepresented characteristics in others. Interestingly, especially the dimensions format and moderation, added based on the practitioners' feedback in the focus groups (Stein et al., 2023b) showed little variation in the dataset. This could indicate the practitioners' interest in these topics due to difficulties in implementation in practice, such as a lack of practical solutions. Utilizing cluster analysis, we identified significant design variations across commercial (Archetype 1), hybrid (Archetype 6), and digital, non-commercial involvement projects (Remaining Archetypes), as the majority of projects in our sample. Further exploration of this predominant category yielded the identification of a total of eight design clusters that we described as preliminary – partially experimental – archetypes of DIP projects. Although especially experimental archetypes can't be considered robust due to their small sample size, the archetypes show promising alignments with prior research in DIP. This includes e.g., distinctions between resource- and intelligence-oriented participation, patterns of a joint increase of communication structures and participation degree, or distinctions between organization- and citizen-centric design (Haklay, 2013a; Van Dijk, 2012). Encompassing classical economic crowdsourcing, research-oriented, and policy-driven initiatives alongside emerging mixed archetypes that transcend traditional boundaries, the illustrative scenario analysis underlines the suitability of the taxonomy to guide emerging developments in the digital government sector.

With our research we provide multiple practical and theoretical contributions. For practitioners and policymakers, the DIP taxonomy already captures design knowledge that can provide guidance within the DIP domain (Stein et al., 2023b), yet its instantiation as a web application makes this knowledge more easily accessible to them. The web application supports practitioners in the development of new initiatives by guiding them through key design considerations and providing inspiration for their implementation based on various participatory practices. In a similar way, existing projects can be evaluated by raising awareness of one's own design decisions and highlighting alternatives. Beyond this guidance, the web application facilitates evidence-based research on digital participation by collecting further DIP data and empowering practitioners to participate in capturing design descriptions of their projects. The DIP archetypes provide a starting point for practitioners, making common design choices tangible and enabling comparison while breaking down the diversity of DIP approaches. As a theoretical contribution, our research activities strengthen the evaluative rigor of the initial DIP taxonomy. In addition, while the web application instantiation of the taxonomy can serve as an example to other practitioner-oriented taxonomy developments, the archetypes as operationalized models, grounded in empirical data, offer a valuable counterpart to theoretical typologies, enabling a deeper understanding of current DIP practices and a point of reference for their evolution over time.

As a limiting factor of our research, for subcycle (a), we note that in favor of brief study duration, little information was collected on the test user group, preventing us from insights into user-specific experiences. Instead, we relied on the sufficient diversity of project initiators on mit:forschen!, which was documented, e.g., in terms of institutional or thematic backgrounds (Moczek et al., 2021). Regarding the illustrative scenario analysis, our approach to data collection ensured standardized coding but was limited to the interpretation of publicly available data and may, therefore, contain biases or incomplete information. In particular, dimensions such as incentives for participants and initiators can externally only be evaluated to a limited extent. In addition, our project sampling method cannot strive for representability in the field of DIP projects. Thus, rather than aiming at uncovering all design types within the DIP landscape and relying exclusively on external descriptions, we contribute preliminary archetypes as a starting point and enable continuous collection of further (internal) design knowledge via the DIP web application.

In conclusion, our research aims to provoke more engagement with the topic of project design for digital involvement not only by researchers but also by practitioners. As the DIP field is developing rapidly, creating rigid models and overviews is not feasible. Instead, structures should be created that promote the continuous exchange between research and practice. With the DIP taxonomy, we want to encourage research areas such as digital government, citizen science, and crowdsourcing to share experiences, rather than to separate themselves from one another because emerging practices in one discipline might be already established in others.

# 4. A Citizen Science Application of the DIP Taxonomy

## 4.1. Introduction

Citizen science, defined as the active involvement of the public in academic research, has gained significant prominence in the research landscape in recent years. This growth is evidenced by the rising number of scientific publications on the topic (Bautista-Puig et al., 2019; Kullenberg and Kasperowski, 2016), the expansion of citizen science projects, and a heightened level of institutionalization at both national and international levels (Vohland et al., 2021a). Alongside this quantitative increase, there has been notable diversification in both the application contexts and methodologies of citizen science. While traditional and effective crowdsourcing approaches in biology persist (Pettibone et al., 2017; Tauginienė et al., 2020), new projects across diverse scientific disciplines have emerged (Tauginienė et al., 2020; Wiggins and Wilbanks, 2019), employing a variety of participatory methods (Schleicher and Schmidt, 2020). A central theme in citizen science research has been the extent of participant involvement, with frameworks developed by scholars such as Haklay (2013a) and Shirk et al. (2012) delineating various levels of participation. These classifications range from projects where participants primarily engage in data collection to those involving participant-driven data analysis and further to initiatives enabling citizens to conduct independent research. Although the mapping of the citizen science landscape has advanced alongside the field's expansion, gaps remain in understanding the key design decisions that form how citizen science projects operate in practice. Beyond the participation level, several of these decision areas, which we call design dimensions (cf. Nickerson et al. (2013)), have received limited scholarly attention. This includes, for example, communication strategies within projects, incentive structures designed to promote participation, or feedback mechanisms, which remain underexplored both theoretically and empirically. While individual studies have provided experiential insights and strategies for implementing these aspects, namely design characteristics (cf. Nickerson et al. (2013)), systematic investigations into their prevalence and structure across projects are lacking.

However, investigations into these design dimensions hold particular relevance given the ethical commitments intrinsic to citizen science. Questions arise regarding the extent to which projects adhere to established principles of good practice within citizen science, as outlined in documents such as the European Citizen Science Association's (ECSA) Ten Principles or the Austrian Quality Criteria for Citizen Science (European Citizen Science Association, 2015; Heigl et al., 2020). For instance, to what extent do citizen science projects fulfill the ethical requirement of providing feedback to participants? Furthermore, given the potential of citizen science to democratize knowledge production (Irwin and Horlick-

Jones, Tom, 1995), it is important to examine whether current practices align with this democratic ideal. This raises the question of communication directionality within projects (Hecker and Taddicken, 2022): do most citizen science initiatives support bidirectional communication to facilitate the democratization of knowledge production? Additionally, understanding the motivations for adopting participatory approaches and the specific incentives designed to foster engagement offers valuable insights into both the ethical and practical foundations of citizen science initiatives. Beyond ethical considerations, a detailed analysis of the various design choices in citizen science is critical for enhancing the integration and efficacy of this research approach. Although an increasing number of training resources and guidelines are available (García et al., 2021), citizen science project designs are often developed or modified iteratively throughout the project lifecycle. Given that citizen science spans a wide range of disciplines, projects often address diverse research questions and employ varied research methodologies (Moczek et al., 2021; Monzón Alvarado et al., 2020). However, it remains unclear whether this thematic and methodological diversity impacts fundamental design decisions. Therefore, a systematic delineation of design dimensions and characteristics could be valuable not only for evaluating existing practices but also as a practical planning resource for researchers initiating new projects.

In response to these gaps, the present study applies a taxonomy for participatory processes developed by Stein et al. (2023b) to a sample of 60 German citizen science projects. Originally developed for participatory projects in fields such as politics and business, this taxonomy includes design elements applicable to citizen science, including some that are not traditionally associated with the field. This study thus aims to address the following research questions in the context of evaluating German citizen science projects through the lens of this taxonomy:

- Research Question 1: *What design characteristics do citizen science projects exhibit in an overarching taxonomy for participatory concepts and how are their preferences distributed?*

- Research Question 2: *Given the diversity of citizen science practices, to what extent can distinct design clusters be identified?*

- Research Question 3: *Regarding the multidisciplinary nature of citizen science, is there an association between a project's disciplinary focus and its assignment with specific design clusters?*

## 4.2. Theoretical Background: A Taxonomy for DIP

In science, politics, and economics alike, the inclusion of citizens in formerly exclusive processes and structures is gaining increasing relevance. For citizens, various participation opportunities exist, particularly in the digital realm, running parallel to one another. While the academic field discusses this involvement as citizen science (Haklay et al., 2021), research on political participation examines the phenomenon of e-participation (Macintosh, 2004; Sanford and Rose, 2007) and in the private sector context, crowd-X processes are a present practice (Estellés-Arolas et al., 2015; Howe, 2006b). Despite differing research fields and definitions, (digital) participation projects often share similar challenges regarding their design and implementation. In order to better explore and classify participation projects across domains and thus enable cross-domain learning, Stein et al. (2023b) aggregate these projects under the general term DIP and develop a taxonomy of their key characteristics. They define DIP as "projects that utilize digital platforms for the involvement of multiple external individuals in a defined participation process" (p.5) and derive 19 dimensions for describing DIP through an iterative process of theory and practice. The dimensions encompass three classifications determining the level of participation within a project, five dimensions delineating the project's implementation process, two dimensions concerning incentives for both participants and organizers, four dimensions depicting communication structures within the projects, three dimensions addressing project stakeholders, and two dimensions focusing on the projects' outcomes (Stein et al., 2023b). Although the taxonomy aims to analyze the design of participation projects independent of domains, potential advantages may emerge when employing the taxonomy within specific domains.

In the realm of citizen science research, numerous endeavors have been made to theoretically distinguish between various initiatives of citizen science practices. Early frameworks such as those proposed by Shirk et al. (2012) and Haklay (2013a) employs the level of participation in projects as a basis for differentiation. Shirk et al. (2012) conceptualize five project models ranging from contractual and contributory to co-created and collegial projects, while the typology of Haklay (2013a) encompasses four types of citizen science projects, spanning from crowdsourcing and distributed intelligence to participatory and extreme citizen science, reflecting higher levels of participation. Although these frameworks primarily focus on the degree of participation, they implicitly address various associated design choices, such as the project initiator (e.g., Shirk et al. (2012)) or the type of activity conducted within the project (e.g., Haklay (2013a)). More recently, scholarly attention in the field of citizen science has shifted away from rigid definitions, adopting a more open and inclusive approach. The ECSA has outlined ten principles of good citizen science practices, offering general guidance to practitioners regardless of specific project types (European Citizen Science Association, 2015). Haklay et al. (2021) emphasize the instrumental nature of citizen science definitions, advocating for the recognition of the plurality within citizen science endeavors. Consequently, they refrain from prescribing a singular, concrete definition, opting to promote open, fact-based discussions surrounding citizen science projects.

While considerable effort has been invested in crafting theoretical definitions of citizen science and offering recommendations for its implementation, relatively little attention has been devoted to making the actual design decisions of citizen science projects tangible. Spasiano et al. (2021) endeavor to develop a transdisciplinary framework encompassing citizen science practices, methods, and associated issues. However, their analysis primarily draws upon review articles from academic journals. Similarly, Stein et al. (2023c) conducted a design review of 16 multi-project citizen science platforms, yet their emphasis remains on the platform features rather than investigating their utilization in initiatives. Conversely, research analyzing real-life projects and their design choices has been undertaken by scholars such as

Monzón Alvarado et al. (2020) and Moczek et al. (2021). Monzón Alvarado et al. (2020) examined 36 Mexican citizen science projects across seven dimensions, encompassing geographical, temporal, and thematic scope, as well as sources of income, participant types, and models of participation tasks. Their research makes the distribution of different levels of participation in real-life projects tangible, while also linking these design choices to the geographic scope of initiatives. However, several other design choices remain unexplored due to limitations imposed by the availability of public information on the projects. Similarly, Moczek et al. (2021) analyzed 79 German citizen science initiatives concerning their funding, project roles and staffing, levels of participation, task organization, diversity, and inclusiveness. By engaging directly with projects on the “mit:forschen!” (formaly Bürger Schaffen Wissen) platform, Moczek et al. (2021) leveraged insights from project initiators regarding their projects’ design and volunteers’ involvement. While this has already provided many insights into the organization of citizen science projects, aspects of their exact implementation and how these relate to the classically considered levels of participation remain unexplored. Furthermore, the approaches lack the possibility to classify citizen science projects in the field of participatory practices. However, recognizing parallels and differences to other forms of participation could be helpful for practitioners to supplement their own design practice or to distance themselves as a discipline from certain characteristics consciously.

### 4.3. Methodological Approach

To address our research questions, we made use of the DIP taxonomy (Stein et al., 2023b) to classify citizen science projects in Germany. Taxonomies present a suitable approach to structurally describe a target domain by focusing on similarities and differences (Nickerson et al., 2013). While, to the best of our knowledge, no design taxonomy for citizen science projects is currently available, we refrained from developing such one, adopting a broader perspective by using the DIP taxonomy. Acknowledging the increasing relevance and diversification of participation formats in political and economic practices, enabling comparability of citizen science practices to other participatory paradigms seems beneficial. Furthermore, citizen science itself is experiencing a diversification of approaches and methods across national and disciplinary boundaries (Haklay et al., 2021), favoring broader definitions in the community. However, while digital citizen science formats have been trending throughout recent years, there is still a large number of projects that operate analogously. In the DIP taxonomy, the dimensions “Format” and “Implementation” (Stein et al., 2023b) are specifically oriented toward the digital domain. To encompass the diversity of analog projects, we modify the DIP taxonomy by incorporating additional characteristics labeled “analog” and “none”.

To obtain a sufficiently large project sample, we collaborated with the citizen science meta-platform “mit:forschen!”, which has been an anchor to citizen research before (e.g., Moczek et al. (2021)). mit:forschen! is the central citizen science platform in Germany. The platform has been online since 2014 and lists German citizen science projects. For a project to be included on the platform, registration is necessary. As part of the registration process, the project is reviewed by two independent experts and checked to see if it meets the platform's criteria. The following criteria guide this evaluation:

- Adherence to citizen science principles: Projects must pursue a scientific research question or aim to establish scientific infrastructure.
- Active citizen involvement: Citizens should be actively engaged in significant phases of the research process, such as developing research questions, collecting data, analyzing results, or disseminating findings.
- Transparency: The roles and contributions of all participants must be clearly and openly communicated.
- Accessibility and support: Projects must have a dedicated German-language website and appoint a contact person to address inquiries and assist participants.
- Regional and linguistic focus: Projects should be conducted in German and have a clear regional focus within Germany.
- Non-commercial orientation: Projects primarily serving commercial purposes are excluded.

The platform started with 14 projects; at the time of the survey, 240 projects were registered on the platform. Drawing on the experiences of Monzón Alvarado et al. (2020) and Moczek et al. (2021), we opt for a direct engagement with project initiators to accurately characterize projects in line with the taxonomy's dimensions. Accordingly, we conducted a survey to incorporate their internal perspectives on their projects. However, the application of the theoretical construct through multiple, diverse practitioners can lead to subjective and ambiguous classifications (Stein et al., 2023b). Therefore, we provided a web application, featuring detailed descriptions of the taxonomy's characteristics and facilitating the application process (Stein et al., 2025b). Between June and August 2023, we contacted all active citizen science projects ($n = 240$) on mit:forschen!, receiving $n = 66$ project classifications through the web application. After removing duplicates, we identified 60 unique project descriptions suitable for further analysis. We additionally enriched project data by manually categorizing the projects into the five research domains "Humanities, social and cultural sciences", "Engineering and planning sciences", "Mathematics and computer science", "Medicine and health sciences", "Natural sciences" and a remaining category of "Others".

## 4.4. Results

In the following, we will first report on descriptive results by dimension before giving insights into the clustering results and associated differences between research domains. An overview of the descriptive results is presented in Figure 4.1.

| Dimension | Sub-Dimension | Characteristics | | | | |
|---|---|---|---|---|---|---|
| Degree of Participation | D1 Extent of participation | Information sharing (0.40) | Consultative (0.45) | Democratic (0.28) | | |
| | D2 Participation offer | Single task (0.15) | Multiple tasks optional (0.63) | Multiple tasks mandatory (0.22) | | |
| | D3 Type of participation | Active-effort (0.87) | Active-ressources (0.17) | Passive (0.02) | | |
| Implementation of Participation | D4 Format | Digital (0.48) | Analog/ digital (parallel) (0.33) | Analog/ digital (sequential) (0.17) | Analog (0.13) | |
| | D5 Implementation | Asynchronous-web-based platform (0.62) | Asynchronous-mobile application (0.27) | Synchronous (0.30) | None (0.10) | |
| | D6 Structure of participation | Team work/ participation (0.37) | Individual work/ participation (0.77) | | | |
| | D7 Time requirements | High (0.22) | Low (0.17) | Self-selected (0.72) | | |
| | D8 Prerequesites | Domain knowledge (0.18) | Domain-specific equipment (0.08) | Assumes preconditions (0.80) | | |
| Incentives | D9 Incentives for participation | Self-related extrinsic (0.20) | Self-related intrinsic (0.60) | Impact-related (0.87) | | |
| | D10 Reasons for the participatory design | Acceptance & legitimation (0.42) | Funding (0.08) | Access and ressources (0.77) | Value-based (0.33) | Profit maximisation (0.02) |
| Communication | D11 Direction of communication | One-sided (0.03) | Two-sided (0.43) | Multi-sided (0.60) | | |
| | D12 Suggestions feedbacked | Expert feedback (0.70) | Crowd feedback (0.43) | No feedback (0.13) | | |
| | D13 Community building | Yes (0.57) | No (0.38) | | | |
| | D14 Moderation | Crowd (0.12) | Individual from crowd (0.15) | Organization/ expert (intern) (0.65) | Organization/ expert (extern) (0.03) | None (0.25) |
| Project Stakeholder | D15 Project driver | Crowd (0) | Individual from crowd (0.10) | Organziation/ expert (0.90) | Equal partnership (0.08) | |
| | D16 Project owner | Crowd (0.02) | Individual from crowd (0.08) | Organization/ expert (0.95) | | |
| | D17 Target group | Open (0.83) | Restricted (0.15) | Closed (0.03) | | |
| Gains and Outcomes | D18 Project outcome | Product (0.20) | Knowledge (0.93) | Decision (0.08) | Sharing things (0.60) | |
| | D19 Publicity of the outcome | Public (0.98) | Accessible for the participants (0.10) | Non-public (0) | | |

Figure 4.1.: Overview of the taxonomy's dimensions; In parentheses, the sample's distribution onto its characteristics.

### 4.4.1. Descriptive Analysis

**Degree of Participation:** In assessing the degree of participation within the projects, three sub-dimensions were evaluated. Firstly, project initiators could indicate the extent of participation across three levels (Somech, 2002): "Information Sharing" (where the crowd provides specified input but lacks decisive power), "Consultative" (where the crowd freely generates input without decisive power), and "Democratic" (where input generation and decision-making are jointly undertaken). Additionally, initiators could specify whether volunteers engage in one or multiple (mandatory) tasks within the participation offer category and whether participation was passive or involved active effort or resource contribution. Overall the vast majority of projects indicated an informative and/ or consultative character, with the optional offer of multiple participatory and effort-based tasks. However, except for passive tasks, which were only indicated by one project in the sample, all characteristics were selected by at least 10% of the projects, indicating a substantial variety across projects.

**Implementation of Participation**: In terms of implementing participation, the taxonomy assesses the participatory format, ranging from digital, mixed to analogous formats, their implementation with asynchronous or synchronous tools, and the work structure with individual and team-based options. Furthermore, time requirements and participation prerequisites with domain knowledge or equipment specifications are assessed. In evaluating this domain, we found a dominance of low-threshold but also tool-supported projects: The majority of projects indicated no necessary participation preconditions, while work is structured individually with self-selected time expenditures. Simultaneously, a minority of projects indicated a completely analogous conduction without any digital tool support. Following the first dimension, a variety of project design characteristics could be found, with all characteristics except equipment prerequisites being selected by more than 10% of the projects.

**Incentives**: Regarding incentives, both the incentives of the project initiators and the participants were assessed. Regarding the participants' incentives, all incentive types, from self-related (intrinsic/ extrinsic) to impact-related, were present, although extrinsic motivators remained in the minority. In terms of initiators' incentives, economic incentives such as funding or profit maximization were rarely present ($< 10\%$). Rather, the majority of projects indicated a need for access and resources.

**Communication**: To assess communicative structures within the project design, the taxonomy assesses whether the communication is one, two, or multi-sided, and whether and who gives feedback to participants' contributions and moderates discussions. Finally, project initiators can indicate whether community building is supported in the project. Overall, the evaluation of our sample shows that a majority of projects have dialogical structures in place, including multi-sided communication and community-building features, as well as expert-based feedback and moderation. While a substantial fraction of projects builds on crowd feedback, the crowd is typically little involved in moderation aspects ($\leq 15\%$). One-sided communication and external moderation are rarely present ($< 10\%$), yet some projects implement no feedback or moderation ($\geq 15\%$).

**Project Stakeholders**: In terms of project stakeholders, the taxonomy evaluates who initiated and drives the project, as well as who the intended target group is. For this domain, the projects in the sample showed particularly little variation. Most projects are initiated and driven by experts or an organization while being open to the general public. Some projects target a restricted participant group. However, all other characteristics, such as non-expert initiation, crowd-based drivers, or equal partnership constructs, are indicated by less than 10% of the projects.

**Gains and Outcomes**: Regarding project outcomes, the taxonomy assessed the type of project outcome as well as their publicity. As for project stakeholders, in this domain, there is particularly little variation across projects: Almost all projects indicate public knowledge outcomes, while the majority additionally specifies the sharing of certain things as outcomes.

### 4.4.2. Cluster Analysis

The cluster analysis yielded two distinct clusters with an average silhouette coefficient of 0.36. Further subdivision of projects resulted in lower silhouette coefficients, notably declining sharply at the points of two and five clusters. For the distinction of two clusters, a larger cluster of 42 projects (cluster 1) and a smaller cluster of 18 projects (cluster 2) were observed. In the following, the two clusters will be characterized based on their average choices across the dimensions of the taxonomy. A visualization is provided in Figure 4.3.

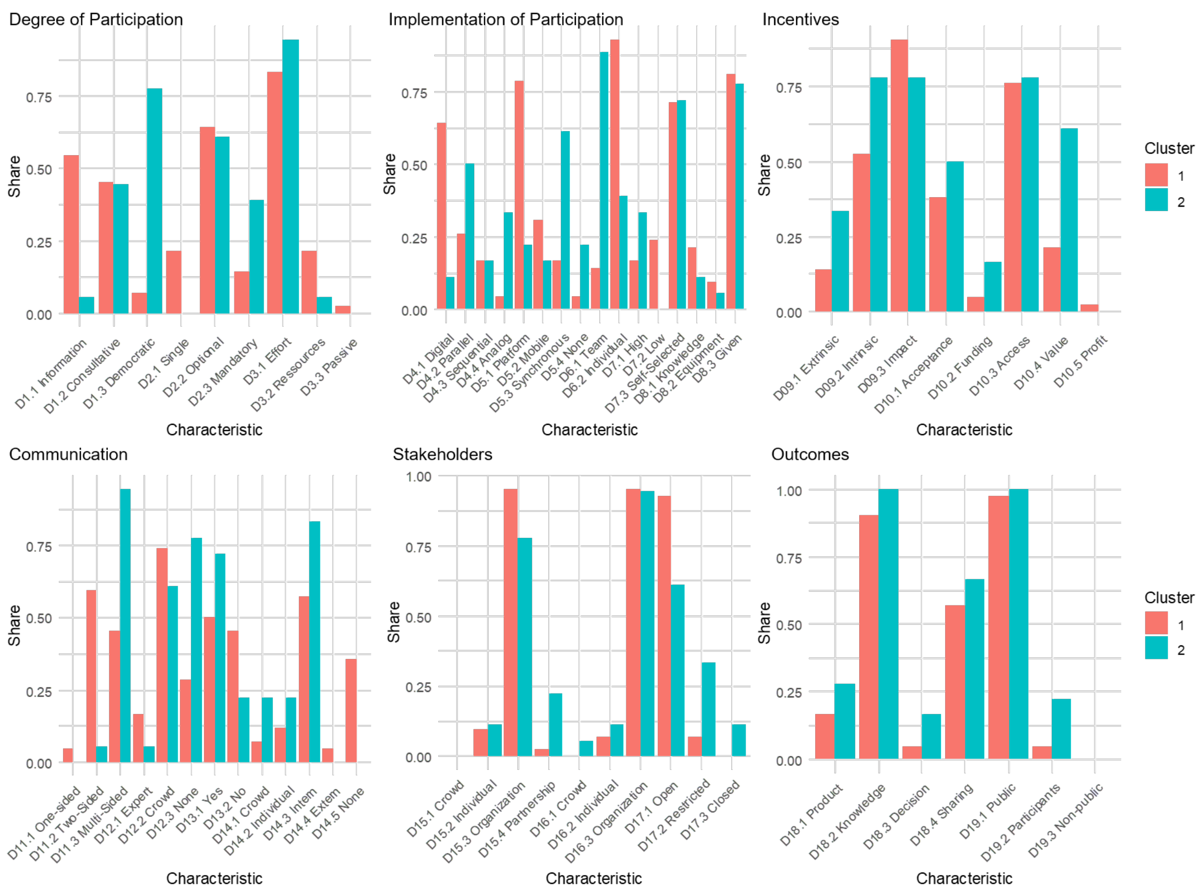


Figure 4.2.: Comparison of the design clusters regarding the share of their projects that employ the design characteristics.

**Cluster 1**: Projects in cluster 1 tend to exhibit a lower degree of participation, as shown by high values for informative and consultative projects, as well as projects offering only a single or passive task(s).

Their implementation is predominantly characterized by digital methods and individual work, with all projects indicating low time requirements falling into cluster 1. Compared to cluster 2, projects in cluster 1 emphasize impact-related volunteer incentivization, while reasons for their participatory design are often related to access/ resources and acceptance/ legitimation. Communication structures within cluster 1 are varied, but overall, there is a lower degree of communication and less focus on community engagement than in cluster 2. For instance, almost all projects with single/ two-sided communication and external/ no moderation are grouped in cluster 1, and expert feedback is more frequently mentioned than crowd feedback. Regarding stakeholders, projects in cluster 1 exhibit similarities, indicating an open but expert-centric character. Most projects are initiated and driven by experts, while the target group remains open. There is a mixed picture in terms of outcomes, although most outcomes are made accessible to the public.

**Cluster 2**: Projects in cluster 2 demonstrate a high degree of participation, with the majority of projects indicating a consultative or democratic character, inclusive of active, effort-based tasks. All projects within cluster 2 offer multiple tasks to participants. Regarding implementation, projects in cluster 2 are primarily characterized by collaborative work analogously or with both digital and analogous components. Many projects are conducted in digital and analog formats in parallel, utilizing team-based approaches facilitated by tools for synchronous work formats. Time requirements vary, ranging from high to self-selected. Incentive structures within cluster 2 projects are diverse; however, participants are more frequently incentivized by self-related motives compared to cluster 1. Additionally, funding and value-based participation are enumerated more frequently as reasons for participatory design. Communication within cluster 2 projects typically exhibits pronounced, community-oriented structures. Almost all projects facilitate multi-sided communication and ensure feedback and moderation by the crowd and experts. The majority of projects actively support community building. The stakeholder structure within cluster 2 projects presents a varied picture, with experts predominantly serving as initiators and drivers, but other constellations, such as equal partnerships, are also present. Similarly, the target group for cluster 2 projects displays characteristics ranging from open to restricted or closed. Finally, regarding gains and outcomes, projects in cluster 2 exhibit variability, but in comparison to cluster 1, more projects produce decisions as outcomes, and outcomes are more frequently restricted.

**Exemplary Projects:** To illustrate the nature of the two clusters, we highlight two exemplary projects from the sample. The ZOWIAC project[1] has been assigned to cluster 1 in the clustering process. The project investigates the distribution of invasive species in Germany, such as raccoons and American mink, and their impact on native ecosystems. Citizens contribute by reporting sightings through the ZOWIAC app or website. Verified data are displayed on distribution maps, supporting future conservation measures

[1] Information on the ZOWIAC project can be found here: https://www.mitforschen.org/projekt/zowiac-forschung-gebietsfremden-raubtierarten

and monitoring potential health risks related to parasite transmission. On the other hand, the "Digital Active Women" project[2] has been assigned to cluster 2. It aims to better align the digital information and advisory needs of women who recently migrated to Germany with the services offered by local authorities, counseling centers, and migrant organizations. Women with up to seven years of residence act as co-researchers, evaluating digital services and creating recommendations through workshops and an online survey. Project findings are shared with stakeholders to improve targeted digital outreach and enhance these women's social participation.

### 4.4.3. Project Domains

The samples' distribution onto thematic domains is visualized in Figure 4.3 on the left side and set into relation with the cluster affiliation in Figure 4.3 on the right side. Overall, a dominance of the domains "Humanities, social and cultural science" and "Natural sciences" is apparent, enriched by a few instances of medicine- and engineering-related projects in the sample. Investigating relations between project domain and cluster affiliation, we find that cluster 1 primarily consists of projects in natural sciences (0.48) and humanities, social, and cultural sciences (0.36), while cluster 2 is primarily composed of projects in humanities, social, and cultural sciences (0.50) and from medicine and health sciences (0.28). Applying a Fisher's- Exact Test, we find that the domain distribution and the cluster affiliation are dependent ($p < 0.01$).

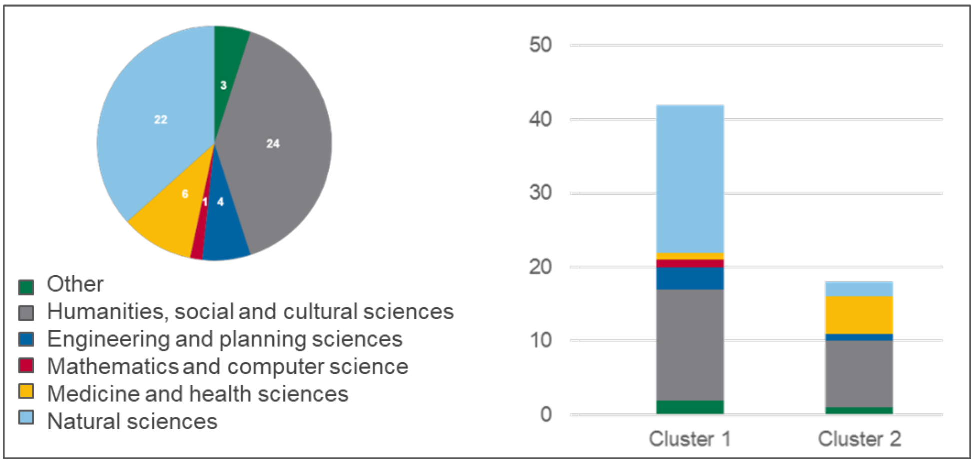


Figure 4.3.: Pie chart of the disciplinary focus of projects in the sample (left) and visualization of their proportion in the two design clusters (right).

[2]Information on the "Digital Active Women" project can be found here: https://www.mitforschen.org/projekt/digital-active-women-wie-beratungs-und-informationsangebote-besser-ankommen

## 4.5. Discussion and Conclusion

This article presents insights from a survey conducted with 60 German citizen science projects, applying a taxonomy-based characterization of their project design. Aiming to extend insights into how different citizen science projects are structured, implemented, or incentivized and to position their design in the broader landscape of participatory approaches, we utilized a comprehensive taxonomy for (digital) involvement projects covering 19 distinct design dimensions that apply to citizen science projects and beyond. In the following, we will discuss our survey results and their implications for answering our research questions and position the main contributions and limitations of our work.

Applying the DIP taxonomy to citizen science projects, we generally find a broad diversity of design characteristics adopted by the various projects. Yet, several characteristics seem absent or rarely present in the citizen science sample. Specifically, this encompasses 1a) passive types of citizen participation, 2a) profit maximization-related incentives for project initiators, 3a) one-sided communication styles, 4a) external moderation, 5a) settings of equal partnership between citizens and researchers, 6a) crowd-initiated and 7a) crowd-moderated projects, 8a) closed target groups, 9a) decision outcomes and finally 10a) non-public outcomes. Similarly, some characteristics are employed by almost all projects in the sample, namely 1b) Organization/expert-initiated and 2b) driven projects, 3b) knowledge outcomes, and 4b) public outcomes. The under-/ overrepresentation of these characteristics could either hint towards inherent traits of the citizen science landscape (compared to other participatory approaches) or inherent traits of the citizen science meta-platform mit:forschen!. For example, the demand for knowledge outcomes and their public provision follows directly out of the ten ECSA principles, advocating for a core characteristic of citizen science projects. Similarly, the demand for feedback and the active involvement of citizens, specified in the principles, could explain the general absence of one-sided communication. Interesting in this regard is that a substantial number of projects indicate the absence of any feedback in their projects, challenging principle number five. With much research being conducted by non-profit institutions such as universities, the absence of profit maximization is likely also to be inherent to citizen science projects. The absence of closed target groups, on the other hand, could be a result of sampling on the platform mit:foschen!, as citizen science projects with closed target groups might benefit less from the listing on a citizen science meta-platform. For other absent characteristics, we could have expected broader adoption based on the citizen science literature, such as settings of equal partnership or the activation of the crowd as initiators or moderators, as they are specifically described in citizen science typologies (e.g., Shirk et al. (2012)). The initiation and drive through experts/ organizations must be presumed inherent to citizen science projects, although in this case, reality does not seem to live up to the expected principles, which foresee citizens in various roles such as project drivers.

Despite the under-/ overrepresentation of certain characteristics in the citizen science sample, there are multiple design dimensions where the sample proves a high heterogeneity of project design. We find that this heterogeneity cannot be perfectly explained by the existence of several distinct design clusters, but rather, a variety of design characteristics seem to be freely combinable. Yet, our results indicate the existence of two loose design clusters differentiating projects with lower degrees of participation from projects with higher degrees of participation. In that sense, our results support the differentiating role that the degree of participation has in the citizen science literature (Haklay, 2013a; Shirk et al., 2012). Yet the insights from the clustering solution enable the identification of potential connections between the degree of participation and further details of their design, thus extending traditional typologies. On the basis of existing typologies, we describe our two design clusters as "crowdsourcing research design" (cluster 1) and "participatory research design" (cluster 2). Projects in the crowdsourcing research cluster exhibit comparable characteristics to projects described as "Crowdsourcing" or "Volunteer Sensing" in the typology of Haklay (2013a). In contrast, projects in the participatory research cluster resemble the descriptions of "Participatory Science" or "Extreme Citizen Science". As in the theoretical typology, we find differences in the communication structures between the two clusters, but additionally, differences in the implementation of the participation or the task scope become apparent. These are interesting links to further explore in upcoming work to determine whether different design configurations might be more successful for different degrees of participation.

Finally, in our sample, we find a significant relation between the cluster affiliation and the project's topic domain, which might be an important insight in light of the multidisciplinary nature of citizen science. Coming from various research disciplines, each with its distinct methodological approaches, requirements for data, data quality, and much more, our results might indicate that different design configurations are more reasonable for certain disciplines than others. On the other hand, results could also indicate that there is too little exchange on the design of citizen science projects between different disciplines, and thus, much room left for mutual inspiration and learning. Certainly, either way, the finding needs to be further discussed in the research community to consider best practices within the given multidisciplinarity.

### 4.5.1. Theoretical and Practical Contributions

Overall, with the insights from our study, we can realize several theoretical and practical contributions. As descriptive work, we create an important basis for the research community in quantifying different design characteristics in citizen science projects. We thus enable a better characterization of citizen science projects in Germany and a differentiation from other participatory approaches. Due to the large number of dimensions considered, we provide new insights into possible correlations between different design approaches, especially with the frequently considered degree of participation, and at the same time, identify gaps for further research where distributions cannot yet be explained.

For practitioners, recording the status quo can be an important tool for reviewing the fulfillment of self-imposed goals and principles and, at the same time, making targeted adjustments or addressing identified opportunities or challenges. For new project initiators in particular, the recording of common practices can be a useful guide when planning the design of their own initiatives – especially since a one-size-fits-all approach does not seem to exist, but depending on the project and discipline, different design configurations are more common than others.

### 4.5.2. Limitations and Future Research

Certainly, there are several limitations of our work to be considered, some of which can show important directions for future research. First, there are several aspects to be discussed in the consideration of our data. Due to the limited sample size, as well as the focus of citizen science projects in Germany on the platform mit:forschen!, the representativeness of the results is limited, especially in the international citizen science landscape. Furthermore, although employing the taxonomy web application, we tried to ease the taxonomy usage; our approach cannot exclude subjective interpretations of the survey participants. In this sense, the decision against the enforcement of mutual exclusivity for multiple characteristics of one dimension also resulted in the taxonomy not being followed precisely as it was originally intended, and in some cases, statements were made that would have required further explanation. Finally, a limiting aspect of our descriptive approach is that we cannot make statements about the design configurations' quality. Since we have no insights into the reasons behind the choice of specific design options, nor do we know the success achieved by the citizen science projects, we can only reflect the common decisions made by project initiators. Alongside continuing our quantitative, taxonomy-based approach, such as expanding the sample to include additional citizen science projects from various sources, this limitation is an extremely exciting direction for future research. A qualitative study, for instance, could be conducted to understand the motivations behind the choice of different design options. Were they the most suitable ones, the most cost-effective, or the simplest ones? Were alternative options considered at all? By complementing the quantitative results with these insights, important conclusions could be drawn for citizen science advisory practices. Expanding the data to include a metric for project success would also be an interesting next step. However, we foresee significant challenges in representing the diverse goals of citizen science projects with appropriate metrics.

### 4.5.3. Conclusion

In conclusion, with this work, we aim to provoke a discussion among practitioners and the research community on the design of citizen science projects in a sense that is open to being inspired by the variety of different approaches and design options within and beyond the field of citizen science. While we believe that defining desirable community standards and principles is certainly important, it is likewise important to quantify the status quo to ask critical questions and identify potential challenges in practice.

# Part III.

# Positioning and Simplifying Data in Digital Involvement

**Abstract Part III**

*Part III, "Positioning and Simplifying Data in Digital Involvement", examines the critical role of data in digital involvement and the affordances for inclusive participation. As information forms the foundation of any participatory process, effectively communicating project-related data is essential in datafied societies. Thus, to inform the design of digital involvement platforms a nuanced understanding of the relevance and impact of data in the context of participation is required. This part begins by exploring the historical evolution and societal diffusion of data as a foundation for the complexities of contemporary data-driven societies. It further incorporates a survey study on citizens' perceptions of datafication and digital involvement tools to complement the theoretical perspective. Building upon these insights, this part investigates the effects of varying data representations on citizen's involvement. E-participation in urban planning serves as a compelling use case for this research, given its reliance on complex spatial and geographical data and the critical role of involving laypersons in designing sustainable and livable environments. It thus exemplifies scenarios where simplifying data representation is pivotal for fostering democratic processes in a datafied society. To address this, the following part explores simplification strategies through virtual reality (VR) and 3D visualizations in two consecutive research endeavors. The research on VR-based involvement was conducted in collaboration with Jonas Fegert and presented at the 31st European Conference on Information Systems (Stein and Fegert, 2024). Similarly, research on 3D visualizations was undertaken in partnership with Alicia Wittmer, Lukas Buß, and Jonas Fegert. The work was presented at the 58th Hawaii International Conference on System Sciences (Stein et al., 2025a) and nominated for a Best Paper Award. In this part, titles, tables, and figures from the original publications have been renamed, reformatted, and re-referenced to align with the structure of the dissertation. Additionally, chapter and section numbering were adjusted, and formatting, abbreviations, and references were standardized for consistency.*

# 5. The role of Data (Literacy) in DIP

## 5.1. Introduction

In the critical pursuit of shaping digital involvement in a datafied society, a thorough understanding of the roles that data plays in this society serves as a fundamental basis. This includes an understanding of its current relevance and significance in public discourse, from which the necessary requirements and competencies of those involved in the discourse can be derived. This chapter, therefore, lays the theoretical foundations for considering data (literacy) in digital involvement. By tracing the historical development of data and its societal purpose, one gains a deeper understanding of how datafication has evolved and what characterizes today's complex data-driven society. In doing so, this chapter arrives at a positioning of data literacy as a societal competence and examines what the term encompasses and how it is relevant to the general public.

Beyond the scientific-theoretical framework, the role of data and data literacy in society is also shaped by citizens' perceptions. Their perspectives are essential for designers of digital processes and tools to ensure that they do not fail to meet actual needs. According to established IS theories such as the "Unified Theory of Acceptance and Use of Technology" or the "Technology Acceptance Model", the perceived benefit of technology plays a decisive role in the users' intention to use it (Davis, 1989; Venkatesh et al., 2003). The experiences and attitudes of citizens with regard to the relevance of data and data-related skills in their lives are, therefore, a decisive basis for the successful design of tools. In the absence of meaningful studies on this topic, in this chapter, a short survey is conducted with a mixed sample of 512 citizens investigating their perception of data presence, competence relevance, and the requirements of participation tools. Combined, the theoretical findings and practical insights provide a basis for designing supporting DIP tools in Parts II and III of this dissertation.

## 5.2. Development of Data in Public Discourse

Data, as defined by the Cambridge Dictionary, constitutes "information, especially facts or numbers, collected to be examined and considered and used to help decision-making or information in an electronic form that can be stored and used by a computer" (McIntosh, 2013). This two-part definition illustrates the fundamental change that data and its role in society have undergone as a result of digitalization. Yet data has been an influencing factor in society for a long time. As early as the Upper Palaeolithic, people began recording data to memorize and count events by digging marks into walls and sticks (Alaimo and Kallinikos, 2024). It is also known from the Mesolithic that traders and government officials used clay markers as records of possessions and transactions (Alaimo and Kallinikos, 2024). In the epochs that followed, the power of data to govern emerging empires developed. The systematic census of the Roman Empire to facilitate tax collection, for example, is an early demonstration of data utilization to practice control and power on citizens (Hintz et al., 2018). It becomes apparent that from the beginning, data was used to establish a unified social truth, but only a minority had the power to shape it. Factors that distinguished this minority could be, for example, status and access to recording tools. During the years to follow, these differentiating factors adapted through the changing (im)material basis of data but also through the professionalization of data-related skills and various professions in the years to follow.

In the Renaissance, during the scientific revolution, data gained importance as records of external observations became the basis of science. Initially, this new form of science was promoted and practiced by amateurs and "gentleman scientists". Science and its data were thus part of the reality of life for curious citizens, such as the ecological research of farmers or hunters. It was only in the late 19th century that science professionalized, emphasizing the role of official scientists affiliated with universities and research centers and diminishing the influence of amateurs in the field. (Levy and Germonprez, 2017; Miller-Rushing et al., 2012) The Renaissance was also the time when the discipline of statistics blossomed: In the 17th century, John Graunt, a self-taught draper living in London, introduced fundamental statistical concepts such as statistical association or inference. With his publication, which was based on a systematic study of the London mortality lists, he replaced speculation with structural quantification and conclusions. The very linguistic rootedness between *state* and *statistics* indicates the fundamental role that data products have played in political decision-making since the Enlightenment. Data was given a new epistemic role by being accepted as something (externally) given and its measurement and social meaning became closely linked to expertise and institutions. They became part of the concept of society by using rates such as immigration or crime to characterize societies and thus contributed to a new formation of reality. At the same time, they gave institutions new legitimacy and autonomy, reinforcing the historic relationship between the control of data and societal influence. (Alaimo and Kallinikos, 2024; Connor, 2024; Radermacher, 2024)

The 19th and 20th centuries were characterized by industrialization and digitization. This, in turn, had an enormous impact on the collection and use of data in society. The technological advancements fueled an increase and professionalization of bureaucracy in political and economic institutions. Expanding the collection of data on one's citizens or workers, enabled new forms of monitoring and surveillance (Hintz et al., 2018). The introduction of Hollerith's punch card tabulating machine marked the beginning of machine data processing. It laid the foundation for a decoupling of data and its information output and also meant that data was no longer visible as such without the necessary equipment such as cards, hole punches, tabulators, and sorters being owned and operated. In the following century, these changes were extended by the theory of electronic data processing and the introduction of the computer, and computer science. The detachment of data from its original context and the transformation into machine-processable tokens, influenced by the language and materiality of machines, contributed to both the formalization and specialization of data practices. Advances in the scope and complexity of data collection and computations have led to an expansion of their role in economic and organizational processes. Digital data remain artifacts of knowledge, but recursive interactions between data and their cognitive and communicative functions enable novel dynamic developments. (Alaimo and Kallinikos, 2024)

Nowadays, in the 21st century, personal computers, the Internet, and digital communication are the norm, revolutionizing data and information generation, access, and sharing for every individual. Big data and artificial intelligence (AI) introduce never-seen opportunities for data-based insights for those who have the skills to properly use it. Data and information have never been easier to access, process, and share by anyone, yet they have never made it so easy for economic and political actors to accumulate power. In the 2010s, the global surveillance and espionage scandal, leaked by Edward Snowden, and the Cambridge Analytica scandal exemplified how the massive collection and analysis of data enable actors to exercise surveillance and economic or political influence in secret (Hintz et al., 2018). The non-transparent collection of data and the complexity of data analyses make it even more difficult for the layperson to understand common practices, which in turn fuels misinterpretation, misuse, and falsification of data in public discourse. The new possibilities and threats surrounding data in the 21st century became apparent in the outbreak of the Covid-19 pandemic in the year 2020: As the former Director-General of Eurostat and Chief Statistician of the European Union Walter Radermacher describes it, the pandemic "has been accompanied by an 'infodemic', a flood of data, some of varying quality, that confuses rather than informs the layperson" (Radermacher, 2024, para.1). Data Journalism, a field that has been continuously growing since 2008, remarked a new peak demand for data-led news stories, becoming vital to influence public opinion and government policies (Tong, 2024). Showcasing the power of digital connectivity and modern technology numerous participatory initiatives contributed to the collection and analysis of important health data (Tan et al., 2022). Simultaneously, the increasing skepticism regarding experts of all sorts fueled the negation or relativization of official facts and the trust in one's own calculations and interpretations for parts of the population (Radermacher, 2024). The flood of misinformation complicated fact-checking activities, even in journalism (Tong, 2024). Paired with a lack of data literacy in society (Bhargava et al., 2016; Debruyne et al., 2021), this resulted in a critical threat to the fight against the Covid-19 pandemic (Tan et al., 2022) and the basic democratic order.

## 5.3. Data Literacy as a Societal Competence

Following the disruptive technological changes of the 21st century and their impact on the public role of data, the increasingly demanding skills to understand and work with data become vital not only for experts and professionals. Scholars are concerned with defining data literacy as a competence framework and positioning data literacy as a societal competence. With their data literacy framework *Future Skills* Schüller et al. (2019) create a comprehensive guideline for the competencies associated with data literacy, highlighting its wide range from the establishment of a data culture to the ability of acting data-driven. Their fine-granular framework, which differentiates necessary knowledge, skills, and attitudes for different adoption levels of competence, is aimed at supporting the design of course curricula and learning objectives in educational institutions, and private and public organizations. Central to the framework is the arrangement of competencies around the data value creation process: Showing the abundance of competencies necessary to encapsulate information about a real-life system into a final data product, and reversely, decoding data products properly into the characteristics of the very systems they represent, showcases the mammoth task of equipping professionals, decision-makers and citizens with the relevant skills to exchange information among each other.

To highlight the centrality of these competencies for the general public, the data literacy framework by Yates et al. (2021) goes one step further, defining *Data citizenship*. Differentiating three main dimensions, namely data thinking, data doing, and data participation, they outline the intertwined nature of data literacy and active citizenship. Data thinking reflects on the necessity for citizens to view and analyze everyday situations from an angle of data, whereas data doing encompasses practical skills of managing and analyzing data. Finally, data participation focuses on the proactive engagement with and about data from an individual and network perspective (Yates et al., 2021). Critical to the data citizenship framework is also the extension of the data literacy term in light of the increasing threat of online dis-, mis-, and misinformation (Carmi et al., 2020). In this context, Soßdorf et al. (2024) propose a *Synergistic Literacy Model* for combating disinformation, linking data literacy to critical media literacy. They argue for the important role of using theoretical literacy models to conceptualize and evaluate digital tools, promoting learning and empowerment for citizens in a datafied society.

## 5.4. Citizen Perspectives on Data (Literacy)

While from a theoretical perspective, the relevance of data in the public discourse and the importance of data literacy as a societal competence is undisputed, the experiences and attitudes of citizens are central to any measures undertaken in these regards. Yates et al. (2021) operationalized their data citizenship framework with a UK-representative study, identifying six data citizenship personas. While inquiring attitudes and skills of data literacy in the public, their main focus resided on media usage and on citizen's awareness of the use of their personal data. For the German realm, comparable studies exist assessing media usage, and citizens value perceptions on their personal data (e.g., Behrens et al., 2014; EOS/ Kantar, 2020). However, attitudes regarding the general relevance of data and data literacy skills in their lifes, remains unexplored. The Federal Ministry of Education and Research (BMBF) recently funded a project on the first large-scale representative study assessing citizens' data literacy in Germany [1], yet results are only expected in 2025.

Thus, to gain a first impression of citizens' perspectives on the prevalence of data and their need for

[1] Project DaLi executed by the LIfBi, see *www.lifbi.de/DataLiteracy*

literacies to cope with this, in this chapter, a brief survey study is conducted focusing on the current post-pandemic situation in Germany. Drawing from the design of a representative study in Germany, investigating important skills for a digital future (Börsch-Supan, 2017) and the definition of data literacy (D'Ignazio and Bhargava, 2016) a questionnaire is developed, to inquiry about the presence of data, the necessity for data literacy skills both in private and public life and associated requirements for digital involvement infrastructure (see Table 5.1). The survey was part of a larger experiment on digital tools for participation in urban planning, which is reported in Chapter 6 of this dissertation. Utilizing the panel provider Prolific (Palan and Schitter, 2018), between December 2023 and January 2024, the survey was distributed to participants located in Germany. Based on the fulfillment of an attention check, 512 survey questionnaires were further processed in the analysis. The sample constituted 241 female and 271 male participants between 18 and 71 years old ($Mean = 29.55$). Not all participants originally came from Germany, as indicated by 27.15% participants with a foreign mother tongue. In terms of education, 140 participants had a degree higher than a bachelor/ pre-diploma, 155 had a bachelor/ pre-diploma, 114 had "Abitur" as the highest German high school diploma, 61 had a lower school diploma, and 42 participants indicated other or no details regarding their education. The results of the survey are reported in the following and visualized in Figure 5.1

**Presence of data:** Participants reported considerable presence of data both in the public discourse as well as in their private lives. On average, the presence of data in the public discourse was rated with 5.05 points on a 7-point Likert scale ($SD = 1.28$), whereas participants indicated an average value of 4.81 ($SD = 1.64$) for the presence of data in their private life. Differences between the two are significant, yet negligible in effect size ($p \leq 0.001$).

**Necessity of data literacy skills:** Regarding participation in the public discourse, participants agreed on a high necessity for all data literacy skills ($Mean \geq 5$). On average, the skill of being able to argue with data was valued highest ($Mean = 5.90$, $SD = 1.00$), followed by the ability to read data ($Mean = 5.90$, $SD = 0.91$), the ability to analyze data ($Mean = 5.70$, $SD = 1.09$) and finally, working with data ($Mean = 5.50$, $SD = 1.15$). Regarding the importance for private success, participants also agreed on a high skill importance ($Mean \geq 5$), although reported values were slightly lower. On average participants ranked the ability to read data as most important with 5.68 points on the 7-point Likert scale ($SD = 1.10$) followed by the ability to argue with data ($Mean = 5.51$, $SD = 1.24$), the ability to analyze data ($Mean = 5.42$, $SD = 1.32$) and finally, the ability to work with data ($Mean = 5.27$, $SD = 1.31$). Effect sizes of differences between the individual skill importance in public and private life, as well as within the skill importance in one category, are small to negligible.

**Requirements for involvement tools:** For the design of involvement tools, participants on average, emphasized the need to support understanding of relevant participation data most ($Mean = 6.03$, $SD = 0.88$), closely followed by the need to provide relevant data ($Mean = 5.94$, $SD = 0.94$). Supporting one's skill improvement and autonomy in handling data was supported less, but yet received much approval by citizens ($Mean = 5.32$, $SD = 1.29$; $Mean = 5.47, SD = 1.14$). While the effect sizes for differences between

| Topic | Items | | Sources |
|---|---|---|---|
| Presence of data in public and private life | I often encounter data and figures in the public discourse. | | Self-developed |
| | Dates and figures are not very present in my private life. | | |
| Skills needed for public and private life | What skills will be crucial for coping in the world of tomorrow? What do you need to be able to participate in public discourse? | Be able to read data, i.e., understand what data is and what aspects of the world it represents. | Adapted from Börsch-Supan (2017) and D'Ignazio and Bhargava (2016) |
| | | Be able to work with data, i.e., create, record, cleanse and manage data. | |
| | | Be able to analyze data, i.e., filtering, sorting, aggregating, comparing and performing other analytical operations. | |
| | | Being able to argue with data, i.e., using data to support a larger narrative designed to convey a specific message to a specific audience. | |
| | For each skill, please decide to what extent you believe it will be critical to your personal success and well-being. | Be able to read data, i.e., (...). | |
| | | Be able to work with data, i.e., (...). | |
| | | Be able to Analyze data, i.e., (...). | |
| | | Being able to argue with data, i.e., (...). | |
| Requirements for involvement systems | Involvement systems should provide me with all the data and information that is relevant to the participation process. | | Self-developed |
| | Involvement systems should help me to understand the data and information relevant to the participation process. | | |
| | Involvement systems should help me to improve my understanding and handling of data and information. | | |
| | Involvement systems should enable me to handle data and information independently. | | |

Table 5.1.: Survey questionnaire.

the indicated needs were small to negligible for most pairs, a significant difference with a moderate effect size could be found for the comparison of supporting understanding versus skill improvement ($d = 0.544$, $p \leq 0.001$).

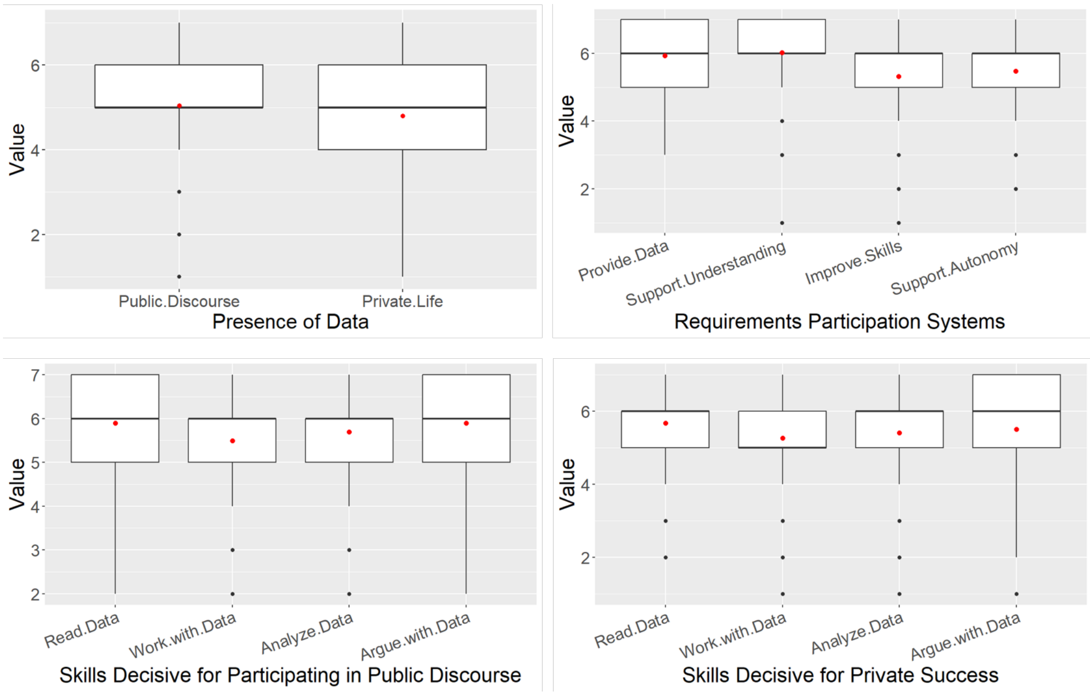


Figure 5.1.: Boxplots of the survey results with red points indicating sample mean.

**Demographic Factors:** To assess potential differences in opinions within the citizen sample, we evaluate the influence of age, gender, educational level, as well as differences for people with foreign mother tongues using a joint regression. The results are visualized in Table 5.2. Overall, we find that the demographic factors explain little of the variability in the survey data ($Ad\,justedR^2 \leq 0.046$). For the factor of sex, we find that male participants reported significantly higher presence of data in their lives ($\beta = 0.36$, $p \leq$0.001), stressing both the necessity of data literacy skills for public participation ($\beta = 0.18$, $p \leq$0.05) and even more so for private success ($\beta = 0.35$, $p \leq$0.001) than female survey participants. Furthermore, participants that had a different mother tongue than German reported a significantly lower presence of data in their lives than native German speakers ($\beta = 0.18$, $p \leq$0.05). Finally, the factor of education significantly influences the reported presence of data, as well as the reported requirements for involvement systems. Compared to participants with Abitur as their highest educational qualification, participants with degrees above Bachelor level reported significantly higher presence of data in their lives. Simultaneously, participants with an educational degree lower than Abitur reported significantly higher values for the requirements of involvement systems compared to participants with Abitur ($\beta = 0.18$, $p \leq$0.05). Finally, the residual category of other educational qualifications yielded significant differences to Abitur-participants, which, however, can not be interpreted further.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Data Presence | Skills Public | Skills Private | Involvement Systems |
| | (1) | (2) | (3) | (4) |
| Sex: Male | 0.360*** | 0.180* | 0.351*** | −0.113 |
| | (0.107) | (0.075) | (0.092) | (0.068) |
| Age | −0.010 | −0.008 | −0.004 | 0.008 |
| | (0.007) | (0.005) | (0.006) | (0.004) |
| First Language: Non-German | −0.259* | 0.033 | 0.127 | −0.146 |
| | (0.125) | (0.087) | (0.107) | (0.079) |
| Education: Bachelor/ Vordiplom | 0.114 | 0.122 | 0.165 | 0.088 |
| | (0.152) | (0.106) | (0.130) | (0.096) |
| Education: Degree higher than Bachelor | 0.387* | 0.163 | 0.126 | 0.165 |
| | (0.160) | (0.112) | (0.138) | (0.102) |
| Education: Degree lower than Abitur | −0.131 | −0.108 | −0.288 | 0.253* |
| | (0.194) | (0.136) | (0.166) | (0.123) |
| Education: Other | −0.072 | 0.211 | 0.450* | 0.280* |
| | (0.221) | (0.155) | (0.190) | (0.140) |
| Constant | 4.981*** | 5.794*** | 5.283*** | 5.417*** |
| | (0.222) | (0.156) | (0.191) | (0.141) |
| Observations | 512 | 512 | 512 | 512 |
| $R^2$ | 0.047 | 0.028 | 0.059 | 0.035 |
| Adjusted $R^2$ | 0.034 | 0.014 | 0.046 | 0.022 |
| *Note:* | | | | *p<0.05; **p<0.01; ***p<0.001 |

Table 5.2.: Regression results for averaged values of data presence, skills needed for public participation, skills needed for private success, and requirements for involvement systems.

## 5.5. Discussion and Conclusion

In this chapter, a theoretical approach to trace the development of data in public discourse throughout history and to position data literacy as a multifaceted set of competencies essential for all members of society was taken. A survey was used to contrast these insights with perspectives from citizens on the role of data and data competencies in their public and private lives.

By assessing the roles and characteristics of data over time, it becomes evident that data has evolved significantly in terms of its figuration, methods of processing, and professionalization as a field. However, from its inception to the present day, data has enabled the objectification of events, creating a unified "truth" that legitimizes decisions. The historical evolution of data offers insight into the formation and authority of political, scientific, and private institutions, demonstrating how actors can use the collection and analysis of data as a means to exert control and influence. This role alone motivates the democratization of data access, understanding, and usage to counteract potential misuse and prevent unilateral advantage. While advancements in the 21st century have partially expanded data accessibility, the sheer volume, recursive interactions, and novel analytic capabilities have introduced new challenges to its democratic use, even leading some citizens to reject information altogether. Numerous data literacy frameworks showcase the complexity and diversity of skills to properly handle today's data-based tasks. Consequently, simplifications and abstractions are urgently needed to facilitate inclusive participation in complex decisions. Simultaneously, mastering data-related skills is crucial for autonomous public and private engagement in a datafied society — a notion strongly supported by citizen perspectives.

Our study demonstrates that respondents overwhelmingly recognize the presence of data and the importance of relevant skills in both private and public life - interestingly with little differences between the two. This supports the notion that in the datafied society, the lines between public and private life blur. Few demographic differences are to be found, and additionally, limited variations in perceptions across different areas of competence, highlighting the importance of a broad spectrum of data-related skills. When evaluating requirements for involvement systems, there are moderate differences in the extent to which individuals believe involvement systems should support their data literacy. While a majority supports all extension states, on average the highest approval focuses on the requirement of enhancing understanding of relevant data.

Overall, this chapter offers an understanding of the societal role of data and the complexity of associated literacies as a foundation for further research in this dissertation. It further motivates the integration of data-related needs into the design of digital involvement systems. Yet, when using the insights, there are some limitations to acknowledge. First, the chapter focuses only on a selection of key milestones in the history of data, giving a summary rather than an exhaustive overview of its evolution. Second, the presented study is not intended to represent a representative sample of Germany. It offers initial perspectives from a diverse participant pool, pointing toward promising avenues for future research. For example, although there was great agreement among the respondents, some demographic differences warrant further investigation. The study provides limited information to interpret these fully, but differences based on gender, education, and language offer interesting starting points for future studies on group-specific data literacy needs and challenges. Future studies could delve into how data literacy requirements and preferences vary among these demographics, and investigate whether certain formats, content delivery methods, or digital tools might more effectively address specific group needs. Targeted findings could help inform design features within involvement systems and tailor support that warrants diverse literacy backgrounds.

In summary, examining the historical development and role of data, alongside the positioning of data literacy as a societal competency, provides a vital foundation for understanding the challenges of designing digital involvement in a datafied society. Insights into citizen perspectives underscore the importance of training diverse data-literacy competencies. They also support the idea that involvement systems should cater to a range of needs — from basic data provisioning and comprehension support to more advanced scenarios for skill enhancement and fostering autonomy.

# 6. Case Study: DIP in Urban Planning

## 6.1. Enablement Factors for XR Participatory Urban Planning

### 6.1.1. Introduction

To face the complex challenges of designing sustainable cities, urban planning projects are increasingly relying on citizen participation (Allianz Vielfältige Demokratie/ Bertelsmann Stiftung, 2018; Wolf et al., 2020). Citizens are not only local experts in their environments but also the most affected by planning decisions. Their perspectives are crucial for developing new, sustainable solutions and the success of any identified measure depends on their acceptance (Allianz Vielfältige Demokratie/ Bertelsmann Stiftung, 2018; Amado et al., 2009). Thus, providing citizens with a voice in planning and implementation of urban areas is essential for creating more sustainable cities that incorporate their needs (Amado et al., 2009; Videira Lopes and Lindstrom, 2012; Wolf et al., 2020). E-participation, especially, has gained recognition as a means to achieve the *2030 UN Sustainable Development Goals* (Palacin et al., 2021). A significant challenge in enabling successful participation in urban planning lies in the effective communication of often intricate construction plans or abstract visions for the future (Engel and Döllner, 2012; Wolf et al., 2020). Providing comprehensible information is the foundational level of citizen participation (Arnstein, 1969; Palacin et al., 2021). For further degrees of participation, it is crucial that those involved can understand given information and implied consequences (Engel and Döllner, 2012; Videira Lopes and Lindstrom, 2012). Therefore, supporting participants' abilities to understand and contextualize information as a basis to derive action, is of central importance in participation processes (Hayek, 2011). However, this is not trivial. Information can include various data types (Billger et al., 2017), and while visualizations must aid citizens in the necessary literacy to understand and use this data, they must also facilitate unbiased participation (Hayek, 2011). Additionally, they need to be affordable and deployable by local governments, which are often the participation organizers.

Recently, some urban planning initiatives have turned to 3D visualization and XR technologies (Eilola et al., 2023; Wolf et al., 2020). XR has gained much attention, fueled by companies such as Meta, announcing the Metaverse and, among others, its usefulness for concretizing construction planning (Tech at Meta, 2021). In comparison to approaches like 2D maps, XR is said to provide alternative access to complex planning data (Engel and Döllner, 2012). However, it is also linked to considerable costs, e.g., the acquisition of technology, the training of personnel, and user support. Beyond the hype, it becomes imperative to verify whether the usage of XR technology can genuinely contribute to improving participants' literacy in the context of urban planning. Current research reports mixed results on retained information and their correctness or confidence in one's own decision (Fegert et al., 2020; Reinwald et al., 2014; Van Leeuwen et al., 2018). It seems not yet clear which aspects of XR usage for participation processes in urban planning are influential, as factors like human support can impact the participation process (Van Leeuwen et al., 2018). Thus, deriving effects from the literature is challenging, as initiatives and reporting are very heterogeneous (Eilola et al., 2023) and different aspects are measured when investigating participants' literacy. As information builds the basis for participation and decision-making (Arnstein, 1969), researching how XR can contribute to literacy seems relevant.

We thus propose the overarching research question: *How does the usage of XR affects participants' literacy in the context of participatory urban planning and what specific factors are crucial for successful XR participation formats?* In this work, we want to lay the necessary foundation for answering this research question, by contributing a theoretical mapping of prior research and a pretest of an experimental study design. By presenting this preliminary work, we aim to inform the design process of further research studies and encourage a more nuanced consideration of XR for participation in urban planning to develop practical guidelines for its effective use.

The first part of this chapter is structured as follows: First, the analysis of current literature will be reported, where we investigate differentiating factors of XR usage and aspects of participants' literacy, utilizing the framework by Eilola et al. (2023). Second, the exploratory experiment with n=41 participants is presented, where we investigate how effects on literacy aspects could be measured. We compare VR participation with a Head-Mounted Display (HMD) with a paper-based process and test the scenarios both in supported and in unsupported situations to differentiate the effects of given support.

### 6.1.2. Theoretical Background

**E-Participation and Urban Planning:** E-participation, a form of e-democracy, utilizes information and communication technology to enable citizens to participate in policy decisions (Macintosh, 2004). While other forms of e-democracy focus on the electoral processes (Macintosh, 2004), e-participation evolves around concepts of collaboration (Sanford and Rose, 2007). In the application of e-participation for urban planning, information and communication technology usually consists of visualization technology or geographic information systems (Sanford and Rose, 2007). In general, urban planning is understood as the design and regulation of spatial development in cities and rural areas, including dimensions such as physical forms, economic functions, or social activities (Schubert, 2015; Wolf et al., 2020). Participatory formats in urban planning can include formal, mandatory instruments, or informal participation (Wolf et al., 2020) and support different tasks in the planning process, e.g. informing and motivating, generating ideas, or making decisions (Hayek, 2011). The urban planning use case poses a specific challenge to participation, as the understanding of spatial and geographic data presents an issue for many people (Engel and Döllner, 2012; Wolf et al., 2020), implying that the participation has to be regarded as a learning process (Wolf et al., 2020). Participation involves processing information, extracting and contextualizing it (Hayek, 2011; Wissen et al., 2008), which can include quantitatively measurable data, e.g., climate, noise, or speed data, but also subjective data on topics such as aesthetics, well-being or safety (Billger et al., 2017). Therefore, required literacy to participate in urban planning includes aspects of data literacy e.g., the understanding and interpretation of data, or the derivation of action (Schüller et al., 2019), such as idea generation, or decision making (Hayek, 2011).

**XR for Participatory Urban Planning:** With the introduction of the Metaverse, the term “XReality”, has been coined to account for the spectrum of immersive systems to provide formats of digital reality (Dwivedi et al., 2022; Rauschnabel et al., 2022). The umbrella term covers, but strictly separates, VR and augmented reality (AR), conceptualizing them each on an individual continuous spectrum. The differentiation is based on the presence of the physical environment: While with AR, the physical environment is part of the user’s view, with VR, it is entirely replaced. Therefore, XR does not conceptualize “X” as “extended”, which linguistically fails to capture VR, but as a placeholder for any format of digital reality (e.g., assisted, mixed, atomistic/ holistic virtual) (Dwivedi et al., 2022; Rauschnabel et al., 2022). In participatory urban planning, a variety of formats and technologies are used to display 3D planning data (Eilola et al., 2023); therefore, the concept of XR seems suitable for discussing their usage. In pilot studies, authors report a positive effect of XR on participants’ motivation and engagement level in urban planning processes (Fegert et al., 2020; Van Leeuwen et al., 2018), and its potential to reach new participant groups (Reinwald et al., 2014). However, impacts on users’ literacy are not clear. Some studies report positive impacts of XR on the general understanding of planning content (Fegert et al., 2020; Reinwald et al., 2014) or on imagining dimensions (Fegert et al., 2020) and positions (Reinwald et al., 2014). Additionally, studies find that XR eases participation, supporting the participation process (Schrom-Feiertag et al., 2020). Other studies, however, find positive effects on the vividness of participants’ memories, but not on their accuracy or correctness (Van Leeuwen et al., 2018). Instead of the visualizations’ degree of immersion, perceived competence increase depended positively on the presence of a support team (Van Leeuwen et al., 2018). Effects on participants’ data-driven actions, in the form of self-confidence in their own voting, were also not found (Van Leeuwen et al., 2018). Additionally, depending on the data quality and representation, studies found that models in the immersive representation can feel confusing or distorted (Rzeszewski and Orylski, 2021).

Notably, these contributions investigate a variety of literacy aspects ranging from perceived to actual knowledge gain and abilities. However, to fully understand individual findings and their implications, we need to structurally differentiate investigations and potential influencing factors. A framework for analyzing the usage of 3D visualizations in participatory urban planning is given by Eilola et al. (2023), highlighting four important factors: the 3D visualization production, the user interface, the communicative engagement, and the planning context. We use this framework to analyze the aforementioned studies, allowing us to structurally describe and compare their respective XR usage for future work (see Table 6.1). It becomes apparent that there are multiple differences in the design and testing of XR, proving it challenging to understand individual effects on literacy aspects. Based on the analysis, it thus seems necessary to examine the varying factors individually, while reporting on the results in a structured manner.

| Source | 3D visualization production | User interface | Communicative engagement | Planning context | Evaluation |
|---|---|---|---|---|---|
| Fegert et al. (2020) | Realistic visualization of buildings, objects, animals | AR visualizations with tablet vs. VR with HMD; including moving items | Participation with support, non-digital communication | Artificial Use Case on zoo construction, future scenario | Between-subject, perceived and actual knowledge gain, willingness to donate |
| Reinwald et al. (2014) | Abstract visualization of objects | AR visualizations with tablet vs. 3D visualization on paper/ 2D maps | Participation with support, non-digital communication | Real use case of street design, future scenario | Between-subject, perceived knowledge gain |
| Schrom-Feiertag et al. (2020) | Abstract visualization of buildings, objects, persons | VR visualizations with HMD, including text display, survey function, traffic simulation | Participation with support, in-app and non-digital communication | Real use cases of street and railway station designs, future, hypothetical scenarios | Only demonstration, Ease of participation |
| Van Leeuwen et al. (2018) (Field study) | (Photo-) Realistic visualization of buildings, objects | VR visualizations with HMD vs. multiple others | Participation with and without support, digital and non-digital communication | Real use case of park reconstruction, future scenarios | Between-subjects perceived knowledge gain, vividness, vote certainty |
| Van Leeuwen et al. (2018) (Lab study) | Photo-realistic visualization of buildings, objects | VR visualizations with HMD vs 3D visualizations on a tablet | Participation with support, no further communication | Artificial use case of park reconstruction, future scenarios | Between-subjects perceived and actual knowledge gain, vividness, vote certainty |

| | | | | | |
|---|---|---|---|---|---|
| Rzeszewski and Orylski (2021) | Different detail and quality levels | VR visualizations with HMD and AR on smartphone; including text display, multiple other features | Participation with support with no further communication | Artificial use cases of historical, current, future scenarios | Within-subject, perceived usefulness |
| Experiment Approach | (Photo-) realistic visualization of buildings, objects | VR visualizations with HMD vs. 3D renderings on paper; including text display and voting option | Participation in a supported scenario vs. non-supported scenario | Artificial use case on court yard reconstruction | Between-subject, perceived and actua knowledge gain, willingness to vote, certainty of vote, vividness |

Table 6.1.: Analysis of studies examining impacts on participants' literacy based on Eilola et al. (2023) including the evaluation form. In grey, analysis of our own preliminary study.

### 6.1.3. Methodology

In answering the overarching research question, the analysis of the current literature has shown that a differentiated consideration of individual factors is necessary to investigate their potential influence on aspects of literacy in the context of urban planning. At the current state, insights are not sufficient to support a well-founded hypothesis for most of the identified factors, therefore, an exploratory approach can be employed (e.g., Hayek, 2011). To investigate how effects of identified factors could be measured going forward, we conduct a preliminary experiment, initially focusing on the effect of the user interface and commutative engagement on a set of literacy aspects identified in the literature.

**Experimental Design:** The experiment is designed as a four-treatment, between-subject lab experiment. In this way, we can keep a similar planning context and control external factors that might influence participation in real-life. Varying aspects of the user interface (VR with an HMD vs. paper-based participation) and the communicative engagement (with or without human support), we end up with four treatment groups being a VR treatment group under supervision (VR-Supervision), a paper-based treatment group under supervision (PB-Supervision) and a VR and paper-based treatment group without supervision (VR-/PB-Control). In the literature, we find different attempts to investigate participants' literacy, including perceived and actual knowledge gains, vividness, willingness to do a certain action, and certainty to do so (e.g., donation, vote). However, in the user interface studies differ in whether presented information is exclusively visual or also textual. We thus include all these aspects to gain a differentiated picture of potential effects on participants' abilities. The questionnaire can be seen in Table 6.2.

**Use Case and Design of the Participation Tool:** We utilize an artificial urban planning project: Participants envision themselves to be employees of a company that reconstructs their courtyard. They may choose between the construction designs of a tiny house, a food truck, or a playground, while for each of the designs, they receive 3D renderings and information. For the experiment, the use case should include both visual and textual information. Therefore, we design six information variables being dimensions, shadow casting (visual information enclosed in the renderings); environmental sustainability, user benefits (textual information); design coherence and accessibility (visual and textual information). The dimensions represent both quantifiable and subjective criteria as typically the case in urban planning (Billger et al., 2017). The visualizations and textual information are equally provided in the paper-based and the VR-tool (see Figure 6.1). On paper, participants see several 3D renderings of the objects, obtain information as text, and can vote for their favorite design. In VR, participants can see the objects from different angles, obtain information in the form of text and audio, and also vote on the designs. The 3D environment is based on 360-degree photographs and rendered in Unity. The experiment's characteristics are summarized and compared to previous work in Table 6.1 (grey).

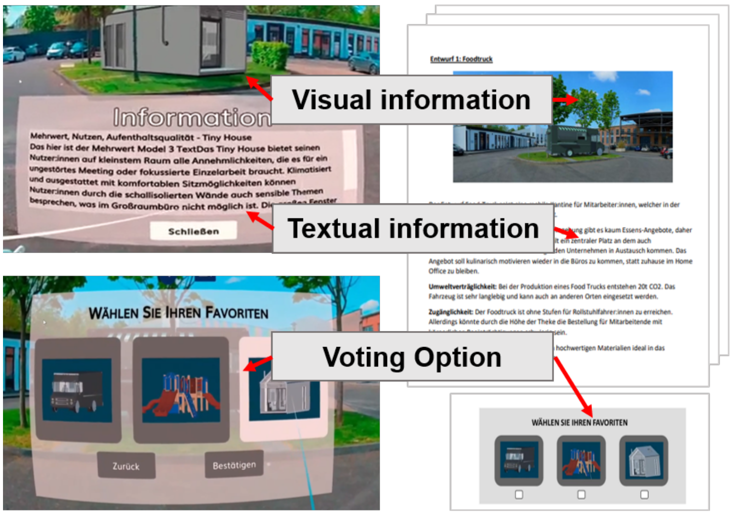


Figure 6.1.: Participation in the VR (left) and on paper (right).

**Conduction and Evaluation of the Experiment**: The experiment was conducted on December 12 and 13, 2022, with 43 participants receiving monetary compensation for their participation. Participants were randomly assigned to the treatment groups based on the registration time. All participants received an introduction to the scenario before they were instructed to participate, using the tool to inform themselves and vote for their favorite design. After participation, tools were collected and participants received a final questionnaire. For the supervised group the initial introduction was personal, and they had the possibility to ask questions throughout, including technical aid with the tool. In the unsupervised

| Dimension | Item | | Source |
|---|---|---|---|
| Actual knowledge gain | 1 | Please rank the three designs according to their dimensions. | Self-developed |
| | 2 | Please rank the three designs by their shadowing. | |
| | 3 | Please rank the three designs by their environmental impact. | |
| | 4 | Please rank the three designs according to their accessibility for wheelchair users. | |
| Perceived knowledge gain | 1 | With the support of the participation tool, I could easily imagine the dimensions of the building projects. | Adapted from Fegert et al. (2020) |
| | 2 | With the help of the participation tool, I was able to imagine how the designs would fit into the environment of the Software4You buildings. | |
| | 3 | The participation tool helped me to better appreciate the utility of the designs. | Self-developed |
| Vote | 1 | After using the participation tool, I feel more willing to vote on the structural designs. | Adapted from Fegert et al. (2020) |
| | 2 | Please indicate how confident you are with your vote. | Adapted from Van Leeuwen et al. (2018) |
| Vividness | | Please evaluate the representations of the construction project presented in the participation tool in terms of vividness. | Adapted from Peukert et al. (2019) |

*Except for items of actual knowledge gain, measured with a ranking, all items have been measured on a 7-point Likert scale.*

Table 6.2.: Experiment questionnaire.

treatment, participants were isolated in an experimental cabin, receiving initial instructions via text and a brief video introduction to the operation of the HMD. The evaluation was undertaken with the statistical software R and it considered n=41 participants, based on the fulfillment of two attention checks. Of the 41 participants, 11 were female and 30 were male. The participants were university students of different disciplines and predominantly under thirty years (97.6%). In particular, technical courses of study (e.g., engineering, computer science) were represented, sporadically also other disciplines (e.g., business administration, humanities). The group had little prior knowledge with XR ($Mean = 3.58$ on a 7-point Likert scale) and urban planning ($Mean = 2.15$, on a 7-point Likert scale). To test treatment randomization, Kruskal-Wallis tests were conducted for participants' prior experience, and chi-squared tests were conducted for age and gender. No significant differences were found ($p_{XR} = 0.790$ $p_{urban-planning} = 0.707$ $p_{Age} = 0.377$ $p_{Gender} = 0.288$), thus treatment randomization can be considered successful. For the examination of group differences of the other variables Kruskal-Wallis tests were conducted, due to a violation of the assumption of normal distribution.

### 6.1.4. Preliminary Results

The results are reported by dimension and summarized in Table 6.3. In terms of actual knowledge gain, results differed strongly between the information variables. While 80% of the participant could sort the proposals according to their accessibility, less than 32% could indicate the right dimensional ranking. No clear tendency can be noticed, whether textual or visual information was easier to remember and likewise, no significant differences between the treatments were found. This was different for the case of perceived knowledge gain. For visually transferred information, significant differences between the treatments were found ($p_{perceived1} = 0.059$, $p_{perceived3} = 0.008$), while for written information results did not differ significantly. For perceived knowledge gain on dimensions and the environmental fit, participants in the cabin treatments each rated their learning experience better than participants in the supervised groups, and participants in the VR treatments better or the same as participants in the paper-based treatments. Interestingly we measured no linear relationship between participants' actual and perceived knowledge gain on the dimensions of the designs. For the vividness of the designs, impressions were slightly higher for both cabin-treatments compared to their supervised counterpart, however effects were not significant. Interestingly, it was not consequently higher for the VR treatments than for the paper-based treatments. Finally, average willingness to vote was higher in both VR treatments compared to their paper-based counterpart, however, differences were not statistically significant. For the certainty of the vote, no clear tendencies could be found.

| **Dimension and Item** | | **Treatment** | | | | | | | | **Kruskal Wallis** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **PB-Supervision** | | **PB-Control** | | **VR-Supervision** | | **VR-Control** | | |
| | | Mean | SD | Mean | SD | Mean | SD | Mean | SD | |
| Actual knowledge gain | 1 | 0.3 | - | 0.2 | - | 0.5 | - | 0.27 | - | 0.5249 |
| | 2 | 0.7 | - | 0.6 | - | 0.5 | - | 0.73 | - | 0.7081 |
| | 3 | 0.9 | - | 0.4 | - | 0.4 | - | 0.64 | - | 0.0766 |
| | 4 | 0.9 | - | 0.8 | - | 0.9 | - | 0.64 | - | 0.3815 |
| Perceived knowledge gain | 1 | 4.8 | 1.75 | 5.3 | 1.64 | 5.3 | 1.42 | 6.45 | 0.69 | 0.0589 |
| | 2 | 5 | 1.15 | 5.3 | 1.77 | 6 | 0.67 | 6.45 | 0.52 | 0.0083 |
| | 3 | 5.8 | 1.03 | 5.5 | 1.43 | 5.5 | 1.18 | 4.91 | 1.97 | 0.8393 |
| Vote | 1 | 5.3 | 1.16 | 4.7 | 1.89 | 5.8 | 0.63 | 6.18 | 0.75 | 0.1211 |
| | 2 | 5.2 | 0.63 | 4.3 | 1.42 | 4.8 | 0.92 | 5.55 | 0.69 | 0.1034 |
| Vividness | | 4.2 | 1.4 | 4.4 | 1.65 | 4.2 | 1.4 | 4.64 | 1.21 | 0.8989 |

Table 6.3.: Descriptive statistics for the experiment variables, grouped by treatment.

### 6.1.5. Discussion and Conclusion

XR technologies are highlighted as means to enable civic participation in urban planning helping to design sustainable cities of the future (Eilola et al., 2023; Schrom-Feiertag et al., 2020; Videira Lopes and Lindstrom, 2012). Yet, validating their capability to effectively inform users requires a more nuanced assessment. In the current literature, variations in the design, usage and evaluation of XR make it difficult to isolate individual effects on participants' literacies. Therefore, we analyzed five experimental studies in depth, providing a basis for comparison and targeted testing of individual factors. We also conducted an experiment to explore how effects can be measured on a set of literacies, initially focusing on the user interface and commutative engagement. By providing textual and visual information and voting options on an urban planning project equally on paper and in a VR tool, in a supervised and unsupervised format, our experimental design allows us to compare individual effects of the medium VR and the presence of support on perceived and actual knowledge gain, vividness and voting behavior. Due to the exploratory character of our study, our findings are yet limited in expressiveness but can still point towards streams for further research. First, opposing Van Leeuwen et al. (2018), the provided support in our case did not significantly improve aspects of participants' literacy, but rather visual, perceived knowledge gain and vividness were even better in the cabin treatments. For practical implications, this is interesting, as providing human support implies additional efforts for organizers, which might be avoidable for certain use cases (e.g., young target groups, planning contexts with low complexity). Besides, cabins provide a form of isolation to the users, which might positively affect their visual perception. If reproducible, this might also have implications for the practice of e-participation, as cabins could be integrated, e.g., in on-site showings. Striking about our results is certainly the absent effects of the VR tool on actual knowledge gain in comparison to the paper-based tool. This resembles results of Van Leeuwen et al. (2018), however contradicts assumptions that could have been made based on Fegert et al. (2020). Although users of the VR treatment perceived their knowledge for visual items higher, their actual knowledge did not seem to be positively influenced compared to the paper-based group. This might have important implications for further evaluations of XR urban planning tools, which should differentiate between perceived and actual knowledge effects. For voting, our results look promising, towards a positive effect of VR, although we could not establish statistical significance. Like Van Leeuwen et al. (2018) we could also not find significant effects for the vote confidence. To further explain these differences in the effects on participants' abilities, it could be useful to consider established cognitive learning theories (e.g. Makransky and Petersen, 2021) in the evaluation of participation processes, building a conceptual literacy model for this use case.

Overall, mean values for all items and groups positively exceeded the neutral level of 4, indicating a positive participation experience across treatments. This is interesting, as the effort and costs in building and using the VR tool significantly exceeded those of a paper-based participation tool. As such, cost-benefit considerations could possibly speak against VR under certain circumstances, motivating further research. However, it is important to consider two major limitations of our study, when interpreting its results. First, the preliminary nature of the study results in limitations regarding the conclusiveness of observed effects and the transferability to more complex construction projects. Second, since we used a student panel to have an ideally homogeneous group, we omit other target groups of e-participation projects. Thus, in real-life projects where planning scenarios might be more complex, or participants have linguistic, physical, or skill-related barriers, XR tools or support teams could produce other effects.

The results and limitations of our research, offer numerous starting points to inform further research. They aim to inform the derivation of hypothesis for the examined factors while inviting studies to also individually test the effects of other factors where current studies differ (e.g., 3D visualization production or planning context). Overall our insights indicate that XR might not automatically be a better solution than traditional alternatives to enable informed participation. This motivates, once again to deepen our understanding of the specific design and implementation factors that are crucial to add value to participants' literacy when utilizing XR. Only in doing so can we fully exploit the potential of XR for citizen participation, empowering citizens and experts to jointly develop sustainable city solutions.

## 6.2. 3D Visualization Types for E-Participation in Urban Planning

### 6.2.1. Introduction

As urban spaces continue to grow, designing them as sustainable and livable environments for all residents becomes increasingly challenging (Bedi et al., 2023; United Nations, 2018). Therefore, urban planning remains a central application field for e-participation (Le Blanc and United Nations, 2020). Multiple initiatives have demonstrated the importance of citizen involvement to improve decision-making while establishing mutual understanding and trust (Allianz Vielfältige Demokratie/ Bertelsmann Stiftung, 2018). However, a crucial element of effective involvement remains the communication of complex spatial and geographical data in an inclusive and comprehensible manner (Engel and Döllner, 2012; Wolf et al., 2020). Among the various visualization techniques, virtual 3D city models stand out for their potential to provide accessible representations of urban environments (Engel and Döllner, 2012). Therefore, the Association of German Cities specifically recommends their usage to support citizen participation (Deutscher Städtetag, wegewerk, 2017). On the other hand, despite an increase in their availability (Lei et al., 2023), particularly high-resolution 3D models are expensive and time-consuming to create and rarely available for free (Girindran et al., 2020). It is, therefore, not surprising that in research and practice, the 3D models used in e-participation greatly differ from one another in terms of data sources or model generation and, as such, in their level of detail, realism, and resolution (Eilola et al., 2023; Lei et al., 2023). This raises the question of how differences in presentation affect the participation process, especially considering that visualizations must facilitate an informed and unbiased involvement for all participants (Hayek, 2011). While there is already an ongoing scientific debate about the potential and impacts of new technologies such as AR and VR for e-participation (Fegert et al., 2020; Reinwald et al., 2014; Schrom-Feiertag et al., 2020; Van Leeuwen et al., 2018), substantial evidence-based knowledge to guide the implementation of underlying 3D visualizations is still missing (Billger et al., 2017; Eilola et al., 2023). We thus pose the research question: *How do properties of 3D visualizations affect their suitability for processes of e-participation in the context of urban planning?*

This work addresses the research question in a two-step research approach using qualitative and quantitative research methods. First, we conducted qualitative interviews with stakeholders of urban planning projects ($n = 14$), discussing the potential of low-detail versus high-detail planning visualizations for e-participation from an expert perspective. Second, we adopt a user respectively citizen-centric perspective, comparing 3D city models differing in their level of detail, realism, and resolution. In an online experiment ($n = 512$), we examine how users experience the visualizations in terms of their motivating, informative, and trust-building capacities, inquiring about their suitability for different participatory scenarios. By adopting expert and citizen perspectives and combining qualitative and quantitative research methods, we provide comprehensive insights into the different effects of 3D visualizations as an important basis for their use in participatory urban planning and the design and evaluation of traditional and immersive e-participation systems.

### 6.2.2. Related Work

**E-Participation in Urban Planning:** Urban planning is concerned with the spatial development in cities and rural areas and can span physical as well as economic or social dimensions (Schubert, 2015; Wolf et al., 2020). Making use of ICT, e-participation, enables civic participation in public policy-making (Macintosh, 2004; Pristl and Billert, 2022; Sanford and Rose, 2007) and has been discussed for years in the context of urban planning. In this context, ICT is usually employed for forms of visualizations and geographic information systems (Sanford and Rose, 2007). Participation formats in urban planning can vary in their formality level and focus on different stages and tasks in the process (Eilola et al., 2023; Hayek, 2011; Wolf et al., 2020). The Spectrum of Public Participation generally defines five participatory levels from information to empowerment (International Association of Public Participation, 2007). As a baseline, they all require citizens to process and contextualize information (Hayek, 2011; Wissen et al., 2008), implying that a trustful information basis needs to be created. For visualizations to evoke trust, an interplay of credibility and reliability, and the user's familiarity and understanding must be given (Pandey et al., 2023).

**The Usage of 3D Visualizations:** In participatory planning, 3D visualizations typically refer to georeferenced visualizations, representing the real world (at least in parts), such as 3D city models or 360° panoramic street views (Eilola et al., 2023). Conducting a literature review, Eilola et al. (2023) mapped common characteristics of 3D visualizations in e-participation, highlighting a rich variety of approaches. Yet they noted a lack of standardized reporting and rigorous suitability evaluation. Lei et al. (2023) confirmed this heterogeneity from a technical perspective, evaluating 40 open datasets and noting a lack of thematic and attribute content for many models. To describe model complexity, the international standard differentiates between four levels of detail (LOD) from basic block- and form-based structures (LOD1, LOD2) to detailed visualizations (LOD3, LOD4) (Biljecki et al., 2016). Their influence on aspects of the e-participation process has been investigated by several scholars. Studies suggest more detailed visualizations enhance participant motivation through higher immersion (Onyimbi et al., 2017; Wu et al., 2010) and improve clarity and understanding (Imottesjo and Kain, 2022), which could promote a more trustful information basis. Furthermore, high-detail visualizations seem to be useful at various participation levels, facilitating understanding, critical assessment, and stakeholder interaction (Imottesjo and Kain, 2022; Lafrance et al., 2019; Onyimbi et al., 2017; Wu et al., 2010). Yet, some authors highlight the usefulness of both abstract and detailed visualizations, with their individual (dis-) advantages (Hayek, 2011). Abstract visualizations appear to be useful, especially in the initial stages of planning projects (Billger et al., 2017), facilitating focus on certain attributes and data aspects (Eilola et al., 2023), which might reduce information overload.

**High-Detail Visualizations in Practice**: The choice of high-detail 3D visualizations involves considering costs and benefits as their high-quality generation is usually expensive and time-consuming (Girindran et al., 2020). With a constrained budget, high-detail visualization, in comparison to low-detailed ones, may suffer in quality, either in the level of resolution or realism. Rzeszewski and Orylski (2021) utilized a public LOD2 city model, manually rendering textures with Google Earth data to increase the detail level, with the result that participants negatively reported on the models' distortion. Similarly, Imottesjo and Kain (2022) used an open data, photogrammetric mesh model with poor close-up resolution, which was received insufficient for understanding and interacting. Procedurally generated city models, produced by systems such as "CityGen" offer a cost-effective approach to producing high-detail, high-resolution visualizations by automating parts of the city modeling process based on a set of predefined rules and inputs (Deng et al., 2023; Kelly, 2006, 2007). Originating from the gaming industries, their application is also proposed for participatory planning scenarios (Roumpani, 2022). However, since models do not guarantee the actual realism of all components, their trusting capacity as an information basis can be impaired.

**Research Gap and Hypotheses:** Building on this theoretical foundation, we identify the need to further elaborate on the (dis-) advantages of 3D visualizations with varying levels of detail while integrating a more practical orientation by considering the factor of quality for high-detail city models. We derive the following hypotheses of their suitability for the context of participatory urban planning (Hypotheses 1-4):

Hypothesis 1: Highly detailed compared to low-detailed 3D city models

a. increase participants' motivation,
b. provide a more trustful information basis and
c. introduce a higher information overload for participants.

Hypothesis 2: Highly detailed 3D city models are equally suitable for all participatory levels, while low-detail ones are more suitable for initial participation levels.

Hypothesis 3: Detailed, low-resolution, compared to detailed, high-resolution 3D city model

a. demotivate participants,
b. provide a less trustful information basis and

c. introduce an equal information overload for participants

resulting in their diminished usefulness for all participatory levels.

Hypothesis 4: Detailed, low-realism, compared to detailed high-realism 3D city models

a. equally motivate participants,

b. provide a less trustful information basis and

c. introduce an equal information overload for participants,

resulting in their diminished usefulness for all participatory levels.

### 6.2.3. Methodological Approach

To answer our research question, we follow a two-step research approach. First, we set out to gain qualitative insights by conducting expert interviews with stakeholders from the urban planning domain. Second, we aimed to quantitatively contrast expert perspectives from the participants' points of view, conducting an online experiment.

**Expert Interviews:** We conducted 14 semi-structured expert interviews, following Kaiser (2014). We invited a variety of stakeholders, including city administrations *(Urban planning: B1, B4; Traffic planning: B2; Green space planning: B3; City council: B6; Children and youth representative: B7; Citizen participation representative: B5)*, business and research *(Planning offices: B9, B14; Traffic engineering: B11)*, and associations and cooperatives *(Sustainability association: B8; Building cooperative: B10; Association for the blind and visually impaired: B13; Disability advisory council: B12).* Interviews were conducted and recorded digitally and took place between March and April 2022. The transcription and evaluation were conducted with MAXQDA, following Dresing and Pehl (2018) and Kuckartz and Rädiker (2022). While the interviews covered various aspects of participation in urban planning (e.g., incentives, inclusion concepts), this work focuses solely on the visualization-relevant parts: Experts were shown an abstract and a detailed building design visualization and were asked about their understanding and thoughts on the suitability of these visualizations for informing and collaborating with laypeople in an e-participation application.

**Online Experiment:** Using a within-subject approach (Charness et al., 2012) we designed an artificial urban planning scenario for the participatory re-design of a public space ("Karlsplatz") in Munich. Including aspects of quality into the comparison, we created four 3D models of the public space, with varying levels of detail, realism, and resolution (see Figure 6.2). Model A, the highest in detail, resolution, and realism, was based on 360° camera images. Model B was a highly detailed, procedurally generated model using open street map data and arbitrarily filling in missing information e.g., textures. For model C photogrammetry was used to create a mesh model based on aerial photos, providing an accurate yet poorly resoluted model of the Karlsplatz. Finally, model D was a low-detail LOD2 city model based on aerial data. Both models, C and D, were provided by the city of Munich. To avoid any confounding effects, all city models were equally accessible through a web browser and contained 5 teleportation points for navigation and no further features. Randomizing the order of the visualizations, for every city model, participants were asked to first explore the model before then proceeding to answer a brief treatment questionnaire (see Table 6.4). Conscientious engagement with the city models was ensured by requiring participants to find a hidden code before they could continue. The treatment questionnaire was constructed utilizing established constructs from the literature, including Pandey et al. (2023) for trust in visualizations, and Metag and Gurr (2023) for information overload. We assessed motivation with different items from Fegert et al. (2020) and based items for perceived usefulness for different participatory levels on the Spectrum of Public Participation (International Association of Public Participation, 2007). All items were assessed on a seven-point Likert scale (1- Completely disagree to 7- Completely agree). Concrete usage opportunities identified by Eilola et al. (2023) could be indicated by participants with a multiple-choice selection. For the conduct of the experiment from December 2023 to January 2024, we utilized the research panel provider Prolifc (Palan and Schitter, 2018), ensuring a diverse research panel and paying monetary compensation. The experiment design was pretested with 30 participants, before proceeding with the data collection. After data cleaning based on the fulfillment of an attention check, the sample used for analysis consisted of 512 participants. Conducting data analysis in R, we assessed internal scale consistency with Cronbach's Alpha, clearly exceeding consistency thresholds of 0.7 for all constructs (Tavakol and Dennick, 2011). Due to the within-subject design, we utilized repeated measures ANOVA to test for inter- and intra-model differences and applied a Greenhouse-Geisser correction in the case of measured sphericity. We applied post hoc analysis in the form of pairwise t-tests and reported adjusted p-values with a Bonferroni correction to account for the paired samples. Effect sizes were assessed according to Cohen (1992). To gain explorative insights into the potential effects of demographic factors, we conducted joint regressions of factors onto the observed variables.

### 6.2.4. Results

#### 6.2.4.1. Expert Interviews

The results of the expert interviews generally showed a strong preference towards more realistic representations, while also highlighting some practical challenges imposed by this choice. According to the interviewees, a more realistic representation is easier to understand (B2, B4, B5, B9, B11, B12) while enabling a more detailed assessment, which is especially relevant when aspects of physical accessibility should be considered (B12, B13). One participant highlighted the necessity of details when voting on specific elements as part of the participation (B3). Abstract visualizations might limit participants' imagination (B9) or impede orientation and understanding (B11), being typically rather aimed at experts than lay people (B8, B14). On the other hand, those responsible for planning also see an increasing risk of manipulation as the level of detail increases (B1, B3, B6). Especially when considering different planning alternatives, minor differences in environmental parameters, such as weather and colors,

| **Construct** | **Item** |
|---|---|
| 1. Motivation, Enjoyment<br>*Cronbach's Alpha: 0.95* | 1.1 The model motivates me to take part in the urban planning project. |
| | 1.2 After using the model, I have a stronger interest in the urban planning project. |
| | 1.3 I enjoyed interacting with the model. |
| 2. Perceived Usefulness for Participation<br>(Inform, Consult, Involve, Collaborate, Empower)<br><br>*Cronbach's Alpha: 0.97* | 2.1. The model provides a good basis for informing citizens about urban planning projects and helping them to understand construction problems, alternatives, and/ or solutions. |
| | 2.2. The model provides a good basis for citizens to give feedback on plans, alternatives, and/ or decisions in urban planning. |
| | 2.3. The model provides a good basis for working with citizens in urban planning projects to understand and address their concerns. |
| | 2.4. The model provides a good basis for developing ideas for urban planning together with citizens and selecting desirable solutions. |
| | 2.5. The model provides a good basis for citizens to make final decisions on urban planning. |
| 3. Information Overload<br><br>*Cronbach's Alpha: 0.93* | 3.1. When looking at the model, I feel overwhelmed by the amount of details shown about the Karlsplatz site. |
| | 3.2. The model gives me more information about the Karlsplatz site than I can actually process. |
| | 3.3. When looking at the model, I am confronted with too many details about the Karlsplatz location. |
| 4. Trust<br>(Familiarity, Clarity, Credibility, Reliability, Confidence)<br><br>*Cronbach's Alpha: 0.90* | 4.1. I am familiar with the representation of this model. |
| | 4.2. I understand what the model is showing me. |
| | 4.3. I believe that the model shows the real world. |
| | 4.4. I would rely on the representation of this model. |
| | 4.5. I would feel confident using this model to make a decision. |
| 5. Usage Opportunities<br>(Features)<br><br>"It would be useful to be able to..." | 5.1. … read information. |
| | 5.2. … give feedback. |
| | 5.3. … place objects. |
| | 5.4. … discuss with others. |
| | 5.5. … see myself as an avatar. |
| | 5.6. … read the opinions of other participants. |
| 5. Usage Opportunities<br>(Scenarios)<br>"The model would be useful to..." | 5.7. … show future visions of the city. |
| | 5.8. … show the current situation. |
| | 5.9. … show hypothetical scenarios. |
| | 5.10. … show historical reconstructions. |

Table 6.4.: Treatment questionnaire with measured consistency scores for constructs.

could significantly influence their perception (B1). Furthermore, visualizing unconcrete future planning scenarios in detail could create false expectations for their implementation (B2, B5, B9, B10). Therefore, the representation should be chosen in accordance with the planning progress (B1, B2, B10), with abstract representations being introduced especially in an early planning phase to facilitate transparent communication and the collection of ideas (B1, B2, B3). In summary, experts assume that the detail level

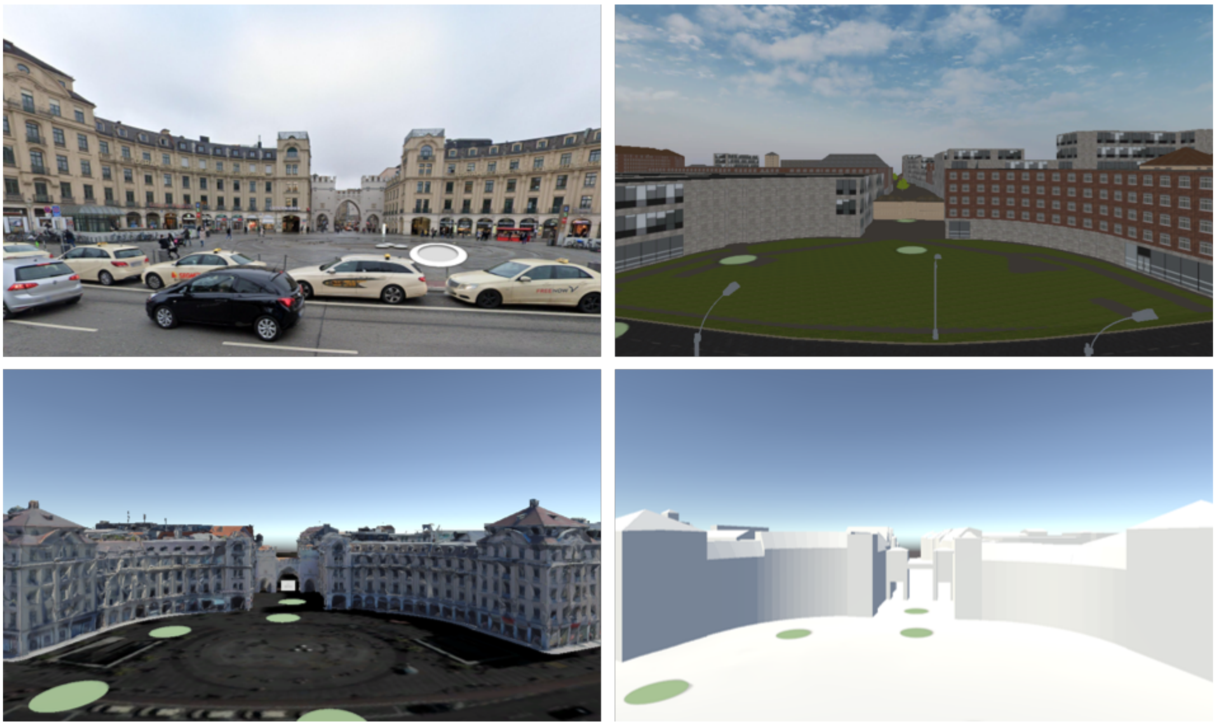

Figure 6.2.: 3D city models of the experiment: Model A (top-left), model B (top-right), model C (bottom-left), model D (bottom-right).

of supporting 3D models plays a vital role in provoking (mis-) understanding for participants in urban planning initiatives. The appropriate level of detail could depend on several factors, e.g., the skills of the viewer, the activity and level of participation, but also the general context of the urban planning scenario.

#### 6.2.4.2. Experiment

The results of the online study are reported by research dimension: motivation, perceived usefulness for different participatory levels, information overload, and trust. An overview of their summary statistics is given in Table 6.5 and results are visualized in Figure 6.3. Finally, we report on the exploratory analysis of indicated usage opportunities and participants' demographics. Regarding the panel's composition, the study participants consisted of 241 female and 271 male participants, between 18 and 71 years old ($Mean = 29.55$). As a prerequisite, we required all participants to be fluent in the study's language German, however participants still came from multiple different countries. Overall for 27.15% of the participants, German was not their native language. While 44.53% of the participants have visited the Karlsplatz, the location of the e-participation scenario, only 12.10% currently live or have lived in Munich. In terms of education, 140 participants had a degree higher than a bachelor/ pre-diploma, 155 had a bachelor/ pre-diploma, 114 had "Abitur" as the highest German highschool diploma, 61 had a lower school diploma, and 42 participants indicated other or no details regarding their education. A limited group of 32 participants had prior experience with participatory urban planning.

| Dim | Item | | A | B | C | D |
|---|---|---|---|---|---|---|
| 1. Motivation | 1-3 | Mean | 5.34 | 4.42 | 3.32 | 3.25 |
| | | SD | 1.28 | 1.68 | 1.68 | 1.62 |
| 2. Participation Usefulness | 1 | Mean | 5.47 | 4.96 | 3.69 | 3.66 |
| | | SD | 1.27 | 1.48 | 1.70 | 1.69 |
| | 2 | Mean | 5.34 | 4.86 | 3.61 | 3.62 |
| | | SD | 1.31 | 1.50 | 1.70 | 1.67 |
| | 3 | Mean | 5.37 | 4.79 | 3.59 | 3.59 |
| | | SD | 1.32 | 1.50 | 1.73 | 1.70 |
| | 4 | Mean | 5.32 | 4.88 | 3.60 | 3.66 |
| | | SD | 1.36 | 1.54 | 1.68 | 1.70 |
| | 5 | Mean | 5.16 | 4.65 | 3.38 | 3.42 |
| | | SD | 1.48 | 1.65 | 1.74 | 1.74 |
| | 1-5 | Mean | 5.33 | 4.83 | 3.57 | 3.59 |
| | | SD | 1.23 | 1.44 | 1.63 | 1.60 |
| 3. Information Overload | 1-3 | Mean | 3.76 | 2.38 | 2.74 | 1.92 |
| | | SD | 1.76 | 1.35 | 1.46 | 1.26 |
| 4. Trust | 1 | Mean | 5.80 | 4.70 | 4.02 | 4.01 |
| | | SD | 1.17 | 1.50 | 1.66 | 1.70 |
| | 2 | Mean | 6.19 | 5.50 | 4.74 | 4.58 |
| | | SD | 0.84 | 1.30 | 1.64 | 1.74 |
| | 3 | Mean | 6.49 | 3.71 | 3.24 | 2.64 |
| | | SD | 0.85 | 1.78 | 1.81 | 1.59 |
| | 4 | Mean | 6.08 | 4.29 | 3.27 | 3.30 |
| | | SD | 1.00 | 1.57 | 1.76 | 1.67 |
| | 5 | Mean | 5.75 | 4.40 | 3.12 | 3.15 |
| | | SD | 1.26 | 1.64 | 1.81 | 1.70 |
| | 1-5 | Mean | 6.06 | 4.52 | 3.68 | 3.54 |
| | | SD | 0.75 | 1.24 | 1.41 | 1.37 |

Table 6.5.: Mean and SD for motivation, participation usefulness, information overload and trust, grouped by 3D model.

**Motivation:** Participants reported significantly different perceptions of the motivating capabilities of the four visualizations ($p < 0.001$). While both model A and B are, on average, perceived motivating, this is not true for Model C and D. Pairwise t-tests showed that significant differences are to be seen between all visualization formats ($p < 0.001$), except models C and D ($p > 0.05$). The effects are largest for comparison of model A with the models C and D ($d > 1$), moderate for the models C and D with model B ($|d| > 0.58$), and small for model A with model B ($|d| = 0.50$).

**Perceived Usefulness:** There are significant differences in the mean usability scores averaged across all participatory levels ($p < 0.001$). Differences are significant between all models ($p < 0.001$), except models C and D ($p > 0.05$). Effect sizes are large for comparisons between model A and model C/ D ($d > 0.90$), while they are small for the comparison of model A with model B ($|d| = 0.30$). Comparing model C and D with model B we see moderate effects ($|d| > 0.69$). Overall, there is little variation in the indicated usefulness across different participatory levels. Comparing average scores for the different

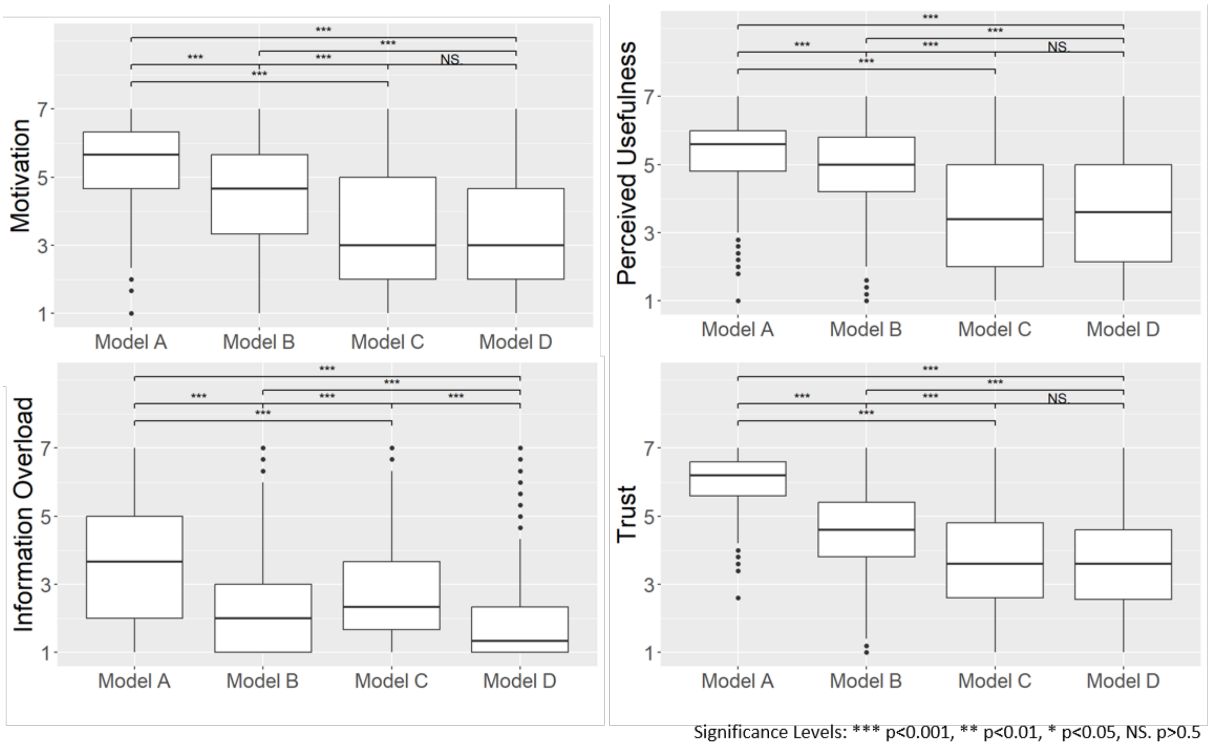


Figure 6.3.: Boxplots for the dimensions motivation, perceived usefulness, information overload, and trust grouped by model including results of pairwise t-tests.

levels within each model, effect sizes for all tests are small to negligible.

**Information Overload:** For the information overload caused by the four city models, significant differences could be noted ($p < 0.001$). Pairwise t-tests confirmed significant differences between all four visualizations ($p < 0.001$). The largest effect is to be seen in the comparison of model A with model D ($|d| = 0.96$), while the effects between model A and all other models are moderate ($|d| > 0.53$). Model D is perceived as less overloading than models C and B, with a moderate and small effect size ($|d| = 0.55$ and $|d| = 0.36$). Likewise, measured differences between model C and model D only have small effect sizes ($|d| = 0.25$). Importantly, participants, on average, negate an information overload for all investigated models.

**Trust:** Investigating participants' trust in the visualizations, significant differences between the models can be seen ($p < 0.001$). Differences are significant between all models ($p < 0.001$), except models C and D ($p > 0.05$). Model A receives the highest scores with large effect sizes in comparison to all other models ($|d| > 1.2$). Interestingly, model A is followed by model B, which shows moderate effect sizes in comparison to models C and D ($|d| > 0.52$). Both models, A and B, receive, on average, a positive mean trust score, while they are negative for the other models. Exploring the different factors, contributing to the trust construct, we see that model A scores best for all variables. The average score of model B is usually closer to the ratings of models C and D except for the item measuring understanding. Differences between the model C and D are small to negligible for all variables.

**Usage Opportunities:** To explore concrete usage opportunities, we inquired about both the usability of certain features within the models as well as their usability for certain scenarios. The most useful feature indicated across all city models, was to read additional information in the model ($\geq 70\%$). For models B, C, and D, in second place, the ability to place objects was mentioned ($\geq 57\%$), while for model A, giving feedback was mentioned more frequently (56%). Interestingly, communication-related features like communicating with others (38%), and reading others' opinions (42%) were the only characteristics that were named more frequently for model A than for other models. The last favored feature across all models was the ability to see oneself as an avatar (between 17% and 23%). In terms of scenarios, for future and hypothetical scenarios, model B scores best (85% and 79%), followed by model D (71% and 70%), C (58% and 60%), and model A (58% and 51%). For depicting the current scenario, on the other hand, model A scores best by far (93%), followed by model B (66%), model C (53%), and model D (44%). Finally, the historical reconstruction scenario generally scored the least across all models. Here model C scores best (48%), followed by model B (45%), model A (44%), and finally model D (36%).

**Demographic Factors:** Investigating correlations between demographic data and observed variables, we find that generally, for all models, demographic details explained very little of the observed variation in motivation, trust, usefulness for participatory levels, or information overload ($AdjustedR^2 < 0.07$). In the following only significant relations ($p < 0.05$) will be reported. For the factor gender, male participants reported more trust in models C and D ($\beta = 0.61$ and $\beta = 0.42$), while for model C, their information overload was lower than for female participants ($\beta = -0.37$). For model D, male participants exhibited higher levels of motivation ($\beta = 0.47$) and usefulness for participation ($\beta = 0.54$). In terms of educational qualification (Base level: Abitur), participants with educational degrees lower than Abitur indicated higher scores of motivation ($\beta = 0.56$), trust ($\beta = 0.33$), and usefulness for participation ($\beta = 0.60$) in model A, together with lower information overload ($\beta = -0.911$). Participants with a higher educational degree than a bachelor, on the other hand, trusted model D significantly more ($\beta = 0.43$). The remaining category of other or unspecified educational qualifications also showed significant differences, but these cannot be interpreted further. Regarding the participant's age, significant effects could be found, especially for model C. For this model, older participants tend to exhibit less motivation and trust, rating it less useful to support participation ($-0.03 < \beta < -0.01$). Furthermore, for model A, older users reported lower information loads than younger users ($\beta = -0.02$). In terms of local proximity to the e-participation location, residents and ex-residents of Munich exhibited lower levels of trust for model B ($\beta = -0.47$) than other participants. Finally, for all models, participants who didn't speak German as a native language reported higher information overloads ($0.45 < \beta < 1.00$), and for model D lower levels of trust ($\beta = -0.29$).

### 6.2.5. Discussion and Conclusion

3D visualizations have become a substantial element in facilitating e-participation in urban planning processes. Despite their prevalence, the academic discourse is yet lacking unified reporting and rigorous evaluations (Eilola et al., 2023). As we observe considerable heterogeneity in the design and implementation of 3D models in practice, it is crucial to understand how their characteristics influence e-participation processes and their participants. Our qualitative study gives insights into experts' perspectives on (dis-) advantages of abstract versus detailed visualizations, while our online experiment with citizens additionally distinguishes detailed 3D city models in terms of their realism and resolution. Overall, many of our findings are consistent with related literature in the field. Both the expert interviews and the online study suggest an advantage of more detailed 3D visualization to facilitate participation. In the experiment, the highly detailed city model based on the 360° images, shows clear superiority in terms of enhancing participants' motivation and trust (Hypothesis 1a, Hypothesis 1b), especially for people with lower degrees of education. Although the model also evoked the highest information overload (Hypothesis 1c), load reports generally remained very low, so that the model was still perceived as most useful across all participatory levels. However, our nuanced consideration of high-detail models suggests that the level of detail is not the only decisive component. Especially, the photogrammetric mesh model clearly falls behind the other detailed visualizations, failing to evoke motivation (Hypothesis 3a) or trust (Hypothesis 3b). Not being considered useful for any participatory level, the model scores mostly equally bad as the low-detail model. The procedural model, on the other hand, scored positive values in almost all categories, although it still remains significantly behind the motivational and trustful capacities of the high realism model (Hypothesis 3a, Hypothesis 3b), even more so for actual residents of the urban planning area. Contradicting our hypotheses (Hypothesis 2c, Hypothesis 3c), both the procedural and the mesh models are perceived as less overloading than the 360° image model. As effects are only moderate, a potential explanation for this might be that the 360° image model included additional visualizations of people, which is a limitation of the study design and could have had an effect on the perceived load of information.

As prior research, the qualitative interviews highlighted some substantial challenges emerging from using high-detailed models, thus proposing more abstract models for certain participation stages and scenarios. Our online study, however, proposes mixed findings in this regard. While we see that the LOD2 model is selected more often as suitable for unconcrete planning scenarios than, e.g., the 360° image model and vice versa, we only find limited support for our hypothesis 2, as the LOD model seems to score equally (un)useful across all participatory levels. Possible explanations for this might hint towards important limitations of the study. Our study utilizes an artificial participation scenario of a public space reconstruction and inquiries into the subjective perceptions of citizens on different models. Considering the low level of prior participation experience, it may have been too difficult for study participants to precisely assess the models' usefulness for different participatory levels, which is also reflected in the particularly high Cronbach's Alpha, suggesting redundancy of items (Tavakol and Dennick, 2011). Still, with our research we are able to realize important theoretical and practical contributions. With our online study, we contribute robust theoretical findings regarding the effects of 3D visualizations with varying levels of detail, realism, and resolution as perceived by a large and diverse set of participants.

These present an important basis for the design and evaluation of traditional, but especially emerging, immersive e-participation systems. For design researchers, insights from our qualitative study and our explorative research into the relations between 3D city models and desired features and scenarios can be an important and encouraging starting point for investigating potential requirement differences in the context of e-participation systems. While our research focuses on participatory urban planning, in light of the increasing relevance of extended reality applications and the metaverse, insights into perceptions of different user groups of 3D urban environments can be helpful beyond the area of e-participation. As a short-term, practical contribution, our research can guide participation initiators, facilitators, and urban planners in the question of what city model to (not) use for their participatory processes based on their available options and key objectives. As an information basis, it can help governmental agencies to prioritize developments or acquisitions, especially in the context of the hype surrounding new technologies. In the long term, we believe our insights should motivate governmental and public institutions to support the collection, preparation, and maintenance of more high-quality geodata, providing them accessible to all practitioners.

With our research contributions and limitations, we see several exciting opportunities for future work: The conduction of a field study would allow for important observations in a more realistic and potentially multi-staged urban planning process while considering mixed visualizations such as a detailed environment model with abstract planning features could hold promising extensions to our considerations.

Concluding with the work presented in this chapter, we provide a robust reference point for evaluating 3D city models for participatory urban planning while hoping to inspire further evaluation studies to continuously improve practices of e-participation.

# Part IV.

# Designing Assistance and Learning in Digital Involvement

**Abstract Part IV**

*Focusing on the critical importance of data literacy in datafied societies, Part IV, titled "Designing Assistance and Learning in Digital Involvement" examines methods and tools for enabling skills and fostering learning among participants of digital involvement processes. Simplifying data representations is a key strategy for inclusive civic participation in political processes. However, other contexts of digital involvement offer the opportunity—or even the responsibility—to create societal value beyond the scope of individual projects by supporting skill development and data literacy among participants. For instance, crowdsourcing platforms have an ethical obligation to assist workers in adapting to changing labor market conditions through education and training. Similarly, citizen science projects are envisioned as mutually beneficial initiatives, providing opportunities for accessible knowledge transfer and skill-building. Therefore, this part explores how assistance and learning can be designed to support data-literate citizens, through the conduct of three studies. The first study investigates learning approaches and their implementation in crowdwork platforms, examining both theoretical frameworks and practical applications. This research, conducted with Anna Soßdorf, Osman Kıvanç Kırıkçı, and Jonas Fegert, was presented at the "Internationale Tagung Wirtschaftsinformatik 2024" (Stein et al., 2024b). The second study focuses on citizen science, analyzing existing and potential opportunities for platforms to enhance participants' data literacy. Conducted in collaboration with Alicia Wittmer, Christof Weinhardt, and Jonas Fegert, this research was awarded "Outstanding Paper" at the 21st International Conference "e-Society" (Stein et al., 2023c) and subsequently expanded into a journal article in the IADIS International Journal on Computer Science and Information Systems (Stein et al., 2023a), which is presented here. The final study examines the design of a CA to support data exploration in citizen science, providing a practical tool to facilitate participant learning and engagement. This research, conducted with Timm Teubner and Stefan Morana, was published in the journal Electronic Markets (Stein et al., 2024c), expanding on the contents of my master thesis at the Technical University of Berlin. In this part, titles, tables, and figures from the original publications have been renamed, reformatted, and re-referenced to align with the structure of the dissertation. Additionally, chapter and section numbering were adjusted, and formatting, abbreviations, and references were standardized for consistency.*

# 7. Learning and Skill Development in Crowdwork

## 7.1. Introduction

In the digital age, crowdsourcing has emerged as a powerful means for organizing work and collaborative endeavors. It enables businesses to innovate products and tackle complex problems (Bayus, 2012; Brabham, 2008), contributes to the creation of research artifacts (Haklay, 2013a; Spasiano et al., 2021), and technological innovations (Altenried, 2020). Unlike traditional work arrangements, digital platforms shape the relationships between task providers and task seekers, thereby reducing the connections to and visibility of workers and their skills (Kittur et al., 2013; Zhao and Zhu, 2014). As a result, in the realm of crowdwork, personalized training, and ladders for continuous skill development, typically provided from the employer side, are mostly absent (Bigham et al., 2017; Suzuki et al., 2016). Instead, the responsibility of learning and upskilling predominantly rests upon the workers themselves (Barnes et al., 2015; Margaryan, 2016). With platforms being criticized for insufficient learning opportunities and support structures (Bigham et al., 2017; Margaryan, 2016, 2019; Schmidt, 2017), their design often even encourages crowdworkers to cheat or produce poor work results (Hedderich and Oulasvirta, 2024). As industries and work profiles are rapidly evolving through new technologies, particularly AI (Colombo et al., 2019), an effective and sustainable crowdsourcing approach is challenged. For example, while crowdworkers' data manipulations played a pivotal role in realizing developments in the AI sector (Altenried, 2020), there is now a looming possibility that these very algorithms may replace similar work (Gilardi et al., 2023). Creating pathways for continuous learning and developing data literacy skills becomes paramount to keep pace with changing industry demands (Schüller et al., 2019; Zarifhonarvar, 2023).

We, therefore, seek to shed light on the research question: *"How can crowdsourcing platforms support the learning and skill development of crowdworkers?"*. To provide theoretical and practical guidance, we break this overarching question down into several research steps. First, crowdwork subsumes various kinds of work, ranging from microtasks like image classification to complex work of online freelancers (Estellés-Arolas et al., 2015; Margaryan, 2019). We thus conduct a structured literature review to provide a conceptual mapping of the research on learning within this broad landscape. Secondly, we apply our theoretical findings in a case study on promoting digital and data literacy, examining the platform "Kaggle" by means of a structured artifact review and a preliminary user study. Kaggle, as the most successful data science competition platform, is attracting both economic and scientific attention in the context of crowdsourcing and AI/ machine learning (ML) (Huang et al., 2022; Tauchert et al., 2020). Unlike other examples of crowdsourcing, Kaggle is specifically reported to facilitate learning in the domain of data-related skills (Baba et al., 2018; Fouda, 2020; Li et al., 2022), making a review of its practical implementation of support mechanisms and user insights as a case study particularly exciting. Overall, we identify five distinct teaching and learning approaches in the literature, four of which are implemented on the Kaggle platform in the context of data literacy learning. By discussing the theoretical concepts and their practical usage in an individual platform, we provide guidance for researching and further developing crowdsourcing platforms as a means to enrich learning opportunities for crowdworkers.

## 7.2. Background

This chapter describes the foundations of our work, including the crowdsourcing concept, and, as a basis for our case study, conceptual background on data literacy (learning) and the Kaggle platform.

**Crowdsourcing** was first described by Howe (2006b) as "the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and general large) network of people in the form of an open call" (p.2). It can subsume collaborative or individual work and heterogeneous crowds, including amateurs, experts, volunteers, or companies (Estellés-Arolas et al., 2015; Howe, 2006b, 2008). The wide adoption of crowdsourcing has led to numerous participation formats subsumed under this term, and authors differentiate between two to five different types (Estellés-Arolas et al., 2015; Kuek et al., 2015; Prpić et al., 2015a; Schader et al., 2012). In general, crowdsourcing systems evolve around the three main components: task assigners, providing a task; work providers, responding to the tasks; and the crowdsourcing platform (Zhao and Zhu, 2014). When it comes to teaching and learning in crowdsourcing systems, different, partly conflicting needs of stakeholders exist. Crowdworkers personally benefit from acquiring new skills and knowledge; however, following training and education outside the paid engagement reduces their time to spend on the job and thus income (Martin et al., 2014). Organizations, on the other hand, are especially concerned about output quality and costs (Matsubara et al., 2021; Zhao and Zhu, 2014). While teaching mechanisms may imply additional costs (Matsubara et al., 2021), they can but don't necessarily have to improve crowdworkers' task performance (Wang et al., 2018). As such, many platforms exclude teaching and learning mechanisms and leave this task to the crowdworkers themselves (Margaryan, 2016, 2019; Schmidt, 2017).

**Data literacy** has emerged as a crucial aspect of digital competencies, encompassing the skills needed to responsibly and self-awarely manage digital information, communication, content creation, safety, and problem-solving (Ferrari, 2013). Developing out of information literacy, competencies of data literacy account for the increased digitalization and datafication (Schüller et al., 2019) being the "ability to read, work with, analyze, and argue with data as part of a larger inquiry process" (D'Ignazio and Bhargava, 2016, p.84). Schüller et al. (2019) summarized data literacy in six core competencies (establish data culture, provision data, analyze data, interpret results, interpret data, and derive actions), highlighting its relevance both on a personal and professional level for participation in society and the labor market.

However, skills associated with data literacy and their importance are subject to constant changes, which makes the investigation of ongoing learning opportunities in crowdwork particularly relevant (Fischer et al., 2020; Schüller et al., 2019).

**The platform Kaggle** describes a “crowdcasting” platform where participants enter competitions to solve a task or a problem for the reward (Estellés-Arolas et al., 2015). Kaggle competitions are focused on data science and ML and enable organizers to create a competition by uploading a specific data set alongside a data science task to be solved by the crowd. Due to the increasing relevance of AI and ML for firms, data science competition platforms have gained recognition as crowdsourcing platforms, with Kaggle as the most successful example being a point of orientation for their design and development (Huang et al., 2022; Tauchert et al., 2020). The platform serves as a case study anchor for researching various aspects of crowdwork, such as collaborations and discussions within the virtual community (Hin, 2020; Huang et al., 2022; Li et al., 2022), organizational drivers and success factors for crowdwork (Tauchert et al., 2020), and code documentation (Wang et al., 2021). For reviewing learning and support mechanisms, data science competition platforms, in contrast to other crowdsourcing examples, are particularly interesting as classroom activities indicate their positive impact on user learning (Baba et al., 2018; Chow, 2019; Polak and Cook, 2021). However, further research is required to evaluate their effectiveness for platform crowdworkers and to identify the specific subset of data literacy skills they support (Tauchert et al., 2020). Motivated by the platform’s relevance in the realm of crowdwork and data science, our case study focuses on Kaggle, investigating respective research gaps.

## 7.3. Methodological Approach

To answer the research question, a three-step mixed-method approach was chosen. The individual steps (1-3) will be reported in the following and are visualized in Figure 7.1.

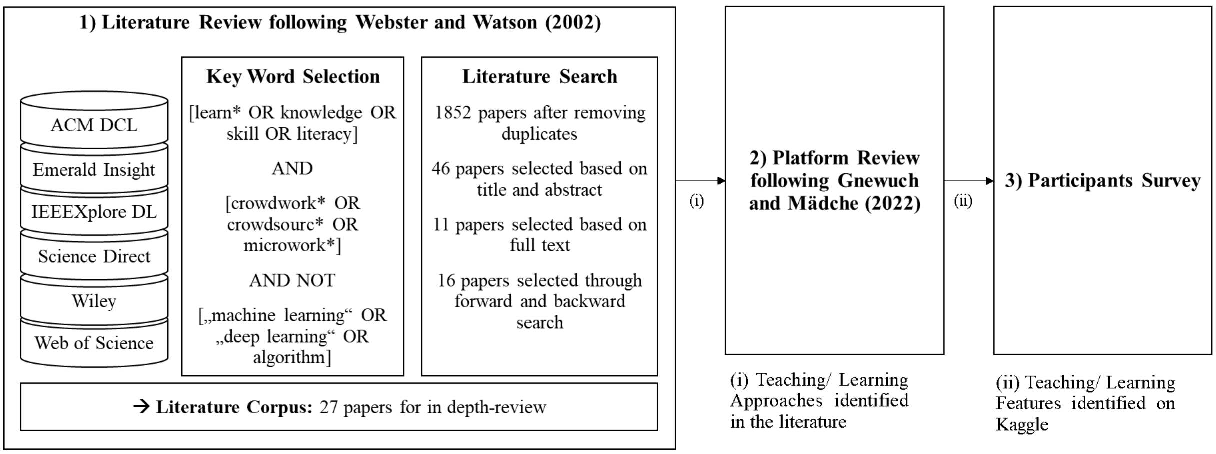


Figure 7.1.: Overview of the methodological approach.

**1) Literature Review:** First, we conducted a structured literature following Webster and Watson (2002). Due to the high interdisciplinarity of the combination of crowdsourcing and learning, we query several databases (see Figure 7.1) to match our search string in title, abstract, or keywords. We limited the search to publications from 2007 to 2022, taking the coining of the term crowdsourcing in 2006 as a starting point. The search string consisted of alternative descriptions for the term “learning” and “crowdsourcing”, excluding certain technical terms, to remove papers reporting solely on technical developments based on crowdsourcing efforts. After duplicate removal, the thematic fit was checked based on screening title, keywords, and abstract, and inclusion and exclusion criteria (Maturity of the publication, e.g., no early-stage drafts; reporting on crowdwork as defined by Howe (2006b), e.g., no communities of knowledge sharing; focus on the learning of crowdworkers, e.g., no crowdsourcing activity to create educational material) were applied on full texts. Finally, a forward and backward search was conducted, resulting in a literature corpus of 27 papers that were classified according to three dimensions (Teaching/ Learning Approaches, Skills/ Content, and Research Methods), following a bottom-up approach.

**2) Artifact Review:** Second, we conducted an artifact review of the Kaggle platform, following steps one, five, six, and seven of the structured artifact review by Gnewuch and Mädche (2022). Steps two, three, and four have been excluded due to our focus on one platform only. Based on our literature review, the objective of our practical artifact review of Kaggle was to determine which theoretical teaching/ learning approaches, are employed by the platform and through what features they are practically implemented (step 1). This objective determined the scope of our review and guided the data extraction that was conducted by two independent reviewers. The reviewers created user accounts on the platform, investigated platform elements, extracted features relevant to the review’s objective, and documented them in a spreadsheet format (step 5). To ensure the persistence of the findings, screenshots of the identified features have been archived in archive.today[1] (step 6). Finally, the results of both reviewers have been synthesized and compared to the findings from the literature review (step 7).

**3) User Survey:** Third, a first user survey with 69 participants was conducted as a pretest to contextualize review findings with the experience of Kaggle users. The survey included (a) items on the personal usage of Kaggle (categorical variables), (b) usage and helpfulness for learning of the identified platform features (binary variables), (c) items on the learning experience (5-Point-Likert-Scale), and (d) areas of data literacy learning (binary variables). Furthermore, we developed three items for measuring crowdworkers’ self-assessment[2] (c) to explore learning from a worker-centric perspective. Finally, we inquired crowdworkers’ perspectives on areas of data literacy learning (d), utilizing the framework by Schüller et al. (2019). The survey was distributed over the general Kaggle forum, as well as in-competition discussion forums in April and May 2023, targeting all active Kaggle users. The survey was based on voluntary and non-compensated participation. Statistical data analysis was conducted in R, utilizing Chi-Squared Tests and Spearman’s Rank Correlation Coefficient for correlation testing. For items (a), (b), and (c), all

[1] https://archive.ph/

[2] Self-assessment items: “How would you rate your overall learning experience on Kaggle?”, “How would you rate your ability to work with data before using Kaggle?” and “How would you rate your ability to work with data after using Kaggle?”

69 response sets have been included in the analysis, while for items (d), 51 response sets were included. This was due to inconsistencies in the responses of participants indicating usefulness for learning, without reporting on the feature usage before.

## 7.4. Results

The following chapter reports on the results of the three research steps, including the structured literature review, the Kaggle artifact review, and the Kaggle survey.

### 7.4.1. Results of the Literature Review

The results of the literature review are presented by dimension (1-3), while an overview can be found in Table 7.1.

| | **Dimensions** | | |
|---|---|---|---|
| | **Teaching/ Learning Approaches** | **Skills/ Content** | **Evaluation Methods** |
| **Characteristics** | a) Doing *(e.g., Kaggle Competitions: Active/ Getting Started, Open end, etc.)* | a) Digital Skills (Soft-/ Hardware-related) | a) Interviews |
| | b) Educational Material *(e.g., Kaggle Learning: Courses, Guides)* | b) Pattern Recognition | b) Survey |
| | c) Comparison/ Examples *(e.g., Kaggle Notebooks, Solution Leaderboards)* | c) Specific Topic Knowledge | c) Literature Review |
| | d) Giving/ Receiving Assessment | d) Reading and Writing Skills | d) Case Study |
| | e) Interaction *(e.g., Kaggle Discussion Forums: General, Topic-Specific)* | e) Creative/ Critical Thinking | e) Field Observation |
| | | | f) Controlled Experiment |

Table 7.1.: Overview of dimensions and their characteristics identified in the literature review. In italics, an overview of Kaggle's feature implementations identified in the platform review.

**1) Teaching/ Learning Approaches:** In terms of teaching/ learning approaches the literature was classified into five main research streams (see Table 7.1), presented in the following. As a first approach a), learning by doing is discussed in seven papers (Anya, 2015; Culbertson et al., 2017; Jackson et al., 2020; Jennett et al., 2016; Margaryan, 2016, 2019; Matsubara et al., 2021). The research indicates that learning by doing can follow a trial-and-error approach without specific platform features, but it can also be enhanced through purposeful task selection or ordering. Literature discusses features like alternating task difficulty or collaboration levels (Matsubara et al., 2021) and enriching tasks by assembling specific task stacks (Anya, 2015). Additionally, Jackson et al. (2020) highlight agent-centered features that allow participants to freely explore rather than guiding them. As a second approach b), ten papers discuss the importance of educational material (Becker et al., 2023; Bigham et al., 2017; Coetzee et al., 2015; De Leon Pereira et al., 2021; Dontcheva et al., 2014; Jackson et al., 2020; Jennett et al., 2016; Lee et al., 2016; Margaryan, 2016, 2019). Research reports that crowdworkers learn through self-study of external resources (e.g., Margaryan, 2019), while some papers investigate features to include educational material directly in the crowdwork platform. This can include interactive instructions and tutorials (e.g., Coetzee et al., 2015), learning sections (e.g., De Leon Pereira et al., 2021), or quizzes (e.g., Becker et al., 2023). In this context, De Leon Pereira et al. (2021) test the influence of differently presenting educational material (text, cartoon, or video); however, they do not find significant differences. Overall, the results of experimental studies show mixed results regarding the learning effect of providing educational material, including positive effects on task performance or content knowledge (e.g., Coetzee et al., 2015) and no or negative effects on accuracy and performance speed (e.g., Bigham et al., 2017). Additionally, (Jackson et al., 2020) report that especially novice users benefit from interacting with educational material. As a third approach c), a range of papers, investigate how crowdworkers learn from comparison and examples (Doroudi et al., 2016; Fantoni et al., 2012; Kobayashi et al., 2021; Mamykina et al., 2016; Margaryan, 2016, 2019; Matsubara et al., 2018; Nakayama et al., 2021; Singla et al., 2014; Zlabinger et al., 2020). Research by Margaryan (2016, 2019) shows that crowdworkers self-report learning from observing others and replicating what they see, while several papers investigate how platform features can induce this behavior. A common approach is to include task examples in the training or work interface (e.g., Zlabinger et al., 2020) or enable comparison after the execution of a task (e.g., Kobayashi et al., 2021) or at the end of a contest (Fantoni et al., 2012). In terms of the example or comparison type, some papers focus on the research of peer contributions (e.g., Fantoni et al., 2012), while others show expert examples (e.g., Doroudi et al., 2016) or AI-generated answers (e.g., Matsubara et al., 2018). Additionally, research investigates how examples and contributions for comparison should be selected and presented (e.g., Mamykina et al., 2016). In terms of the effectiveness of the strategy comparison and examples, for peer comparisons, the literature suggests that the quality of chosen examples is crucial (e.g., Kobayashi et al., 2021). The provision of expert comparisons outperforms peer comparisons in the case of Mamykina et al. (2016) and does not lead to negative effects in any of the studies. In Matsubara et al. (2018), expert comparison is outperformed by AI-provided examples; however, only for the case of learning, not accuracy. As a fourth approach d), six papers mention learning through receiving or giving an assessment of one's or others work (Doroudi et al., 2016; Dow et al., 2012; Margaryan, 2016, 2019; Wang et al., 2018; Zhu et al., 2014). Papers researching the impact of actively assessing contributions for learning focus either on the effect of self-assessment (e.g., Dow et al., 2012) or on the effect of reviewing a contribution from another worker (e.g., Doroudi et al., 2016). Receiving feedback on one's own work is self-reported to enable learning in the works of Margaryan (2016, 2019) and compared to self-assessment in the work of Dow et al. (2012). In terms of features to support this strategy, in most studies, the review process is guided by certain instructions or questions (e.g., Wang et al., 2018). Additionally, Zhu et al. (2014) researched whether organizing interactive group reviews improves the learning experience. Overall, the studies on giving or receiving assessments report mostly positive effects on dimensions such as quality, accuracy, and learning (e.g., Doroudi et al., 2016). Comparing different strategies, Wang et al. (2018) found significant differences between designs of guided self-assessment and between self-assessment and assessing others' work, while Dow et al. (2012) find no significant differences between receiving expert assessment and undertaking self-assessment. Likewise, Zhu et al. (2014) found that undertaking reviews in groups did not significantly improve learning compared to the individual scenario. Finally, e) nine papers investigate the effects of interaction (Becker et al., 2023; Chiang et al., 2018; Coetzee et al., 2015; Jackson et al., 2020; Jennett et al., 2016; Margaryan, 2016, 2019; Suzuki et al., 2016; Ye and Jensen, 2022). Learning through interaction is frequently mentioned by crowdworkers as a learning opportunity (e.g., Jennett et al., 2016) and can be organized differently on platforms. On the one hand, studies investigate how enabling coaching through experts or peers can be integrated into platforms to help people learn. For instance, in the work of Chiang et al. (2018), this is facilitated through exchanging coaching snippets between the crowdworkers, while the platform of Suzuki et al. (2016) brings together mentors and novices via text or video chat. On the other hand, research shows how learning is supported through informal connections in forums (e.g., Becker et al., 2023), chats (Coetzee et al., 2015) or comment options (Jackson et al., 2020). Concerning the effectiveness of interaction strategies, results vary, from no coaching effects on quality or accuracy (e.g., Suzuki et al., 2016) and incidental learning forums (Becker et al., 2023) to significant time improvements through coaching (Chiang et al., 2018) and positive effect of forum activities on chances of winning a contest (Ye and Jensen, 2022). Additionally, Jackson et al. (2020) find that forms of communal learning are especially beneficial in uncertain situations and for non-novice crowdworkers, while Coetzee et al. (2015) revealed mixed effects of chats depending on their purpose (task vs. educational material).

**2) Skills/Content:** With regard to the tasks examined and, consequently, the content and skills to be learned, five main categories were identified (see Table 7.1). First of all, a) a number of papers evolve around learning digital skills. This included a subgroup of papers investigating soft-/ hardware-related skills (Becker et al., 2023; Dontcheva et al., 2014; Doroudi et al., 2016; Suzuki et al., 2016) such as using or developing soft-/ hardware (Dontcheva et al., 2014), web search and development (e.g., Doroudi et al., 2016), or soft-/ hardware testing (Becker et al., 2023). Additionally, a subgroup focuses on labeling or classifying data and images (De Leon Pereira et al., 2021; Jackson et al., 2020; Kobayashi et al., 2021; Lee et al., 2016; Mamykina et al., 2016; Matsubara et al., 2021, 2018; Nakayama et al., 2021; Singla et al., 2014), which can lead on the one hand to pattern recognition skills b), but also the acquiring of domain knowledge c) (Jennett et al., 2016). Furthermore, many papers are centered on reading or writing skills d) (Bigham et al., 2017; Chiang et al., 2018; Culbertson et al., 2017; Dow et al., 2012; Matsubara et al., 2021; Wang et al., 2018; Zhu et al., 2014; Zlabinger et al., 2020). This includes writing, editing, analyzing, or summarizing documents (e.g., Dow et al., 2012), and transcription (e.g., Bigham et al.,

2017) or language learning (Culbertson et al., 2017). Finally, three papers evolve around tasks involving creative or critical thinking skills e) (Coetzee et al., 2015; Fantoni et al., 2012; Zhu et al., 2014). The remaining papers did not specify concrete topics or skills (Anya, 2015; Ye and Jensen, 2022) or clustered broader areas of skill improvement that were not linked to a specific task (Jennett et al., 2016; Margaryan, 2016, 2019).

**3) Methods/ Variables:** In terms of methods and variables, the investigated papers follow six approaches to research crowdworkers' learning. Twelve papers focus on the (self-)assessment of crowdworkers and organizers by conducting a) interviews (Becker et al., 2023; De Leon Pereira et al., 2021; Dontcheva et al., 2014; Jackson et al., 2020; Jennett et al., 2016) or b) surveys (Chiang et al., 2018; Culbertson et al., 2017; Doroudi et al., 2016; Lee et al., 2016; Margaryan, 2016, 2019; Suzuki et al., 2016). Additionally, one paper draws on c) a literature review (Anya, 2015), while another one conducted d) a case study demonstration (Fantoni et al., 2012). The majority of papers, however, draw insights from experiments and field observations. This includes five papers conducting e) field observations on platforms in normal operating mode (Becker et al., 2023; Chiang et al., 2018; Dontcheva et al., 2014; Jackson et al., 2020; Ye and Jensen, 2022) and nineteen papers conducting f) controlled experiments (Bigham et al., 2017; Chiang et al., 2018; Coetzee et al., 2015; Culbertson et al., 2017; De Leon Pereira et al., 2021; Doroudi et al., 2016; Dow et al., 2012; Kobayashi et al., 2021; Lee et al., 2016; Mamykina et al., 2016; Matsubara et al., 2021, 2018; Nakayama et al., 2021; Nguyen et al., 2020; Singla et al., 2014; Suzuki et al., 2016; Wang et al., 2018; Zhu et al., 2014; Zlabinger et al., 2020). Variables used to capture learning in the majority of the experiments focus on measures of task performance and quality, such as time spent on a task, accuracy, or number of corrections (e.g., Bigham et al., 2017). A small fraction of three papers used knowledge tests to measure learning (e.g., De Leon Pereira et al., 2021).

### 7.4.2. Results of the Artifact Review

Presenting the results of the Kaggle review, we contextualize Kaggle's features within the identified teaching/ learning approaches from the literature, summarized in Table 7.1 (see italic text). The centerpiece of Kaggle is its "Competitions", where data science challenges are crowdsourced. Competitions enable crowdworkers to actively work on data science problems, either alone or in teams, facilitating learning by doing. Kaggle offers both paid competitions, and non-paid competitions for beginners, practice, or open-ended exploration. Thus, Kaggle provides purposefully selected tasks to support learning and agent-centered features for self-exploration. Post-competition, the leaderboard allows participants to share their solutions, facilitating comparison and further learning. Within each competition, two additional features can be found being notebooks and discussion forums. Notebooks can be uploaded by peers and present code examples to peers. Additionally, Kaggle experts upload notebooks that serve as training examples. These notebooks can be commented on, copied, and individually run or edited and also be shared with specific crowdworkers to facilitate collaboration. Notebooks can thus be organized in the category of comparison/ examples while also supporting interaction. The discussion forum enables crowdworkers to open up a discussion on topics of the challenge, falling into the category of learning through interaction. Occasionally, the discussion forum is used to share educational material or request assessments from others. Notebooks and discussion forums are featured outside competitions as well. In “Datasets”, crowdworkers can upload their datasets and code examples as notebooks, with a discussion forum enabling interaction on the dataset topic. Datasets can be labeled according to their learning objective. Similarly, “Models” feature notebooks and discussion forums centered around specific data science models. In the “Code” section, crowdworkers can search for specific notebooks independent of a competition, dataset, or model. The “Discussion” section makes all discussions across Kaggle available and searchable. In addition to the discussion forums on competitions, datasets, or models, general discussion forums enable broader exchange on different topics. A section that is specifically intended to help crowdworkers learn is the section “Learn”. In “Learn”, crowdworkers can find educational material in the form of courses and guides. Courses are prepared by Kaggle and consist of tutorial lessons and matching training exercises. When taken successfully, crowdworkers receive a certificate for the course. Guides are curated by Kaggle; however, they present a mixture of materials on a specific topic, such as competitions, notebooks, or external material. While Kaggle provides functionality for four of the five learning approaches from the literature, functionality for giving/ receiving assessments and coaching is little present on the platform. Although crowdworkers can use interaction options to ask for an assessment of their work or comment on the work of somebody else, no explicit functionality for self-/ expert-assessment or review of others’ work is included. Nevertheless, Kaggle motivates asking for help and sharing knowledge by providing medals, e.g., for active participation in forums or notebooks. Likewise, no explicit functionality for coaching has yet been implemented. However, Kaggle announces the “KaggleX Mentorship”, a program that pairs novice crowdworkers with more advanced mentors.

### 7.4.3. Results of the User Survey

Through the Kaggle user survey, a diverse set of Kaggle users could be reached in terms of their membership duration (< two years: 46.4%, between 2-5 years: 36.2%, 5-10 years: 13%, > 10 years: 4.35%) and activity on the platform (< twice per month: 14.5%, 2-4 times a month: 31.9%, > 5 times per month: 53.6%). In terms of participation in competitions, most participants regularly compete in Kaggle competitions (< twice a year: 21.7%, about twice a year: 34.8%, > twice a year: 18.8%); however, there is also a considerable fraction of participants (24.6%) that never participated. Survey participants generally indicated good learning experience, with an overall learning experience between “Neutral” and “Very Good” ($Mean = 4.38$, $SD = 0.64$). Participants rated their data literacy skills prior to working with Kaggle on average 2.83 ($SD = 1.10$), and their abilities after using Kaggle on average 4.26 ($SD = 0.66$). Both the overall learning experience and the skill improvement are positively correlated with their involvement in competitions ($\rho = 0.31$, $p < 0.05$; $\rho = 0.35$, $p < 0.005$). Regarding the areas of learning, significant differences in the frequency of mentions by the participants could be found ($\chi^2 = 71.87$, $p < 0.001$). While for all areas of data literacy, learning has been indicated by at least 20% of the participants (21.74%

for "Aspects of data-driven action"), "Aspects of data analysis" (82.61%), and "Aspects of data product interpretation" (63.77%) have been significantly mentioned more than others. In terms of features used by participants, all inquired features have been used by at least 49% of the survey participants (49.02% both for in-competition forums and leaderboard solutions). Apart from the general Kaggle forum, which was primarily used for contacting survey participants, there were no significant differences in usage between the features ($\chi^2 = 5.30$, $p > 0.15$). Both the usage of the in-competition discussion forum and the leader-board solutions were correlated with the competition participation frequency ($p < 0.05$ and $p < 0.0005$), respectively. In terms of learning capabilities, the Kaggle learning section was most frequently indicated to contribute to learning by its users (92.86%), followed by the in-competition forum and the leaderboard solutions (both 80.00%). Although rated the least helpful for learning, 71.11%, and 68.57% of users that used the General Discussion Forum and the Notebooks, respectively, indicated their importance for learning. Differences in the learning capabilities of the features were not statistically significant ($\chi^2 = 6.55$, $p = 0.161$).

## 7.5. Discussion and Conclusion

In the following, we will discuss our study's results before highlighting its main contributions and limitations and concluding with an outlook for further research. Through a comprehensive literature review, we revealed that a considerable academic corpus exists that can serve as inspiration for the design and evaluation of learning opportunities in crowdwork. We find that five distinct teaching/ learning approaches are implemented through various platform features and span different types and contents of crowdsourcing ranging from microtasks (e.g., image tagging) to professional skills (e.g., Photoshop). This suggests that the creation of learning opportunities is not contingent on a specific task type and that no approach is exclusive to a particular line of work. On the one hand, the discussed platform features vary in implementation complexity and required human resources, allowing platform operators to align approaches with their resources. On the other hand, they differ in their degree of integration into (paid) crowdsourced tasks, suggesting that not all strategies are equally attractive to or adopted by crowdworkers in real-life settings. The review of evaluation measures shows a substantial emphasis on controlled experiments and prototype effectiveness, with a focus on performance metrics. Thus, further studies on operational crowdsourcing platforms, including workers' experiences, are necessary to obtain a more realistic picture of current learning support for crowdworkers. Our case study on Kaggle presents and discusses one such study. The structured artifact review highlights Kaggle's extensive set of learning features employing nearly all teaching/ learning approaches identified in the literature. Notably, many features build on participants' collaboration, such as sharing questions and solutions in forums or through notebooks, with few involving professionals or employees for input. In light of Kaggle's competitive nature, this community focus is striking but aligns with other research highlighting trends toward more team-oriented participation (Huang et al., 2022). In this context, Kaggle's complex incentive scheme, including performance tiers for community participation (Huang et al., 2022), could play an important role in the efficacy of the features.

With our preliminary user survey, we obtain a first impression of the learning experiences of Kaggle users. Although the survey cannot provide concluding insights, it suggests an active usage of all identified learning features and convincing learning experiences by the participants. In addition, it provides initial observations on the potential teaching capabilities of the individual features and on possible interactions between learning and individual characteristics of the crowdworkers. In line with Tauchert et al. (2020), participants' impressions in our survey indicate that learning on the platform primarily focuses on data analysis and interpretation while only touching upon other data literacy competencies. To support crowdworkers in professionalizing as data scientists, the platform may thus need additional exploration of ways to integrate other competencies into the learning offerings.

Our study provides several theoretical and practical contributions: As a first theoretical contribution, we present a synthesis and mapping of the current research state on crowdworkers' learning that provides insights into the broad scope of potential opportunities and enables structural comparisons. Furthermore, it allows for the identification of research gaps and can inform the evaluation of existing platforms, as put to use in our case study. Secondly, we generated structured knowledge on the specific approaches Kaggle implements to enable crowdworkers' learning by intersecting our artifact review with the results from the literature review. This knowledge can be used to analyze learning behaviors on the platform further. We started a consecutive analysis with an initial user survey that enabled us to contextualize our findings while pointing toward areas for further research. Besides the theoretical insights from the literature mapping, our review of an operating platform example serves as an accessible point of comparison for platform operators to get practical guidance on opportunities to design and evaluate learning settings for crowdworkers. For Kaggle itself, we identified limitations in terms of features and content that can inform the platform's own improvement. Furthermore, our insights can help platform operators and crowdworkers adopt new learning strategies also within the frame of current functionalities.

When interpreting our findings, it is essential to acknowledge certain limitations. In our literature review, constructing an all-encompassing search string accommodating the diverse terminology for "crowdsourcing" and "learning" in interdisciplinary contexts proved challenging. Instead, we opted for a more focused search string and supplemented it with a comprehensive forward and backward search. However, it remains a possibility that contributions from disciplines with distinct naming conventions or relevant papers focused on ML content were inadvertently overlooked. For the structured artifact review, despite engaging two independent reviewers, not all facets of Kaggle might have been fully identified, particularly as our review was confined to a limited number of competitions. Furthermore, our initial user survey can only present a starting point for thorough investigations of Kaggle from a user perspective due to its constrained nature. The survey did not strive for representativeness and was short in nature, precluding in-depth exploration of user characteristics such as age, profession, motivations, and detailed usage behavior of different features. The design was adapted to the authors' limited access to Kaggle's user base;

however, for a more comprehensive study, collaboration with platform operators would be necessary. Finally, when transferring insights from our review and survey, it must be considered that Kaggle cannot be presumed to be a crowdsourcing platform where participants heavily depend on platform-generated income. This could significantly impact the adoption and use of features for learning, as some participants may engage with the platform purely out of interest or a desire to acquire or test new skills, and as such, especially incentive schemes not explored in our research must be carefully considered.

Based on our contributions and limitations, multiple streams for further research arise. For the use case Kaggle, we envision a more in-depth analysis of the crowdworkers' behavior and experiences on the platform, e.g., through a survey with a larger user sample or qualitative interviews with crowdworkers. Moreover, we hope to provoke further research on learning opportunities and their implementation on other operating crowdwork systems, especially from a workers' perspective. In this context, investigating how crowdworkers' character traits and preferences might interfere with the effectiveness of different strategies would be an important perspective. Finally, we also see strong potential to investigate how insights from learning in crowdsourcing could produce spillover effects to other forms of digital participation, enabling learning opportunities for the general public. With increasing popularity of formats like citizen science or e-participation that share conceptual overlap with the design of crowdsourcing platforms (Stein et al., 2023b), research findings should be shared to empower citizens in their digital participation through different involvement concepts.

# 8. Promoting Data Literacy on Citizen Science Platforms

## 8.1. Introduction

In the digitized world, the ability to navigate in a data-based environment is a prerequisite for taking part in societal discussions and decision-making (Bhargava et al., 2015; Debruyne et al., 2021; D'Ignazio, 2017). The discourse on the Covid-19 pandemic, involving statistical metrics like the R-value, infection rates, and occupancy rates, exemplifies the importance of data literacy as a crucial skill, not only for professionals but also for the general public (Debruyne et al., 2021; Schüller et al., 2019). However, many citizens are missing basic data skills without the opportunities for continuing education in adulthood (Bhargava et al., 2016; Debruyne et al., 2021), motivating the exploration of creative strategies to involve them in data processes (D'Ignazio, 2017). While in several countries the public interest in science and research has increased during the pandemic (Jensen et al., 2021), the field of citizen science, as a collaborative approach to undertaking scientific research, continues to gain increasing attention (Vohland et al., 2021b). Citizen science can involve non-professionals in different research activities, from the development of research questions to the collection or analysis of research data (Shirk et al., 2012). Citizens have contributed to scientific achievements in a variety of disciplines, such as collecting and sharing geophysical data for earth observation research, or engaging in collective problem-solving and symptom or treatment surveillance in biomedicine (Shirk and Bonney, 2018). While the academic discourse recognizes the ability of citizen science to generate unique insights and power up research workforce (Shirk and Bonney, 2018), it also discusses its potential to improve participants' knowledge and skills, such as scientific literacy or domain knowledge (Jennett et al., 2016; National Academies of Sciences, Engineering and Education et al., 2018). In most forms, citizen science includes the engagement of citizens with data (Bowser et al., 2020; National Academies of Sciences, Engineering and Education et al., 2018), which makes it an interesting use case in the debate on promoting data literacy. In the citizen science literature, it has been shown that learning opportunities depend on a multitude of project characteristics (National Academies of Sciences, Engineering and Education et al., 2018), which complicates broad claims for learning and transformative effects within the variety of citizen science (Bela et al., 2016; De Albuquerque and Almeida, 2020). In this context, digital citizen science, also referred to as online or virtual citizen science, involving citizens by means of ICT, is interesting for multiple reasons. On the one hand, it allows for large-scale participation and, thus, potentially the learning experience of a multitude of hobby citizen scientists (Aristeidou and Herodotou, 2020; Jennett et al., 2016). On the other hand, lately, a shift from individual to generic infrastructure can be observed (Baudry et al., 2022), with some platforms evolving into integrated platforms for the conduction of generic projects (Liu et al., 2021). These multi-project platforms standardize the way digital citizen science projects are designed and conducted (Baudry et al., 2022; Liu et al., 2021), what in turn determines certain project outcomes such as democratic impact or learning (Bela et al., 2016; De Albuquerque and Almeida, 2020). As such, rather than evaluating opportunities for learning and applying data literacy in individual projects, the functionality and design of multi-project citizen science platforms can be an alternative starting point for exploring and shaping potentials and obstacles by design. Thus, we define our research question as follows: *What are the potentials and challenges for promoting citizens' data literacy through digital citizen science?*

Research question 1: What multi-project citizen science platforms currently exist?

Research question 2: How do current multi-project citizen science platforms support the application and learning of data literacy in citizen science projects?

Research question 3: What potentials and challenges do researchers and citizens identify to further develop multi-project citizen science platforms?

To answer our research questions, we draw on the conduction of two independent qualitative studies. Through a structured artifact review, we assess the status quo of the platform landscape, providing detailed insights into the functionality and design of 16 multi-project citizen science platforms. Additionally, we draw from a qualitative interview study with citizens, researchers, and technical experts that elaborates on their perspective on the utilization and design of multi-project citizen science platforms. By combining the results of both studies, we are able to provide a comprehensive assessment of the potential and challenges for citizens to apply and learn data literacy on digital citizen science platforms from the angle of platform design.

## 8.2. Related Work

In this chapter, the concepts of data literacy and (digital) citizen science are introduced and brought together to provide a theoretical basis for investigating data literacy promotion in digital citizen science.

### 8.2.1. Concepts

**Data Literacy** is a wide-ranging concept that is difficult to define and separate from other literacies such as digital or statistical literacy (Bhargava et al., 2015; Gould, 2021; Schüller et al., 2019). It can be described as an objective set of skills, such as the "ability to read, work with, analyze, and argue with data as part of a larger inquiry process" (D'Ignazio and Bhargava, 2016, p.84), but also more generally as the individuals' empowerment to navigate and engage in their data-based environment and society (Bhargava et al., 2015; Schüller et al., 2019). Data, as collected or generated information used to infer about different phenomena (Wise, 2020), is situated in an ecosystem that includes organizational elements such as its producers and consumers or the infrastructure and tools used (Bhargava et al., 2015). Data literacy, therefore, encompasses different competencies (Debruyne et al., 2021; Pedersen and Caviglia, 2019; Schüller et al., 2019), that can be modeled around the pyramid model of data value creation. To guide the education of data literacy, Schüller et al. (2019) developed a framework that describes relevant skills allocated to six core competencies (establish a data culture, provision data, analyze data, interpret results, interpret data, and derive actions). Furthermore, multiple educators have defined best practices for teaching data literacy, emphasizing more than just skill acquisition (Bhargava et al., 2016; Debruyne et al., 2021; D'Ignazio, 2017; D'Ignazio and Bhargava, 2016; Ridsdale et al., 2015; Wise, 2020). They place value on contextualization, such as drawing attention to the relationship between data and its production context (Wise, 2020), working with real-world community data (D'Ignazio, 2017; Ridsdale et al., 2015), or empowering learners to apply skills in their own contexts (Bhargava et al., 2015). Learning in interdisciplinary problems (Pedersen and Caviglia, 2019) and end-to-end data processes are encouraged to explore the impact of inquiry goals or stakeholder interests (D'Ignazio, 2017). Additionally, effective learning tools should be focused, guided, and inviting (D'Ignazio and Bhargava, 2016). Given its complexity, data literacy cannot be taught in a single initiative (Debruyne et al., 2021), so a modular approach should be chosen (Ridsdale et al., 2015).

**Citizen Science** describes an umbrella term for civic engagement in research that is yet lacking a uniform definition (Haklay et al., 2021). The Socientize project characterizes citizen science as "the general public engagement in scientific research activities when citizens actively contribute to science either with their intellectual effort or surrounding knowledge or with their tools and resources" (Socientize Consortium, 2014). While citizen science has been used originally, especially in natural science, nowadays it is a useful means to conduct research in various research domains (Levy and Germonprez, 2017; Pettibone et al., 2017). With the expansion of citizen science into a variety of fields of application, the heterogeneity of approaches has increased (Spasiano et al., 2021). The potential involvement can cover different parts of the research process, from the creation of the research question and hypotheses through data collection to analysis and publication (Shirk et al., 2012). This participation can be organized through citizen science platforms, describing web-based infrastructures that are used to support citizen science initiatives (Liu et al., 2021). Platforms can comprise a variety of infrastructures that either present citizen science activities, display project information, provide material, offer tools, or a combination of the aforementioned (Liu et al., 2021). The emergence of multi-project citizen science platforms offers project initiators the opportunity to set up and conduct their project with its distinct goals while sharing digital infrastructure with other projects instead of developing their own (Baudry et al., 2022). They thus standardize functions, as well as aspects such as community management (Baudry et al., 2022), which makes their design relevant for achieving results such as learning outcomes (Bela et al., 2016; De Albuquerque and Almeida, 2020; National Academies of Sciences, Engineering and Education et al., 2018).

### 8.2.2. Digital Citizen Science for Data Literacy Promotion

Educational initiatives can be assessed within four levels of the Kirkpatrick Model, being the learner's response to the initiative, the learning, the performance, and the result (Horton, 2001). While the first level investigates whether a learner generally enjoys participation, the second and third levels focus on the knowledge and skills acquired that can be applied in the learning setting or transferred to real-life problems (Horton, 2001). The fourth level underlines the focus on the initial goals of the learning, which can be a certain business goal or an impact on the participants' attitudes (Horton, 2001; Schüller et al., 2019). In the case of promoting citizens' data literacy, it is therefore decisive whether participants enjoy their participation in a citizen science project and what data knowledge and skills they gain in participating in the project, which is usually not a learning setting but a real-life problem. Likewise, it should be assessed whether the initiative empowers them to navigate their own data-based environment, which can be seen as an overarching goal of data literacy for citizens (Bhargava et al., 2015; Schüller et al., 2019). In citizen science, a considerable corpus of literature focuses on evaluating learning or learning opportunities in offline or online projects: They often assess general motivations (Jennett et al., 2016; Phillips et al., 2018); however, instead of focusing on data literacy, mostly domain-specific learning, and digital or scientific literacy are evaluated (Aristeidou and Herodotou, 2020; Herodotou et al., 2021; Jennett et al., 2016; National Academies of Sciences, Engineering and Education et al., 2018; Phillips et al., 2018). Although these competencies overlap with data literacy (Bhargava et al., 2015), only some research explicitly focuses on the acquisition of data skills and potential opportunities (Bowser et al., 2020; Golumbic et al., 2020; Radchenko and Maksimenkova, 2016). Other literature investigates the projects' resulting effects and the empowerment of citizens (Bela et al., 2016; De Albuquerque and Almeida, 2020; Haklay, 2013b; Phillips et al., 2018). Overall, educational opportunities seem to be project-specific, depending on the content and design of the projects (Aristeidou and Herodotou, 2020; National Academies of Sciences, Engineering and Education et al., 2018). A different angle for analyzing citizen science projects can be the analysis of underlying digital infrastructure. Many best practices have been developed for citizen science platforms to optimize their design for participation (Jennett and Cox, 2014; Musto and Dahanayake, 2021; Newman et al., 2010; Skarlatidou et al., 2019; Sturm et al., 2018; Wald et al., 2016; Yadav and Darlington, 2016). Linking these principles to their relevance for learning and data literacy education allows us to assess platform designs according to their potential to promote citizens' data literacy. An overview of the identified design principles can be found in Table 8.1.

## 8.3. Methodology

Our research approach comprised the conduction of two independent qualitative studies with the aim of shedding light on both present and future potentials for promoting citizens' data literacy in digital citizen science from the perspective of platform design. First, we conducted a structured artifact review to identify and assess current multi-project citizen science platforms according to their functionality

<table>
<tr><th colspan="2">Platform Design Concept</th><th>Meaning for Data Literacy</th></tr>
<tr><td rowspan="5">Aesthetics and Usability</td><td>Provide simple and clear project main pages (Skarlatidou et al., 2019)</td><td rowspan="4">Provide an enjoyable education initiative (Horton, 2001)<br>Provide a focused and inviting tool (D'Ignazio and Bhargava, 2016)</td></tr>
<tr><td>Ease entry barriers (Jennett and Cox, 2014; Sturm et al., 2018)</td></tr>
<tr><td>Reduce information on the platform (Jennett and Cox, 2014; Skarlatidou et al., 2019)</td></tr>
<tr><td>Standardize naming and navigation (Skarlatidou et al., 2019; Sturm et al., 2018)</td></tr>
<tr><td>Communicate project goals (Jennett and Cox, 2014; Newman et al., 2010)</td><td>Create understanding for inquiry goals and data production context (D'Ignazio, 2017; Wise, 2020)</td></tr>
<tr><td rowspan="3">Data Standards</td><td>Validate user-generated data (Musto and Dahanayake, 2021; Skarlatidou et al., 2019)</td><td rowspan="3">Enable informal learning through the execution of tasks (Jennett et al., 2016)<br>Empower learners to practically apply their skills to real-world data (Bhargava et al., 2015; D'Ignazio, 2017; Ridsdale et al., 2015)</td></tr>
<tr><td>Facilitate entering user-generated data (Newman et al., 2010; Skarlatidou et al., 2019; Sturm et al., 2018)</td></tr>
<tr><td>Enable data analysis and visualization (Musto and Dahanayake, 2021; Newman et al., 2010; Skarlatidou et al., 2019; Wald et al., 2016)</td></tr>
<tr><td rowspan="3">Support</td><td>Provide separate support pages (Skarlatidou et al., 2019)</td><td rowspan="3">Enable formal learning (Jennett et al., 2016)<br>Provide a guided tool (D'Ignazio and Bhargava, 2016)</td></tr>
<tr><td>Provide educational material (Wald et al., 2016)</td></tr>
<tr><td>Provide interactive tutorials and information (Jennett and Cox, 2014; Skarlatidou et al., 2019)</td></tr>
<tr><td rowspan="2">Communication</td><td>Enable communication between participants (Newman et al., 2010; Skarlatidou et al., 2019; Sturm et al., 2018; Wald et al., 2016)</td><td>Enable informal learning through interaction (Jennett et al., 2016)</td></tr>
<tr><td>Enable communication between participants and researchers (Newman et al., 2010; Skarlatidou et al., 2019; Sturm et al., 2018; Yadav and Darlington, 2016)</td><td>Enable two-sided learning for an empowering effect (De Albuquerque and Almeida, 2020)</td></tr>
</table>

Table 8.1.: Key design dimensions identified in the digital citizen science literature.

and key dimensions identified in the literature. For the review, we followed a 7-step methodology for reviewing real-world software artifacts that aims to provide methodological guidance to the review in a systematic way, including problem formulation, artifact search and screening, assessment of practical quality, data extraction, artifact documentation and archiving, and data analysis (Gnewuch and Mädche, 2022). While steps 1-6 have been conducted by one reviewer, for step 7, two independent reviewers have been engaged in coding the documentation. For data extraction and coding, we followed best practices for structured-content analysis of web pages (Saraswat, 1999). Second, based on this depiction of the status quo, we aim to contextualize results with the interests and needs of citizen science stakeholders and uncover potential further opportunities or threats to promoting citizens' data literacy in digital citizen science. As such, we draw from the conduct of semi-structured interviews with citizens, researchers, and technical experts. For the interview preparation, conduct, transcription, and structured content analysis, we followed Kaiser's method for qualitative interviews (Kaiser, 2014). This included the design of an interview guideline and the conduct of a pretest, which was used to test and refine the interview guideline. The transcription and coding were undertaken by a single researcher with the tool MAXQDA, while in the analysis of the results two researchers were involved.

## 8.4. A Review of Citizen Science Platforms

In this chapter, we present the implementation and results of the structured artifact review.

### 8.4.1. Implementation

The structured artifact review was implemented from July to September 2022. It aims to identify the current functionality and design of digital platforms to derive their potential and challenges for promoting citizens' data literacy. Especially interesting to us are multi-project citizen science platforms, as they standardize opportunities for a multitude of projects (Baudry et al., 2022). Therefore, we defined active participation opportunities for citizens (a) and the possibility for generic project creation (b) as inclusion criteria to exclude other platform types (Brenton et al., 2018). In terms of assessing practical usability for the review, we defined as criteria availability in either English or German (c) and the possibility to review them free of charge (d), meaning either the platform must be free of charge or it must offer a demo or link to project examples that can be reviewed. For the software artifact search, we utilized three search directions being: overviews in the citizen science platform literature (Aristeidou and Herodotou, 2020; Brenton et al., 2018; Liu et al., 2021; Luna et al., 2018; Skarlatidou et al., 2019; Yadav and Darlington, 2016) (43 artifacts); the commercial database provider Crunchbase with the filter option 'Citizen Science' (14 additional artifacts); and the EU's and Austria's citizen science information webpages (18 additional artifacts). The initial sample of 75 artifacts included next to multi-project platforms, 32 project overview platforms, five community exchange hubs for educational material and workshops, and 12 single-project platforms. Additionally, six platforms presented several projects but did not allow for the creation of new generic initiatives. Thus, after screening for (a) and (b), 20 multi-project platforms were further reviewed for quality. Excluding four platforms due to language barriers or accessibility issues, 16 platforms for the final review were archived in Archive.Today and can be seen in Table 8.2. For data extraction and analysis, the platforms were assessed according to their participatory functionality based on the concept degree of participation by Shirk et al. (2012) and their design, based on the criteria in Table 8.1, and documented in a concept matrix.

| **Platform ID and Name** | | | |
|---|---|---|---|
| A | Biocollect-Atlas of Living Australia | I | Ispot |
| B | Citizen Science Center Zürich | J | Just One Giant Lab (JOGL) |
| C | CitSci | K | nQuire |
| D | conserve.Io* | L | Pybossa |
| E | CyberTracker | M | SciStarter |
| F | DataCertus | N | Spotteron |
| G | Epicollect5 | O | World Community Grid |
| H | Inaturalist | P | Zooniverse |

* The review of this platform is based on two freely viewable projects and thus limited.

Table 8.2.: List of multi-project platforms for the artifact review.

## 8.4.2. Results

In the following, the assessment results of the 16 multi-project citizen science platforms are reported in detail by dimension. Additionally, a brief overview of the results can be found in Figure 8.1.

| **Dim.** | **Platforms** | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **I** | **J** | **K** | **L** | **M** | **N** | **O** | **P** |
| **Participatory Functions (PF)**: computational power (PF1) data collection (PF2), data analysis (PF3), task assignment (PF4) research questions and approach (PF5) | | | | | | | | | | | | | | | | |
| PF | PF 2,3 | PF 2,3 | PF 2,3 | PF 2 | PF 2 | PF 2,3 | PF 2 | PF 2 | PF 2 | PF 4 | PF 5 | PF 2,3 | PF 2,3 | PF 2,3 | PF 1 | PF 2,3 |
| **Aesthetics & Usability (AU)**: similar project main pages (PA1) eased entry barriers (PA2) reduced information (PA3) easy naming conventions (PA4) search options (PA5) project goal communication on project page (PA6) | | | | | | | | | | | | | | | | |
| PA | AU 1-6 | AU 1-6 | AU 1-6 | AU 2,3 4,6 | AU 3-5 | AU 1, 3-5 | AU 1, 3-6 | AU 1-6 | AU 1-5 | AU 1-6 | AU 1-3 5,6 | AU 1,3 4,6 | AU 1-6 | AU 2-6 | AU 1,24 -6 | AU 1-6 |
| **Data Standards (DS)**: eased data entering (DS1) data analysis and visualization (DS2) data validation (DS3) | | | | | | | | | | | | | | | | |
| DS | DS 1,2 | DS 1,2 | DS 1,2 | DS 1,2 | DS 1,2 | DS 1,2 | DS 1,2 | DS 1-3 | DS 1-3 | - | DS 1,2 | DS 1,2 | - | DS 1-3 | DS2 | DS 1-3 |
| **Support (SU)**: support page (SI1) educational material (SI2) interactive information and tutorials (SI3) | | | | | | | | | | | | | | | | |
| SI | SI 1-3 | SI 3 | SI 1-3 | SI 1,2 | SI 2 | - | - | SI 1-3 | SI 1,2 | SI 1-3 | SI 3 | SI 1 | SI 1-3 | SI 1,3 | SI 1-3 | SI 1-3 |
| **Communication (CO)**: communication between participants (CO1), communication between participants and researchers (CO2) | | | | | | | | | | | | | | | | |
| CO | CO 2 | CO 2 | CO 1,2 | - | - | CO 1,2 | CO 1,2 | CO 1,2 | CO 1,2 | CO 1,2 | CO 1,2 | - | CO 1,2 | CO 1,2 | CO 1,2 | CO 1,2 |

Platforms that meet all criteria in one dimension are shaded

Figure 8.1.: Summary of the review results.

**1) Participatory functions:** To evaluate what aspects of data literacy citizens may apply and improve through their participation in digital citizen science platforms, projects are reviewed for their offer of participation activities. The lowest degree of participation was found for platform O, where citizens can only contribute to research projects through computational power. On the other hand, many platforms, focus on data collection and analysis: Four platforms (D, E, H, I) provide functionality to upload and classify observations, while one platform (F) allows for the upload of complete data sets. Other platforms extend the possibilities for data collection, e.g., to surveys (B, G, P), digital diaries (B), or different types of media (B, G, L). The opportunities for participatory data analysis range from functionalities for transcriptions (B, L, P), mapping or classification of images (B, L, P) to pattern recognition in sound or video material (L). Additionally, one platform offers no code toolboxes for machine learning algorithms (F). Four data collection and analysis platforms (A, C, M, N) stood out through their flexibility, which would potentially allow their usage for other participatory activities. This includes one platform (M) offering a range of third-party tool integrations, another platform (C) with customizable data sheets and possible integrations, a further platform (A) providing multiple individual applications, e.g., for evaluation or learning games, and a platform (N) with functionalities that can be customized on demand. Next to platforms for data collection and analysis, one platform (J) in the sample supports the participatory assignment of tasks, although tasks themselves are not conducted on the platform. This platform enables initiators to structure the project into several phases, including phases such as project conception or prototyping, with the offer of partly challenging tasks. Another platform (K) enables users to define research questions, design a methodological approach, and collect data; however, except for data collection, activities are not undertaken participatory.

**2) Aesthetics and usability:** The platforms' aesthetics and usability play an important role in who can access and enjoy participation in citizen science projects (Skarlatidou et al., 2019), as an important prerequisite to a learning initiative. As such, we reviewed five subdimensions identified in the literature. First of all, we reviewed whether platforms have a consistent concept for their project main pages. Most platforms follow the principle of deploying the same design for every project. Only two platforms (D, N) vary their designs throughout different projects, and one platform (E) has no project overview page by design. Five projects (B, H, N, O, P) give an overview of the projects' progress on the main page, while some platforms include information on contributing participants (B, C, E, H, J, M, N, P) or contributed data (A, C, D, E, G, H, I, K, L, N, P). Additionally, outstanding features that were noted were an overview of currently active participants in the project (I, N, P) or the proposition of similar projects (M). Second, concepts for easing entry barriers to platform usage were reviewed. For this subdimension, some platforms enable participation via the web browser (A, B, C, H, I, J, K, M, P), while others require the installation of software. Above that, three platforms (D, J, N) include either pop-up explanations upon registration or entering a project, and ten platforms provide example pictures or explanations for project tasks (A, B, D, I, J, K, M, N, O, P). Additionally, four platforms (A, J, M, P) indicate the level of difficulty for tasks or indicate the required skills for participating. As a third and fourth subdimension, concepts for reducing text and standardizing naming and navigation conventions were evaluated. To reduce the information load, all platforms utilize pictures or icons, except for platform O, which is primarily text-based. In particular, one platform (E) stood out by replacing texts comprehensively through icons. Additionally, seven platforms (A, H, J, L, M, N, P) use expandable and collapsible text to avoid information overload at first sight. In terms of naming conventions, most platforms try to avoid technical terms and follow easy naming conventions such as 'Project or Community', 'Add' or 'Contribute'. An exception to this was found in the platform J, which uses individual names e.g. 'Spaces', 'Needs', 'Programs' or 'Claps' and technical terms, for instance, in search options. For the platform's navigation, three platforms (B, C, F) allow searching for specific project names and one (E) for a specific web address. Besides that, three platforms enable users to search for names or categories (G, O, P) and three (D, K, L) platforms do not employ any search options. The majority of platforms (A, H, I, J, M, N) however, allow for a broad search according to different characteristics such as age, activity, or organization. As creating an understanding of inquiry goals is important for teaching data literacy, we lastly evaluated the communication of project goals, identifying two different approaches. Three platforms (E, F, I) include general goals on the platform's main homepage, while other platforms include an explanation of project-specific goals on the individual project pages.

**3) Data standards:** In terms of working with scientific data, platforms have been reviewed for their concepts to support entering, validating, visualizing, or analyzing data, as mentioned in the citizen science literature. For collecting project data, most platforms facilitate participation through specified entry masks that include, for example, drop-down menus, checkboxes, or drag-and-drop (A, B, C, D, E, F, G, H, I, K, L, N, P). Additionally, some platforms provide extra help for the classification of data (E, H, I, N, P) in form of common mistakes, help information, or automated proposals. Two platforms do not support data contributions (J, O) and one platform does not have standardized data entry options (M). In terms of data validation, only four platforms present validation concepts: This comprised checking whether all necessary entries were filled (P), flagging contributions through the community (N), a like and reputation system (I), or the differentiation between quality levels (H). For platform D, data validation could not be reviewed. Regarding the analysis of project data, some platforms have integrated data analysis tools (A, B, C, E, F, K, N, O, P), while others focus only on data visualization (D, G, H, I). Analysis tools include data tables (P), sampling options (A), statistics (C, N, P, M), or predictions (A, N). Furthermore, shared codebases for researchers (C, P), machine learning classifications (F), options to compare specific data points or groups (A, C, N), or to follow the total numbers of classifications or solved tasks (C, N, P, B, O), are included in platforms. In terms of visualizations, platforms offer scatter plots (A, N, P), histograms (C, N, P, D, E, G, F, K, M), map visualization (A, C, N, P, D, E, G, F, H, I, M) or plots visualizing data trends over time (C, N). One platform (L) allows for flexible analysis through code integrations without predefining tools automatically, and two platforms (J, M) do not enable data analysis. Besides project data, five projects allow the analysis of data about the projects' progress (B, H, N, O, P), and five projects analyze meta-data about the community (L, M, N, O, P).

**4) Support:** Next to informal learning, digital platforms were reviewed for three aspects of supporting formal learning. The first one comprised a concept for a separate support page. Most platforms have help pages for several topics, such as guidelines on how to build a project (A, B, C, E, F, G, H, I, J, L, M, N, P), or general support pages for citizens (A, D, H, I, J, L, N, O). Additionally, most platforms employ an FAQ page for citizens (A, C, H, I, J, L, M, N, O, P). Five platforms (B, E, F, G, K) do not provide any separate help pages for participants. As a second aspect, platforms were reviewed for their provision of educational material. Implementations of this feature comprised workshops (A, M, P), training sections (A, C, D, H, I, M, P), blog or newsletter articles (C, E, H, I, O, P), and community spaces for learning (J). Six platforms did not integrate educational resources by design, however, project-based integrations are still possible. Third, concepts for interactive tutorials or help information were reviewed. Design features that were employed in this category were pop-up tutorials leading through the platform (J, N, P), tutorials and videos on project pages (H, J, M, N, O, P), or additional information based on hovering over content (A, B, C, K, N, O). Six platforms do not include interactive help information.

**5) Communication:** As a central aspect of empowerment and learning through interaction, concepts for communication between participants, and between participants and researchers were reviewed. Five platforms (A, B, D, E, L) have no concept for a project-based debate of participants, while three platforms (D, E, L) additionally have no possibilities to contact researchers. Five platforms (H, J, M, N, P) enable exchange within the community and with researchers via direct chat options. Nine platforms (C, G, H, I, J, K, N, O, P), allow for communication via forums, either project-based or general forums. Other concepts for community exchange comprised messages on newsfeeds (J, N), comment options (H, I, J, K, M, N, P), or project-based question-and-answer options (M). For contact with researchers or project creators additionally, many projects (A, B, C, F, G, K) offered an e-mail option. One platform in the sample (N) especially stood out due to its broad range of options for community exchange.

## 8.5. Interviews on Potentials of Citizen Science Platforms

In the following chapter, we report on the conduct of a qualitative interview study and a subset of its results, which are used to investigate the research object. The interviews aimed to identify potential and challenges for engaging citizens in scientific research projects via digital platforms and were conducted earlier and independently of the artifact analysis. By revising aspects of the study relevant to our findings from the structured artifact review and the context of applying and learning data literacy, we can contextualize our findings about the current platform landscape and provide guidance for its further development.

### 8.5.1. Implementation

The interviews were implemented from April to June 2021 using a virtual meeting tool. They included 14 participants, who were between 16 and 64 years old and composed of six citizens (C1-C6) and eight professional researchers (R1-R8) out of whom two can be considered technical experts for platform design (R7, R8). We ensured equal gender representation and diversity in terms of prior knowledge regarding the participation in or realization of citizen science projects. The interviewed researchers had diverse backgrounds, such as medicine, sports science, or IS to cover different potential applications of citizen science. After a brief introduction to the topic of digital citizen science, the interviewees were consulted for their perspectives on the offer of different participatory activities according to Shirk et al. (2012), support needs to conduct these activities on a citizen science platform, challenges in working with scientific data, especially with regard to data quality according to Parsons and Lukyanenko (2011), and needs for communication and feedback. Additionally, the interview guideline included the topics of incentives, quality of participation, and aspects of anonymity, which are outside the scope of this chapter.

### 8.5.2. Results

The results of the structured content analysis are reported by dimension, with the respective interview source indicated in parentheses (Citizens C1-C6, and Researchers R1-R8, including two technical experts R7-R8).

**A) Participatory Functions:** To determine potentials for participation and engagement with scientific data, we asked the interviewees whether they could imagine themselves or citizens participating in asking research questions (A1), collecting information and resources (A2), developing hypotheses (A3), developing research methodologies (A4), collecting data (A5), analyzing and interpreting data (A6), and publishing results (A7) via a digital platform (Shirk et al., 2012). For every research step, at least half of the participants expressed interest in participating. While citizens could imagine their participation best in the step of developing hypotheses ($n = 6$), more challenging to them appeared the steps of data analysis and interpretation as well as publishing results ($n = 3$). Researchers could imagine citizens' participation best in developing research questions and data collection ($n = 8$) and the least in developing hypotheses and publishing results ($n = 4$). Obstacles on the side of participants were entitlement (C4, (A1)), adhering to formal structures and wording (C5, (A7)), or motivation and time (C1, C3, (A2); C1, C6 (A4); C5 (A5); C3 (A7)). Additionally, they claimed that they might miss important knowledge (C6, (A1); C4 (A4); C4 (A7)), or feared contributing incomplete or wrong contributions (C3, (A5); C5 (A6)). For researchers, major challenges included the utilization of domain literature (R4 (A1)), containing accessing, evaluating, and reading scientific contributions (R1, R2, R4, R8 (A2)). Additionally, they remarked on the challenge of adhering to formal structures and logical reasoning, thereby ensuring quality (R4 (A1); R2, R5, R6 (A3)), and of missing knowledge necessary for the conduct of tasks (R8 (A1); R4, R7 (A4); R5, R6, R8 (A6)). One researcher added the challenge of coordinating work while being exposed to deadlines (R8 (A7)). On the positive side, participants remarked that enabling participation in more challenging parts of the research process could be interesting and valuable for citizens and researchers (C1, C2, R4, R7 (A6)), generating potentially unconventional and interesting results (R5, R8 (A4), R2, R4, R5 (A6)), while being a great source to impart knowledge (R1, R8 (A4)) and give credit

(R5, R8 (A7)).

**B) Data Tasks:** In terms of working with scientific data, the interviewees were asked about their perception of challenges, especially concerning data quality. Researchers and citizens equally expressed concerns about the quality of data regarding intentional or unintentional manipulations by participants. Citizens could try to adapt their contributions to fit socially accepted perceptions (R2), their own ideas and expectations (C1, C3, C4), or to build up certain political topics (C4, R6). They could also lose interest in working conscientiously over time (R2, R4). The lack of control over whether participants adhere to scientific standards (R4, R5, R8) and the potential incompleteness of contributions (C2) are named as challenges for citizens contributing scientific data. In this context, two researchers noted that challenges might be dependent on the type of data and measurement (R1, R6). Additionally, quality issues might also comprise the scientific relevance of contributed ideas or data, such as generated research questions (R1, R6).

**C) Support:** In A) and B), citizens were additionally asked for needed support to overcome obstacles in participation and data engagement. Participants repeatedly stressed the importance of collaboration and exchange. This could include a direct point of contact in case of questions (C5, R3, R7), the organization of workshops (R1, R5, R6), or functionality for joint development of hypotheses or analyses both between participants and researchers but also the community itself (C1, C3, C5, C6, R4-R8). Furthermore, feedback was frequently mentioned as an important factor (C1, C3, C5, C6, R1, R3-R7) which could be linked to the possibility of reminding participants (R2). When it comes to giving feedback, the participants suggested that it could be either manual or automated (R1, R3) or elaborated through a collaborative review process (R3, R7); however, it should be personal (C3, R5) and transparent (R7). In order to support formal learning, participants ideated that platforms might include tutorial videos (C2-6, R1, R2, R4, R5, R8), short manuals and checklists (C1, C3, C4, R5), or best practices (R7). Materials could be collected in a database (R3, R8), potentially with a recommender system (C4, R3, R8) or a translator, thereby helping with complex scientific language (R8). Additionally, respective knowledge could be transferred by introducing a toolbox to explain, e.g., scientific methods and what they are for (C4, R1, R4) or providing templates (C2, R7). Participants also ideated on the use of chatbots for answering questions (R5, R7) or supporting certain research steps (R1, R3). To help with quality concerns, platforms could introduce control questions (R4), enable comparison with other data (C2, R8), or introduce the need to upload explanations and supporting documents for contributions (C1, C5). However, ethical education might also be important, such as stressing the importance of the validity of contributions and the neutrality of science (C1, C4, R1).

**D) Communication:** For the communication, participants were asked about several communication mechanisms (voting, discussion forum, comment option, chat function) and their suitability for using them on digital citizen science platforms. Overall, all participants showed great support for all different communication mechanisms, underscoring that communication means are indispensable to enable meaningful participation (C2, R4) and that it would be helpful to have everything integrated into a single platform (R5). One citizen highlighted that offering a variety of different communicative means would be important as participants have different preferences (C3), and two interviewees added that it would give important flexibility to the participants to contribute their opinions (C1, R8). The highest endorsement was given to the discussion forum, up-and-down voting, and chat functions. The up and down voting would allow for quick participation with minimal effort (C2, C5, R4, R5) that would help to structure content and enable easy overviews of the opinions and interests of citizens (R2, R4). However, it could also lead to polarization, group building, or the suppression of certain views (R6, R8). A discussion forum was seen as an important means to facilitate collaboration, and the collective creation of innovative ideas (R5, R8), while a direct chat function would improve the building of a community and mutual, quick exchange (R1, R2, R8). Additional participants mentioned video conferences (C1, C3, R4), and one participant noted that audio comments could enable freer participation (C1).

## 8.6. Discussion

Through the conduct of two qualitative studies, we aimed to explore the potential and challenges of digital citizen science to enable citizens to use and extend their data literacy. While the review of 16 multi-project citizen science platforms enables us to draw an extensive picture of the current digital citizen science landscape, the conduct of 14 expert interviews helps us contextualize these findings and elaborate on future directions for multi-project citizen science platforms. As such, we arrive at three main positions:

> **A)** Current digital citizen science projects provide opportunities to engage citizens in collecting and analyzing different types of data while largely neglecting other competence fields in data literacy, although they could potentially be more integrated.

Our review found that all but one platform offered participatory tasks where citizens could engage with data, with a focus on data collection and analysis. For these research steps, most platforms offered extensive functionality to facilitate data contribution or classification and to allow citizens to access some form of data analysis or visualization. Only mechanisms to automate data verification were not common across the platforms, which our interviews suggested could be a challenge given that feedback and quality control are seen as important factors in improving citizens' ability to contribute qualitative data. The review showed that engagement with different types and formats of data is encouraged in many projects and is not limited to research data but also data collected on the projects and its participants. Such engagement can lead to informal learning (Jennett et al., 2016) of data literacy, including aspects of data measurements, variability, and visualization (Kermish-Allen et al., 2019; National Academies of Sciences, Engineering and Education et al., 2018) or interpreting data products to derive recommendations (Golumbic et al., 2020). Data literacy, however, also encompasses skills such as identifying data use cases, measurable objects, and hypothesizing about their relationships (Schüller et al., 2019), which might not all be covered by contributory citizen science projects. To this end, platforms with community inquiry approaches might be promising (Herodotou et al., 2021), of which we could only find one example in our review. Likewise, no platform could be found that formally engaged citizens in final research steps such as interpreting and disseminating the results of the inquiry, which would allow for the application of data literacy aspects such as verbalizing data and data products or identifying data-driven actions (Schüller et al., 2019). From our qualitative interview study, both citizens and researchers believe that collaboration is feasible and interesting in other research activities as well. In particular, for developing research questions, hypotheses, and research methods, citizens showed great interest in collaborating with researchers. This could be an interesting starting point to expand opportunities for citizens to also apply and learn important aspects of data culture competencies, aspects of which were frequently mentioned in our interviews as challenges to participation and data quality.

> **B)** Current digital citizen science platforms are advanced in their aesthetic and user-friendly design while providing diversity for different target groups.

The review suggests that, while varying in their design most platforms provide concepts for different aspects of usability and aesthetics. This can be seen as an important precondition for providing a pleasant learning opportunity and be open to a large group of participants (Skarlatidou et al., 2019). The review showed, that platforms for more literate citizens exist, while others enable working with data almost without any text. This is especially interesting as creative and art-based approaches to conveying data literacy more inclusively are stressed in the literature (Bhargava et al., 2016; D'Ignazio, 2017). Information on the projects' goals is present on almost all platforms' projects' main pages. To further strengthen a positive influence on citizens' data literacy, platforms could additionally think about making them more present during the execution of tasks and highlighting their influence on the data production context or resulting requirements the task's fulfillment.

> **C)** Current digital citizen science platforms provide limited support and communication mechanisms that could be enriched with a variety of innovative functionality.

The conducted review revealed that current multi-project citizen science platforms severely vary in opportunities for support and communication on the platform. Some platforms even present no concept for either support or the communication especially, between participants. This is concerning, as contributory projects can be instrumental, postponing citizens to being data providers incapable of influencing the ways their data is used (De Albuquerque and Almeida, 2020), and studies found that there can exist a large fraction of participants contributing to digital initiatives while not feeling empowered through them (Haklay, 2013b). The close relationship between communicative means and learning opportunities was stressed in the interviews, where the participants frequently mentioned collaboration and feedback as important support aspects. Although there are notable platform examples presenting several support and communication opportunities, the interviews suggested that there is room for improvement to add innovative concepts that are already discussed in other educational settings. This includes, for example, collaboration workflows (Zagalsky et al., 2015), recommender systems (Deschênes, 2020), or chatbots (Okonkwo and Ade-Ibijola, 2021; Pérez et al., 2020). As concerns about poor data quality are one of the key obstacles for policymakers and scientists to support citizen science (Bowser et al., 2020), investing more in functionality to support data literacy education would not only be beneficial for citizens but also for researchers themselves.

## 8.7. Conclusion

In this work, we presented insights from a structured artifact review and a qualitative interview study investigating what potentials and challenges arise for applying and extending data literacy through digital citizen science projects. Grounded in the literature of citizen science and data literacy, we identified and analyzed 16 multi-project citizen science platforms according to their functionality and four design dimensions and contrasted the current conditions with the perspectives of 14 citizens and researchers. By conducting our assessment from the angle of platform design rather than individual projects, we have provided a new view on educational opportunities in digital citizen science, deriving three main positions for data literacy. Additionally, the two studies presented in this chapter can be used in multiple ways by researchers and practitioners. As a theoretical contribution, through conducting a structured artifact review, our work allows IS researchers to compare citizen science platforms and identify structural gaps in terms of research or technology development for digital citizen science (Gnewuch and Mädche, 2022). From a practical point of view, citizen science project initiators and educators can use our work for guidance in choosing a platform that is most suitable for their needs. Although different classifications for citizen science platforms already exist (Brenton et al., 2018; Liu et al., 2021), navigating the large number of digital platforms and determining whether there is an appropriate solution for their project is a challenge to practitioners (Brenton et al., 2018). Additionally, the insights gained through the qualitative study can guide future developments by discussing and contrasting the needs and considerations of both citizens and researchers.

When utilizing the results of our work, the qualitative nature of the studies must be noted, which brings limitations in terms of the generalization of the research results. First of all, the structured artifact review is subject to natural limits in the search process as well as the data extraction undertaken by only one reviewer, which implies that some platforms or functionality or characteristics of the reviewed platforms might not have been detected. Additionally, the platform landscape is constantly evolving and progressing (Brenton et al., 2018; Liu et al., 2021), which implies that the review can only serve as a momentary snapshot that should be enriched in the future with additional platforms or review dimensions. As a next step, additionally identified platforms characterized as community hubs might be interesting to investigate. Although they do not enable active participation, they might hold great potential for data literacy education. Secondly, the qualitative interview study that was used to contextualize findings and give ideas for potential future developments, although including citizens and researchers of different ages, professions, and pre-knowledge, did not strive for a representative sample. As the underrepresentation of certain cultural groups in citizen science projects is seen as a weakness and challenge (National Academies of Sciences, Engineering and Education et al., 2018; Sorensen et al., 2019), it could be valuable to extend the interview study to include a larger cultural variety of citizens within a quantitative research setting. Finally, the studies only investigate the potentials, opportunities, and challenges of data literacy learning rather than actual effects. It is likely that not all learning opportunities are equally effective, nor will they be used equally by all participants. As such, future research is needed to quantify the actual effects on learning and empowerment on the platforms based on different competencies and opportunities to promote them. By providing a starting point for this research, we hope to initiate a more profound discussion on the design of citizen science platforms and related opportunities as well as challenges for promoting citizens' data literacy.

# 9. A CA to Support Data Exploration in Citizen Science

## 9.1. Introduction

The emergence of the digital era has undeniably amplified the profound impact of data on all aspects of our lives. Technological advancements have enabled the collection and storage of large datasets (Clarke, 2016; Twidale et al., 2013). Data-driven business models motivate companies to collect and analyze increasing amounts of data, in some cases even at the expense of their customers' interests (Trzaskowski, 2022). While the increased availability of data can lead to new insights and better decisions, it also accentuates issues of inequality and exploitation (D'Ignazio, 2017). Ownership and literacy of data are critical factors in determining who can effectively utilize data to their advantage (D'Ignazio, 2017). This becomes particularly concerning as data and its products can be misused for personal, political, or economic reasons (Carmi et al., 2020; Pullinger, 2021; Trzaskowski, 2022). The evolving landscape of generative AI poses further challenges with the proliferation of disinformation in the digital realm (Hanley and Durumeric, 2023). In this context, data literacy becomes essential, not just for actively engaging in public debates and decision-making (Debruyne et al., 2021; Radermacher, 2021; Schüller et al., 2019) but also for navigating the digital landscape in general (Carmi et al., 2020). Despite its evident significance, a substantial portion of the population still lacks adequate data literacy, relegating them to passive "data subjects" rather than empowered data users (D'Ignazio, 2017). This data literacy divide perpetuates inequalities and denies individuals the agency to benefit from the data-driven landscape (D'Ignazio, 2017). Consequently, the quest for effective countermeasures becomes a critical pursuit. A promising use case in this regard is citizen science (Twidale et al., 2013), referring to the participation of non-professionals in scientific research activities (Shirk et al., 2012). Historically, citizen science strived to democratize science and counteract social inequity (Irwin and Horlick-Jones, Tom, 1995). Although participatory activities vary, in many citizen science projects, participants can access and work with scientific data (National Academies of Sciences, Engineering and Education et al., 2018). It is, hence, a natural fit when thinking about conveying data literacy and shifting modes of power and agency. Currently, however, several barriers prevent the full realization of educational benefits. First, while many citizen science projects enable participation in data collection, participation in consecutive exploration and interpretation of data is sparse (Monzón Alvarado et al., 2020). The complexity or confidentiality of data and tasks can reduce the offer of activities (Kloetzer et al., 2021). Second, researchers and project initiators are limited in resources, such as funding and time (Kloetzer et al., 2021; Wald et al., 2016) necessary to organize participation and support. This makes current educational citizen science tools such as (peer) mentoring, tutorials and trainings, or curricula potentially unsuitable as they imply additional efforts for researchers or the community (National Academies of Sciences, Engineering and Education et al., 2018). While advisory bodies call for data literacy to be actively addressed in citizen science projects (National Academies of Sciences, Engineering and Education et al., 2018), given the current challenges, appropriate solutions must first be explored. Specifically, to succeed on a larger scale, a flexible yet automated support is required. Therefore, we propose that CAs might be suitable tools for this task. CAs enable the provision of support and information cost-effectively (Kvale et al., 2021) and are used in many educational settings (Okonkwo and Ade-Ibijola, 2021). They can increase learners' motivation and enable students to access content or receive help swiftly (Okonkwo and Ade-Ibijola, 2021). As a support tool for data exploration, a CA could enable citizens to participate in this research step without producing additional work or costs for initiators, such as personal training or mentoring. Also, it would be a more scalable and constantly available solution, quickly providing citizen scientists with information and assistance to enable their participation and learning. Nevertheless, compared to tutorials and curriculums, CAs can provide personalized support that can adapt to the needs of the individual citizen. CAs require conscious design to fit the audience's and the domain's idiosyncratic requirements. Current research on CA design and utilization encompasses aspects of education (e.g., Okonkwo and Ade-Ibijola, 2021; Pérez et al., 2020), working with data (e.g., Alaaeldin et al., 2021; Narechania et al., 2021), and citizen science (e.g., Holowka et al., 2021; Tavanapour et al., 2019). However, the intersection of these three topics remains a research gap. In particular, while CAs for collecting (e.g., Holowka et al., 2021; Lia et al., 2023; Tavanapour et al., 2019) and accessing data (e.g., Narechania et al., 2021; Neumaier et al., 2017; Simud et al., 2020) have received some research attention, the consecutive use case of analyzing data (i.e., an integral part to strengthen data literacy) has not been explored yet. We hence seek to answer the following research question: *How should a CA be designed to support data exploration in citizen science applications?*

We address the research question by applying the DSR approach. Beyond conveying data literacy, we identify the need for motivation and empowerment of citizens in the literature. Based on this, we derive design principles for a CA supporting citizen participation in data exploration and implement them in a prototypical artifact. Evaluating the prototype in an experimental study, we find that using the CA can enhance data literacy and analysis performance among inexperienced users. With this research, we aim to contribute to the ongoing efforts in reducing information disparities and ensuring that data is leveraged for societal benefit. We further identify opportunities for future research by examining the limitations and challenges of the artifact and our research approach.

## 9.2. Related Work

As the foundation of our work, we review the literature on data literacy and its relationship to citizen science. We examine existing efforts to support civic engagement in data analysis and explore the potential of CAs in the domains of data literacy, education, and citizen science to guide our CA design.

**Data Literacy:** Data literacy can be referred to as “the ability to read, write and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use case, application, and resulting value” (Panetta, 2021, para. 3). Rooting back to the notion of information literacy, with an increasing amount of data and the emergence of more and more data-driven professions, data literacy emerged as a buzzword in research and popular press (Schüller et al., 2019). While Gartner’s definition of data literacy focuses primarily on describing a skill set (Panetta, 2021), other definitions also emphasize the ability and motivation to use these skills in one’s environment. Bhargava et al. (2015) define it as “the desire and ability to constructively engage in society through and about data” (p. 24), and Schüller et al. (2019) describe it as an ability needed to navigate the digitalized world and make informed decisions. Therefore, effective data literacy promotion should not only focus on skills but also on empowering and motivating learners to apply these skills in their respective contexts (Bhargava et al., 2015). To guide teaching and evaluation approaches to data literacy, Schüller et al. (2019) developed a data literacy framework that can be tailored to different needs and requirements. They subdivide data literacy into six core competencies: (1) the establishment of a data culture, (2) the provision of data, (3) the exploitation of data, (4) result interpretation, (5) interpretation of data, and (6) the derivation of actions (Schüller et al., 2019). While the framework provides information about the content required to promote data literacy, it does not address how it can be taught. Examples of teaching approaches to data literacy comprise in-person formats such as workshops (e.g., Debruyne et al., 2021; D’Ignazio, 2017), school initiatives (e.g., Bhargava et al., 2016; Gould, 2021), or online formats such as forums, quizzes, and online classes (Jayawickrama et al., 2020). In addition, many digital tools facilitate data-related tasks, such as data collection, processing, and visualization. D’Ignazio and Bhargava (2016) have mapped tools such as Excel, cartoDB, or infogr.am in view of their flexibility and expertise requirements. However, they pointed out that current tools emphasize output creation rather than learning. They derive four design principles for pedagogical learning tools: targeted focus, guidance, inviting design, and tool expandability (D’Ignazio and Bhargava, 2016). These principles should ensure that tools ease barriers to learning and quickly get users started with activities. While being invited to follow appealing first activities, users should find additional information on more demanding practices (D’Ignazio and Bhargava, 2016). Other authors stress that teaching approaches should encompass multiple pathways for users to choose from – according to their needs (e.g., Bhargava et al., 2015). These should be agile, adaptive, and focus on what is effective and meaningful for the learners, such as working with community data (Bhargava et al., 2015; D’Ignazio, 2017).

**Citizen Science:** Citizen science is defined as “the general public engagement in scientific research activities when citizens actively contribute to science either with their intellectual effort or surrounding knowledge or with their tools and resources” (Socientize Consortium, 2014, p.6). Originally used in the natural sciences, today citizen science has proven useful across many different fields (Pettibone et al., 2017). With the expansion of citizen science, the heterogeneity of participation approaches has also increased (Shirk et al., 2012; Spasiano et al., 2021). The most common project types are contributory projects that focus on participatory data collection (Bowser et al., 2020; Monzón Alvarado et al., 2020). The participatory analysis and interpretation of data is less common and usually occurs in co-created or collegial citizen science projects (Shirk et al., 2012). However, citizens have increased access to raw data, for instance, through open (government) data platforms. They could support public institutions in drawing important insights when given access to (gamified) toolkits supporting data utilization (Krishnamurthy and Awazu, 2016; Simonofski et al., 2022; Wirtz et al., 2022). On the citizen science platform Zooniverse[1], for instance, a strong focus in data analysis is put on participatory image classification (Bonney et al., 2016; Simpson et al., 2014). Moreover, for individual citizen science projects, digital tools such as Google Spreadsheet are prepared but often used only in a classroom setting, where additional support and teaching are provided (e.g., Kjelvik and Schultheis, 2019; Shah and Martinez, 2016). Within citizen science projects, learning happens either on a micro level (e.g., through active participation and the execution of tasks) or on a macro level (e.g., by sharing videos or online tutorials) (Jennett et al., 2016). However, the education of citizen scientists and, thus, the provision of educational tools is rarely the focus of citizen science projects. In a study exploring participant motivation and retention in digital citizen science projects, Wald et al. (2016) reported that for most projects, scientific outcomes were the focus while "educational and social benefits [...] were incidental" (p. 562). For researchers, the main barriers to learning are the necessary temporal, technical, or monetary resources, as well as breaking down complex tasks (Kloetzer et al., 2021). Conversely, participants can be prevented from learning due to a lack of confidence, skills, money, or time (Kloetzer et al., 2021). In addition, project design itself can negatively influence learning by including too little feedback or interaction (Kloetzer et al., 2021). These obstacles should be a starting point for technical solutions supporting participatory data exploration.

**CAs:** CAs are applications that allow users to interact with them in a natural language and can either be text or speech-based (Janssen et al., 2020; Rapp et al., 2021). They are also discussed in certain domains under the terms chatbot or chatterbot (Bittner et al., 2019). Offering automation where priory human resources are needed, CAs can be a low-threshold solution, leading to large cost reductions (Kvale et al., 2021). However, maintaining user satisfaction can pose major challenges. Studies on customer service chatbots indicate that factors such as problem resolution, answer precision, and concreteness drive customer satisfaction (e.g., Kvale et al., 2021; Van der Goot et al., 2021), while errors and a lack of functionality can quickly deteriorate it (e.g., Van der Goot et al., 2021). Likewise, studies on CAs in the workplace suggest that CA adoption depends on user characteristics such as individual tech savviness (e.g., Gkinko and Elbanna, 2023). The possible application fields for CAs range from economics

[1] www.zooniverse.org

(e.g., finance or e-commerce) to personal applications such as health or emotional support (Rapp et al., 2021). It has been shown that good CA design highly depends on the domain it is built for (e.g., Bittner et al., 2019). Further research on the transferability of design knowledge between contexts is necessary (Diederich et al., 2022). For instance, while the usage of social cues is encouraged in some CA applications (e.g., Holowka et al., 2021; Tavanapour et al., 2019), it can have detrimental effects when the reliability of information is essential (Stieglitz et al., 2022). Thus, a one-size-fits-all approach to CA design is unrealistic – context, stakeholders, and unique value propositions must be considered (Janssen et al., 2020). To guide the design of a CA to support citizens in data exploration, different domains provide interesting insights. In the following, we shed light on insights from the application of CAs for education, (big) data-related work, and citizen science (see also Table A.4 of the Appendix).

In the domain of education, research differentiates between teaching- and service-oriented CAs. While the former describes CAs targeting knowledge generation, service-oriented CAs provide administrative services, such as introductory or library services (Pérez et al., 2020). When interacting with learners, CAs usually act as "teacher, student or colleague" (Tamayo-Moreno and Pérez-Marín, 2016, p.1). In this role, CAs have proven beneficial as they allow for integrating multiple content into one tool and parallel access by multiple users (Okonkwo and Ade-Ibijola, 2021). The possibility of receiving immediate help on demand is convenient with positive effects on learning motivation (Okonkwo and Ade-Ibijola, 2021). CAs also proved suitable for closing learning gaps between mainstream learners and learners from certain minority groups (Pérez et al., 2020). However, through a structured literature review, Pérez et al. (2020) have identified boredom and user frustration (e.g., through lengthy messages and inadequate replies) as common impediments. Teaching CAs can target various topics and domains, with language learning being a prominent use case (Pérez et al., 2020). Another use case, closer related to data literacy, is math education where CAs have been used (Anh and Ngan, 2021; Nguyen et al., 2019). CAs in data-related work environments usually focus on data provision and depiction for non-technical skilled employees (e.g., Alaaeldin et al., 2021; Narechania et al., 2021; Simud et al., 2020). CAs can conduct tasks such as generating database queries and visualizations in or based on natural language (e.g., Hoon et al., 2020; Narechania et al., 2021; Neumaier et al., 2017; Simud et al., 2020). They can also support decision-making by explaining analytic tools and key performance indicators for a given dataset (e.g., Alaaeldin et al., 2021). Another important application is the identification of relevant datasets, domain-specific scientific tools, and methods (e.g., Keyner et al., 2019; Neumaier et al., 2017; Zhang et al., 2018). Overall, CAs in data-related work environments mainly focus on overcoming modern databases' technical complexities through natural language interfaces. Since supporting the understanding of data and analysis methods is not the focus of publications, fundamental data literacy remains a prerequisite for users.

Within the domain of citizen science, the usage of CAs for different activities is not yet a common practice. First and foremost, CAs have been used for quantitative and qualitative data collection in citizen science projects (e.g., Holowka et al., 2021; Isacco et al., 2018; Lia et al., 2023; Tallyn et al., 2018;

Tavanapour et al., 2019). They enable participants to answer questionnaires, upload text, pictures, or geotags (e.g., Isacco et al., 2018; Lia et al., 2023; Tallyn et al., 2018) and can, in return, provide guidance or encouragement and support or share data directly with experts or the community (e.g., Holowka et al., 2021). Advantages of their use in data collection can include personalized feedback and the ability to conduct further inquiries when observations are incomplete (Portela, 2021). Additionally, they can provide citizens with data or visualizations (Portela, 2021). Other work explores the advantages of CAs facilitating the ideation process by collecting, structuring, and presenting ideas (e.g., Tavanapour et al., 2019) or supporting the community and its interaction (e.g., Athreya et al., 2018; Portela, 2021). Overall, the potential of using CAs for citizen science seems to be not yet exploited. For example, we could not identify literature presenting a CA used for training citizen scientists, although the suitability of CAs for educational purposes has been proven in other domains (Okonkwo and Ade-Ibijola, 2021).

**Research Gap:** The related literature on data literacy, citizen science, and CAs offers valuable insights into the possibilities and challenges associated with designing activities and tools for enhancing data literacy. However, it underscores a significant research gap at the intersection of these topics: the design of tools for active participation in data exploration. While the data literacy literature provides crucial insights into tool design for learners (e.g., D'Ignazio, 2017; Schüller et al., 2019), it emphasizes the need for more learning-oriented tools embedded in a meaningful context for the user (Bhargava et al., 2015; D'Ignazio and Bhargava, 2016). The citizen science literature addresses this context and discusses participants' learning and tools to support projects (e.g., Jennett et al., 2016; Liu et al., 2021; National Academies of Sciences, Engineering and Education et al., 2018). However, it reveals numerous challenges in integrating educational components (e.g., Kloetzer et al., 2021; Wald et al., 2016) and that the data analysis step is often not addressed. The educational CA literature generally discusses opportunities and challenges in using CAs for teaching (e.g., Okonkwo and Ade-Ibijola, 2021; Pérez et al., 2020), but in the specific context of working with data, the focus remains primarily on discovering datasets (e.g., Keyner et al., 2019; Neumaier et al., 2017; Zhang et al., 2018) or making them accessible (e.g., Hoon et al., 2020; Narechania et al., 2021; Neumaier et al., 2017). Likewise, CAs in citizen science do not focus on facilitating data analysis but rather center on qualitative or quantitative data collection (e.g., Holowka et al., 2021; Lia et al., 2023; Tallyn et al., 2018). Additionally, they do not exploit their teaching capabilities. Overall, the existing literature provides crucial insights for designing CAs supporting citizens in data exploration, yet specific design guidelines for this use case are missing. Considering the challenges on the transferability of design knowledge across contexts (e.g., Bittner et al., 2019; Diederich et al., 2022; Janssen et al., 2020), this represents a significant research gap that should be further explored and investigated.

## 9.3. Research Approach

To answer our research question, we conducted a research project following the DSR approach of Peffers et al. (2007). For a "problem-centered approach" (Peffers et al., 2007, p.56), the methodology includes six steps, starting with identifying a problem and deriving a motivation for its solution (see Figure 9.1). First, to elaborate on the problem to solve, we assessed the current state of research by reviewing related literature in the fields of data literacy, citizen science, and CAs and performed a stakeholder analysis based on insights from these domains. In addition, we carried out an expert workshop on data analysis conducted with 12 experts and advanced practitioners, combining ideas from two established requirements elicitation methods, "Introspection" and "Brainstorming" (Sharma and Pandey, 2013). While in an introspection, experts elicit the user needs based on their domain knowledge, in brainstorming, participants from different stakeholder groups are invited to collectively generate ideas (Paetsch et al., 2003). The expert workshop facilitated introspection through a think-aloud session about "conducting a data analysis". Think-aloud sessions, typically known from usability testing, enable researchers to gain insights into participants' thought processes by neutrally observing them speaking out their thoughts aloud while working on a given task (Ericsson and Simon, 1984; Fan et al., 2020). After the individual think-aloud session, brainstorming was conducted in the expert workshop as an open discussion between all participants, guided by the components of the data literacy framework (Schüller et al., 2019). The approach to the conduct of the expert workshop is further described in Section 9.4.1. Second, the following activity of the DSR approach comprised the definition of solution objectives by deriving either quantitative or qualitative requirements for the artifact (Peffers et al., 2007). Therefore, we translated the results from the first activity into atomic user needs as solution objectives and positioned them in the related literature. Thirdly, we approached the actual design and development phase by instantiating our artifact. The two tasks of this phase were outlining the necessary function and design requirements before practically creating the artifact (Peffers et al., 2007). Using our user needs, we determined design specifications, following the schema by Gregor et al. (2020), and implemented them in a prototypical instantiation. The approach to specifying the design principles and implementing the artifact is further described in Sections 9.4.2 and 9.4.3 respectively. To close the first DSR cycle, the artifact should be demonstrated and evaluated. We covered these steps simultaneously by designing and conducting a final experiment with 30 participants guided by best practices to evaluate CAs and data literacy learning (Pérez et al., 2020; Schüller et al., 2019) and DSR artifacts (Venable et al., 2016). Using a between-subjects design, the experiment compared the data analysis performances and (learning) experiences of a participant group guided by the designed artifact with the results of a group that received no explicit guidance but was allowed to use existing material through the web. By combining quantitative (task performance, empowerment, motivation, perceived learning, user experience) and qualitative (user experience, dialogue analysis) insights, the experiment enabled us to evaluate the artifact comprehensively. The approach to the experiment's design, conduction, and evaluation are described in detail in Section 9.4.4

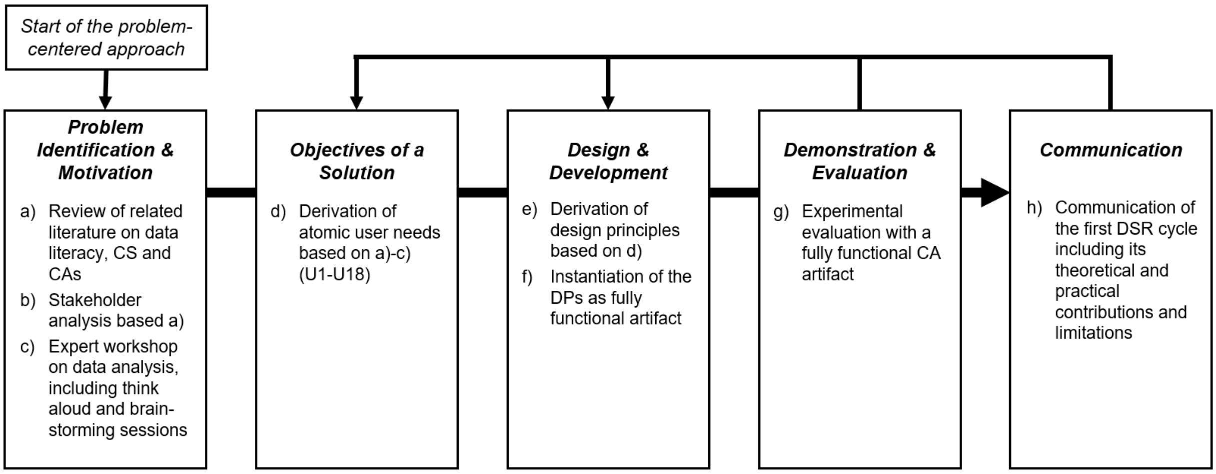


Figure 9.1.: Overview of the DSR approach.

## 9.4. Designing a CA for Public Participation in Data Analysis

### 9.4.1. Problem Awareness and Solution Objectives

We set out to design a CA capable of supporting data exploration in citizen science projects. This design endeavor entails understanding the intricacies of data literacy and citizen science and applying this knowledge to CAs. The review of relevant literature has shown that fostering data literacy requires creative teaching approaches that consider the interests and realities of the target audience (e.g., Bhargava et al., 2015; D'Ignazio, 2017). A challenge in this regard is the complexity of the above concepts and that data literacy is based on a set of competencies rather than a specific skill or technique (Debruyne et al., 2021; Schüller et al., 2019). The intended data literacy content thus needs to be broken down to guide the design of the CA (cf., Expert Workshop). While data collection is frequently seen in citizen science projects (Bowser et al., 2020; Liu et al., 2021; Monzón Alvarado et al., 2020), few projects use participatory data analysis (beyond classification tasks). In general, budgets and time constraints limit such projects for researchers and citizens (Kloetzer et al., 2021; Wald et al., 2016). In addition, maintaining momentum and keeping citizens engaged and active (beyond the first exploration) is crucial but difficult (Wald et al., 2016). Therefore, understanding citizen scientists is key to the design of appropriate support tools (cf., Stakeholder Analysis).

**Expert Workshop:** To understand the specific requirements of the CA (content and design), we invited 12 data analysts first to an individual think-aloud session and second to a brainstorming session. The participants consisted of six Ph.D. candidates and six master students majoring in data science-related fields and presented, thus, advanced to professional practitioners in the field. Sessions were conducted virtually using an online conference tool moderated by one researcher, which had some implications for the workshop conduction. In the think-aloud session, participants were asked to solve several analytical tasks based on a given dataset. Instead of being able to observe their actions physically, in the virtual think-aloud sessions, we used audio, video, and screen sharing to enable as many insights as possible for the researcher. In the following group brainstorming session, participants could then discuss approaches and pitfalls to data exploration and requirements for support based on their own experiences in the think-aloud session. Since in virtual compared to on-site groups lower social presence can be a challenge to the discussion quality, we limited the size of the brainstorming groups to three participants per session to better integrate the individual participants and utilized an online whiteboard to facilitate collaboration (Roberts et al., 2006). The results of the expert workshop were translated into atomic user needs of students in data exploration and tutors in student support, and contextualized with the literature review results (see Table 9.1).

**Stakeholder Analysis:** Most citizen science initiatives cater to broad audiences (Spiers et al., 2019). Common user characteristics include above-average education, above-average income, and above-average seniority (Mac Domhnaill et al., 2020; National Academies of Sciences, Engineering and Education et al., 2018). In addition, citizen scientists tend to "embody the characteristics of autonomy, competence, and relatedness in their hobby" (Jones et al., 2018, p.15). However, citizen science embraces the diversity of participants, and project organizers claim to strive for more diversity in terms of age, gender, and ethnicity (National Academies of Sciences, Engineering and Education et al., 2018). Thus, the level of autonomy and knowledge can be assumed to be heterogeneous. Based on this and research on data literacy (e.g., Logan, 2017; Watson and Callingham, 2004), we distinguish the following (broad) user groups to have in mind for the CA artifact: 1) Beginner Users, 2) Advanced Users, and 3) Professional Users (see Table 9.2).

### 9.4.2. Design Principles

A design principle should include an aim, implementer, and user; a context; mechanisms, and a rationale (Gregor et al., 2020, p.1634). We follow this scheme and propose five principles for the design of a CA for support in data analysis tasks (context), which can be used by researchers and developers (implementer) to create software support for non-expert citizen scientists undertaking their analytical activities (users). Considering user needs U1, U11, U14, and U18, the platform must specify a certain process structure while the user remains free to choose how to follow this path. We find guidance for this requirement in the design principles of Tavanapour et al. (2019), who state that a CA for idea creation must be able to follow a given conversation flow while still being able to lead the process actively and D'Ignazio and Bhargava (2016), who underlined, a data literacy tool should be guided. Additionally, Portela (2021) advises that a CA should include fixed chat commands for user orientation. Therefore, we formulate the first design principle as follows:

| User needs derived from the workshop | | | Concepts in the literature |
|---|---|---|---|
| As a **student**, I want to... | U1 | Understand the data analysis process | Different activities in data value creation (Schüller et al., 2019) |
| | U2 | Understand the dataset (meaning, usefulness) | Competence Obtain Data, Prepare Data (Schüller et al., 2019) |
| | U3 | Know and apply methods of data cleaning | |
| | U4 | Find a start for the data analysis | Low entry point for data literacy tools (D'Ignazio and Bhargava, 2016) |
| | U5 | Understand and select a method for data analysis | Competence Analyse Data, Interpret Data Analysis, Visualize Data, Interpret Data Visualizations (Schüller et al., 2019) |
| | U6 | Note limits and challenges of the data analysis | |
| | U7 | Interpret analysis results and critically question findings | |
| | U8 | Select and design appropriate data visualizations | |
| | U9 | Understand how to interpret and check visualizations | |
| | U10 | Ask questions | CA's answering student's questions (Okonkwo and Ade-Ibijola, 2021) |
| | U11 | Decide on an analysis action and path | Enabling multiple pathways for learners to choose from (Bhargava et al., 2015) |
| | U12 | Get access to data analytics tools | Knowledge and mastery of tools as essential skills (Schüller et al., 2019) |
| | U13 | Get access to assistance and helpful material | CA's integrating multiple contents (Okonkwo and Ade-Ibijola, 2021) |
| As a **tutor**, I want to... | U14 | Steer through the analysis process | Guided data literacy tools (D'Ignazio and Bhargava, 2016) |
| | U15 | Point out missing competencies and skills | - |
| | U16 | Proceed with questions (answer, redirect, collect) | CA's answering student's questions (Okonkwo and Ade-Ibijola, 2021) |
| | U17 | Form an interface to other material/ software | Expandable data literacy tools (D'Ignazio and Bhargava, 2016) |
| | U18 | Support different levels of pre-knowledge | CAs providing individualized support (Okonkwo and Ade-Ibijola, 2021) |

Table 9.1.: Atomic user needs grouped by perspective and contrasted with related literature.

**DP1** In order to structure the analysis process (aim), the system should provide a sorted menu highlighting the individual parts of a data analysis (mechanism), as this enables the user to get guidance on the process and navigate to a specific topic of interest (rationale).

Several user needs (U2, U4) express that beginners' entrance must be eased. Thus, the system should "provide a low entry point" (p. 87) for data analysis (D'Ignazio and Bhargava, 2016). Nevertheless, the

| User Groups | Description |
|---|---|
| Beginner Users | No to little knowledge about data literacy (see idiosyncratic level, informal level (Watson and Callingham, 2004) |
| | New to citizen science or without the typical characteristics of an experienced citizen scientist, requiring guidance and explanations |
| Advanced Users | Experienced citizen scientists or people experienced in data literacy with a fundamental understanding of the data process (see conversational level; Logan, 2017), data analysis, and visualization methods (see inconsistent and consistent non-critical level; (Watson and Callingham, 2004)) |
| | Knowledge is rather basic and incomplete or might date back a long time ago, but users are more autonomous and inform themselves or might have specific questions |
| Professional Users | Familiar with working with data (see critical and critical mathematical level; (Watson and Callingham, 2004)) and working largely to completely autonomously |
| | Group is rather out of scope for the CA |

Table 9.2.: Description of user groups for the CA artifact.

knowledge needed to conduct many such tasks is comprehensive (U3, U5, U7, U8, U9). The stakeholder analysis showed the need to account for different user groups, which is supported by Bhargava et al. (2015), pointing out the necessity to provide "multiple pathways for people with different data literacy needs and capacities to interact within a complex system" (p. 15). Therefore, the platform should equip users with the appropriate background knowledge based on their needs and interests. Tavanapour et al. (2019) to this end, propose a comparable mechanism specifying that CAs should have the "capacity to summarize [...] information [...] and offer further explanations, if requested" (p. 8). We, therefore, formulate our second design principle as follows:

> **DP2** The system should provide a tiered knowledge structure (mechanism) to educate the user efficiently (aim), as this enables the user to determine the depth according to their interests and skills (rationale).

User needs U10 and U16 express that the users should be enabled to get answers to their specific questions, which is a common functionality for teaching CAs (Okonkwo and Ade-Ibijola, 2021). We, therefore, straightforward formulate the third design principle as follows:

> **DP3** The system should allow users to enter questions and process them (mechanism) to get answers to individual questions (aim, rationale).

User needs U12, U13, and U17 imply that the platform should use existing teaching material. To this end, D'Ignazio et al. (2016) point out that data literacy tools should be expandable, bridging the pathway for learners to go from one data literacy tool to the other. We incorporate these findings in the fourth design principle:

**DP4** To efficiently educate the user (aim), the system should provide a combination of self-developed and external materials through embedding or forwarding (mechanism), as users have an interest in a broad offer of learning material (rationale).

Furthermore, the presence of many pitfalls (user needs U6, U9, U15) requires the platform to support users in understanding challenges and avoiding common mistakes. We, therefore, propose that:

**DP5** The system should provide indications and warnings of challenges and common mistakes in time (mechanism) to prevent the user from failing (aim, rationale).

An overview of the design principles and their derivation from the atomic user needs can be found in Table 9.3.

| **Design Principles** | | **Associated User Needs** |
|---|---|---|
| DP1 | Provide a menu that structures the analytics process for the user to get guidance or navigate to a specific topic. | U1, U11, U14, U18 |
| DP2 | Provide a tiered knowledge structure to let the user determine the depth according to their interests and skills. | U2, U3, U4, U5, U7, U8, U9 |
| DP3 | Allow the users to enter questions and process them by either ad hoc answering or forwarding. | U10, U16 |
| DP4 | Provide a combination of self-developed and external materials through embedding or forwarding to educate the user efficiently. | U12, U13, U17 |
| DP5 | Provide indications and warnings of challenges and common mistakes in time. | U6, U9, U15 |

Table 9.3.: Design principles mapped to their respective user needs.

### 9.4.3. Artifact

In the third phase of the DSR process, the formulated design principles are instantiated in an artifact in the form of a CA prototype. The CA provides dataset-independent support to beginner and advanced users in the process of data analysis. It provides process-oriented and knowledge-based advice through messages, pictures, links, and guidance along two workflows:

**Workflow 1:** The data analysis workflow offers guidance for beginners and provides a menu of steps (Figure 9.2), showing different steps of a typical data analysis process (DP1). The menu serves as a central point to which the user returns within the flow. The first step of the data analysis process ("Getting started") reflects DP2, DP4, and DP5. After receiving information on how to get started, the user can request more information (DP2) or browse through external education material (DP4). To address DP5, the CA invites users to analyze their data actively. Upon the user's confirmation to proceed, the bot provides an overview of common mistakes concerning the task the user has just completed (DP5).

**Workflow 2:** The question-and-answer workflow should attract users with basic data knowledge. Here, users can specify topics of interest by asking questions. Upon a request, the CA either recognizes the question as dataset-specific or methodological. In the former case, the bot points out that such questions are out of scope. In the latter case, the bot provides an answer if it recognizes the question. If the question is not recognized, the CA offers to forward the question to a supervising researcher. Upon affirmation, the CA sends questions and contact details to an online spreadsheet, privately accessible to the supervising researcher.

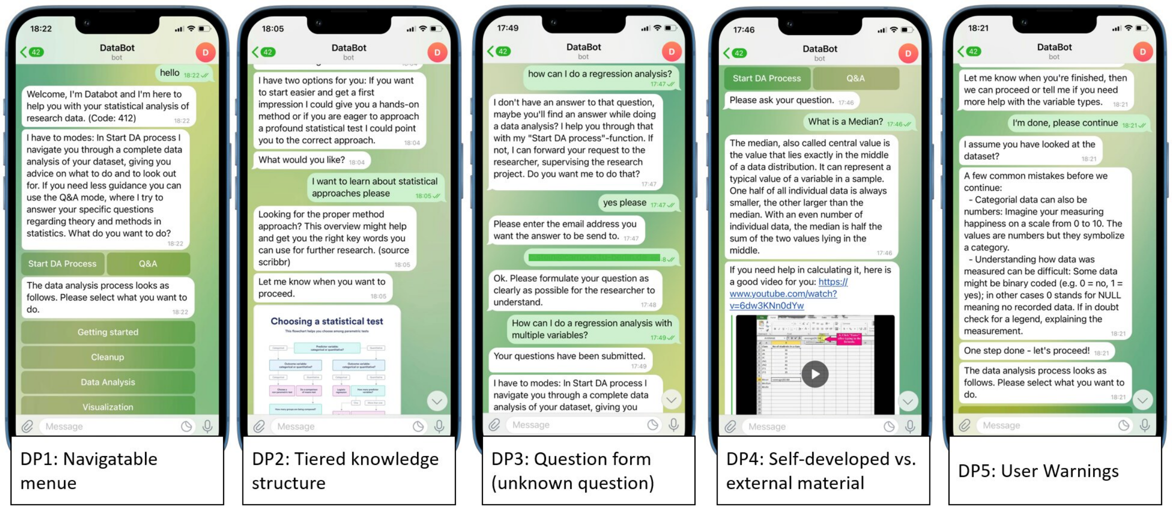


Figure 9.2.: Exemplary conversations with the CA reflecting the implementation of the design principles.

The implementation of our CA is based on the open-source programming framework Rasa Open Source Machine Learning Framework. We used a local setup with the default NLU and core component configuration for the prototype. In addition, Rasa X has been used to facilitate conversation-driven development, repeatedly asking potential users to test the CA in its different stages of development. To lower the barrier of usage, we chose to implement the front end via Telegram (i.e., a popular messenger service).

### 9.4.4. Evaluation

The evaluation phase of a DSR project assesses whether (and if so, how) the artifact solves the problem (Peffers et al., 2007). While for customer service bots it is often sufficient to determine the share of adequate responses, the degree of success in education applications depends on the learning effect generated for the user (Pérez et al., 2020). Thus the evaluation can, for instance, rely on the learner's perception measured through questionnaires or on comparison with a control group, not utilizing the CA (Pérez et al., 2020). For the evaluation of data literacy, Schüller et al. (2019) propose the stage model by Kirkpatrick, which assesses enjoyment, learning success, the learner's behavior, and learning outcomes (Kirkpatrick, 1959). Naturally, behavior and (long-term) learning outcomes can only be evaluated to a limited extent, especially when drawing on online experimental methods. We thus combine approaches from the CA and data literacy perspective in a between-subjects experiment. The experiment consists of an initial questionnaire, the main experimental part, and a post-questionnaire. The initial questionnaire (Table A.2, Appendix) is used to evaluate prior knowledge of the study participants, such as knowledge about citizen science, data analysis, and experience in working with datasets. Additionally, it assesses their motivation for science, including aspects of intrinsic and career motivation, and self-efficacy, using items from the Science Motivation Questionnaire (SMQ), a well-established instrument in pedagogic work and research (Glynn et al., 2011). In the practical part, participants are introduced to the topic of citizen science and the experiment task, which is a data exploration of the data of different formats on passengers on the Titanic (e.g., age (numerical), embarkation port (categorial), survival (binary)). This open-source dataset is free of charge and is known for its use in introductory courses, as well as academic studies (e.g., Ekinci et al., 2018; Gupta et al., 2018), making it particularly suitable for the experiment. The participants receive an extract of the dataset in the form of a .csv and .xlxs file and are asked to complete 12 practical and theoretical data exploration tasks (Table A.1, Appendix). Following a between-subject design, one user group may use our CA artifact for these tasks. They obtain an introduction to the CA and its functionalities and are asked to install it on their device. In contrast, the control group does not receive the artifact. To depict the current status quo, this control group is advised that it is allowed to use all other existing software or learning material, for instance, through the web. The overall task performance of participants is calculated based on all 12 tasks and normalized to the interval [0, 1]. The final evaluation quantitatively assesses user motivation, empowerment, and perceived learning, providing a contrasting introspective view of task performance. To do so, we make use of established survey constructs i.e., interest / enjoyment from the Intrinsic Motivation Questionnaire (Center for Self-Determination Theory, 2022; Ryan, 1982), a second-order construct for empowerment (Kim and Gupta, 2014), and a construct for perceived learning (Alavi et al., 2002) measured on a 7-point Likert scale ranging from strongly disagree (1) to strongly agree (7). In addition, for the CA treatment, user experience is assessed via the User Experience Questionnaire (UEQ) (Schrepp et al., 2017). To get qualitative input, we use further open-ended questions and analyse participants' conversations with the CA. The post-questionnaire can be found in Table 8 of the Appendix.

**Testing**: First, we conducted a set of pretest sessions, in which five users tested the CA treatment and one user tested the control treatment.

**Procedure**: The experiment was executed between December 2021 and January 2022. Participants were recruited through the online platform Prolific (e.g., Palan and Schitter, 2018). Although the experiment was conducted asynchronously, participants could contact a supervising researcher in case of any issues or technical difficulties during the sessions. The average payout was 6.21 GBP per hour, and the average completion time was 46 minutes. From those, participants spent 39 minutes on average processing the analytical tasks.

**Sample**: The sample included $n = 30$ international participants who self-reported fluency in English. Participants were between 19 and 45 years old (23 male, 7 female) and could mostly be allocated to the level of beginner to advanced users: 37.67% of the participants had never worked with datasets before, while the average pre-knowledge of data analysis was indicated with 4.23 points (see Table 9.4). Only one participant indicated a very high level of data analysis pre-knowledge potentially presenting a professional user. The sample was split into $n_{control}$= 10 participants for the control treatment and $n_{CA}$=20 for the CA treatment. This distribution was chosen for several reasons. Firstly, the primary focus of the study was to evaluate the design and effectiveness of our CA by assessing its usage through real users. Thus, by allocating a larger sample to the CA treatment, we collected more comprehensive and diverse insights into use behaviors and perceptions. Secondly, while the experiment design enables us to evaluate user interaction with the CA, interaction with self-selected materials and sources in the control treatment cannot be tracked. Therefore, uncovering different types of users or approaches by opting for a larger sample size was only feasible for the CA treatment. The importance of the control group was to evaluate the relative effectiveness of the CA. Overall, we opted for a smaller sample size in favor of longer experiment duration to ensure that participants were sufficiently engaged with the tasks at hand.

**Randomization Check**: To ensure treatment randomization, the distribution of age, gender, and interest- and knowledge-related factors were evaluated. There were no significant age differences between the two treatment groups, with a mean age of 27.1 years in the CA treatment and 27.5 years in the control treatment ($p = 0.558$). Similarly, the gender distribution was not significantly different, with 20% and 30% female participants in the CA and control treatment, respectively ($p = 0.885$). The two-sided t-test results for the remaining factors involving pre-knowledge and interest also showed no significant differences between the mean scores of the two treatments (see Table 9.4). Therefore, we assume treatment randomization was successful.

| **Treatment** | | **Data Analysis Knowledge** | **Software Skills** | **Dataset experience*** | **Citizen science concept*** | **SMQ** |
|---|---|---|---|---|---|---|
| CA | Mean | 4.400 | 5.700 | 0.650 | 0.100 | 6.058 |
| | SD | 1.536 | 0.923 | - | - | 0.694 |
| Control | Mean | 3.900 | 4.900 | 0.600 | 0.200 | 5.666 |
| | SD | 1.792 | 1.287 | - | - | 0.568 |
| p-value | | 0.433 | 0.060 | 0.797 | 0.465 | 0.134 |

Note: Variables have been measured on a 7-point Likert scale, except for *, which were binary variables. P-values refer to the results of two-sided t-tests.

Table 9.4.: Randomization checks for interests and pre-knowledge of experiment participants.

#### 9.4.4.1. Quantitative Results

We now assess how the availability of the CA affected participants' perceptions of task performance, perceived learning, empowerment, and motivation. The data is shown in Figure 9.3 and summarized in Table 9.5.

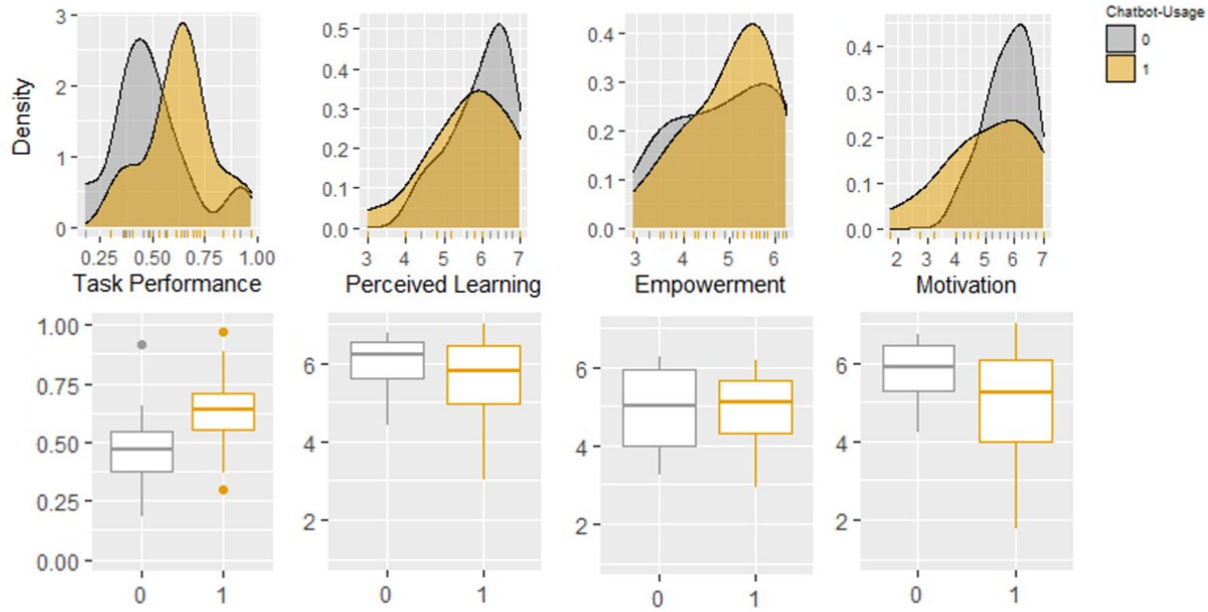


Figure 9.3.: Distribution and score comparison for task performance, perceived learning, empowerment, and motivation grouped by treatment.

**Task performance:** Task performance was measured on a scale from 0 to 1, where 0 indicates no task has been solved correctly, while 1 indicates a participant has solved all tasks correctly. We find that task performance is significantly higher in the CA treatment than in the control condition ($p = 0.048$), with a relative surplus of 14%. Beyond overall task performance, we also assess each task individually. An overview of the results can be found in the Appendix. Notably, for all but two tasks, participants from the CA condition performed better than their counterparts from the control condition.

**Perceived learning, empowerment, and motivation:** Overall, participants indicated substantial perceptions of learning (i.e., 5.75 points on the 1-7 Likert scale), empowerment (i.e., 4.96 points on the 1-7 Likert scale), and motivation (i.e., 5.33 points on the 1-7 Likert scale). For all variables, however, we do not find significant differences between treatment conditions (see Table 9.5).

**Impact of Age, Gender, and Pre-Knowledge:** We do not find any effects of age or gender on any of the target variables. For participants' pre-knowledge, we see a small effect on task performance ($\beta = 0.04$, $p = 0.047$). In addition, we observe small correlations between SMQ (p=0.005) and Software Skills ($p = 0.030$) with perceived empowerment.

**User Experience:** Overall, users sent between five and 44 messages to the CA ($Mean = 16.35$, $SD = 9.670$, $total = 327$). During their interaction with the CA, they followed different paths: Most users

| Treatment | | Task Performance | Perceived Learning | Empowerment | Motivation |
|---|---|---|---|---|---|
| **CA** | Mean | 0.630 | 5.640 | 4.979 | 5.113 |
| | SD | 0.167 | 1.077 | 0.925 | 1.490 |
| **Control** | Mean | 0.488 | 5.960 | 4.925 | 5.775 |
| | SD | 0.197 | 0.799 | 1.102 | 0.786 |
| p-value | | 0.048 | 0.414 | 0.888 | 0.122 |

Note: p-values refer to two-sided t-tests or Welch-tests, respectively.

Table 9.5.: Summary statistics for task performance, perceived learning, empowerment, and motivation grouped by treatment.

(55%) initially followed the data analysis workflow in the proposed order. However, some users (25%) used the menu to jump to the topics of interest directly. Moreover, 20% of users used the question-and-answer workflow rather than the data analysis process path. The different user pathways and transitions are illustrated in Figure 9.4. On average, the CA made between 0 and 3 false intent classifications (Mistake) within the user conversations ($Mean = 0.8$, $SD = 0.834$). Most often, they appeared in the visualization part of the data analysis process. In addition, in three conversations, the CA could not answer a question that would have been in scope (QA other questions). The user experience during the conversations was evaluated with the short UEQ, including eight items. On average, a participant rates an item with 4.65, slightly above the center value. For hedonic items, the rating is lower, with an average score of 4.45; for practical items, the average score is 4.85. The best score can be achieved in the category constructive vs. supporting, while the lowest score is measured for boring vs. exciting.

#### 9.4.4.2. Qualitative Results

Within an open question asking for feedback about the experiment, both groups were highly satisfied and stated that they enjoyed the experience. One participant from the CA treatment stated: “I would like to participate more in studies using this kind of educational bot” (P1) while another denoted it as a “fun experience” (P28). On the other hand, a control group participant stated: “This study was for sure challenging, which I tend to enjoy [...]” (P30). Furthermore, both groups were asked about the support they used and additional support they would have liked. In the control group, the most common support tools used were the Internet (70%) and Microsoft Excel (70%), while some participants stated that they had used a calculator or had drawn on pre-existing knowledge. For the CA group, similar tools were mentioned. However, the usage of the Internet (40%) and Excel (35%) was less frequent. One participant indicated the usage of videos suggested by the CA, while multiple participants mentioned YouTube, leaving open which videos they had watched specifically. Regarding requested support, the control treatment brought forward several ideas: Most frequent proposals were process guidance, more context information, and help with calculations such as formulas or step-by-step explanations. One control group participant formulated: “I would have liked a person to guide me through the process, I did not

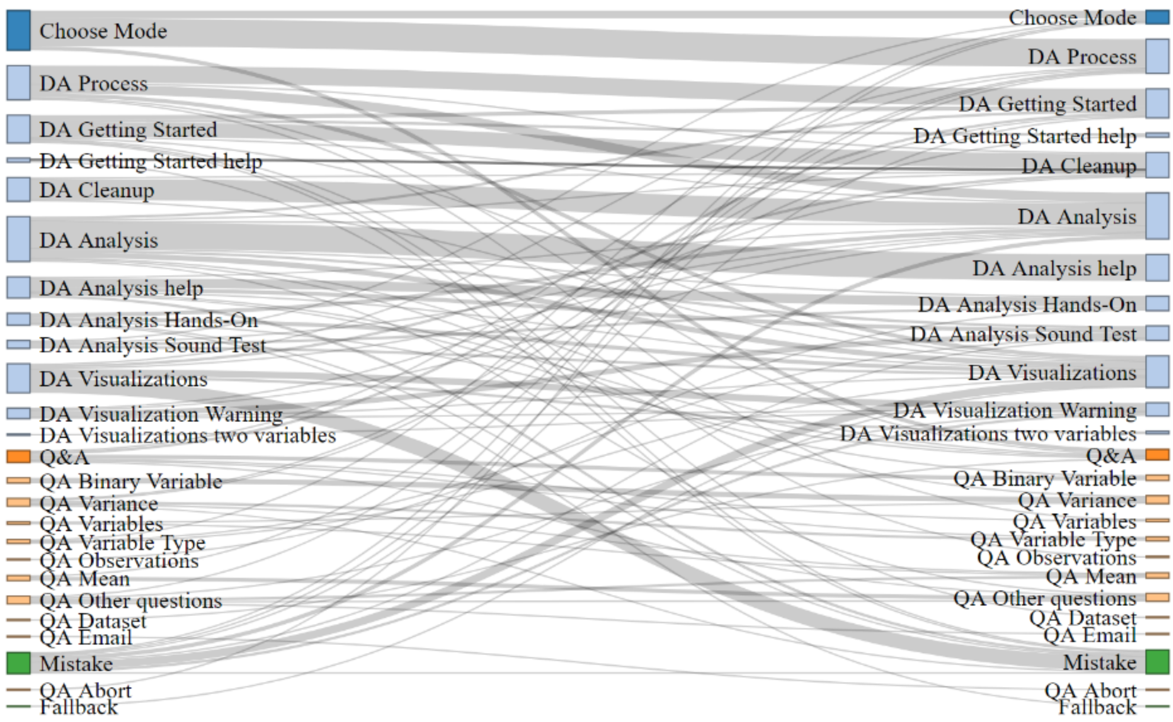


Figure 9.4.: State transitions in conversations with the CA.

know where to start" (P4). Another stated: "I believe a [...] better explanation on how to go [about] the analysis process for a [newbie] would have done justice" (P26). Other ideas from control group participants comprise terminology or Excel explanations and the integration of examples. In contrast, the CA group proposes concrete improvements and extensions to the CA. Most frequently, they formulate the desire to be able to ask dataset-specific questions. Other repeatedly mentioned ideas are automated calculation support, the ability to ask more detailed questions and to have more examples at their disposal. Some participants want supporting graphs, more buttons, and more hints. In addition, users mentioned concrete criticism of the CA's current implementation. One participant indicates that the introductory part should be designed to be more accessible and catchy: "The data bot can be a great tool to get [youngsters interested] in stats, but is missing a proper introduction that can spark the [interest]" (P22). Two participants added that they were confused about the conversation structure, while others thought the CA was too text-heavy. A participant formulates: "I was a little confused on if I had to follow the exact steps the Bot was leading me to, or if I could ask a question totally different from what the bot was telling me" (P3). Overall, participants from the CA treatment had varying levels of satisfaction with the artifact. Positive statements indicated that it provided appropriate support and satisfactory performance. Specifically, one participant stated: "The chatbot was all I needed" (P21). Another said, "The bot was more competent than I expected. Thanks to such bots, anyone can analyze data" (P1). Critical statements indicated different reasons for dissatisfaction. Interesting quotes were, for example: "it is easier to find the necessary info on the internet" (P13), "it would have been better if someone was on a call ready to answer any concerns or difficulties I had" (P28), or "The support was ok, but a crash course or sample problem would have been better" (P30). Finally, participants from the CA treatment group were asked whether and where they could imagine the usage of the CA other than in citizen science projects. Overall, 19 out of 20 participants proposed several application fields, including analysis courses, for instance, in school or university, research, or companies.

In summary, the qualitative results highlight that participants of both treatment groups generally enjoyed the experiment, utilizing different tools for data analysis and having numerous ideas for structuring further support or improving the CA. Sentiments from both treatment groups are highly relevant to evaluate the need for existing functions and identify additional requirements. Moreover, insights into the varying levels of tool satisfaction of the CA treatment group are particularly interesting for understanding user behavior and possible groups.

## 9.5. Discussion

### 9.5.1. Summary

This work presents the outcome of a DSR project that aimed at developing a CA for scientific data exploration with citizens. Building on literature and a qualitative study, we gathered user needs for a CA that assists data exploration. We then established five design principles and implemented them in a prototypical application. The prototype was evaluated by an online experiment and benchmarked against self-organized tools concerning perceived learning, empowerment, motivation, and actual performance in a series of analytical tasks. The experiment's findings provide insights into the CA's design and the particularities of different users, which are reported in the following.

**The CA Design and Effectiveness:** Considering DP1-DP5, the analysis of the CA's conversations revealed that participants used both the data analysis process flow and the question-and-answer flow (DP3), and their usage behaviors demonstrate that the tiered knowledge structure and the freedom to follow the process or navigate to specific topics (DP1, DP2) were effectively implemented. Notably, participants used the question-and-answer flow to ask questions but did not utilize the option to forward a question. This is an important aspect to consider when evaluating the effectiveness of DP3. Regarding the learning material offered, most participants acknowledge the learning support offered by the CA. Some participants also highlighted using the external material provided (DP4). Regarding DP5, the better performance of CA participants in one particular task could indicate that the respective warning issued by the CA was adequate and successful. In terms of effectiveness, overall, participants provided access to the CA performed significantly better than those using self-organized support as a control.

Interestingly, this effect was not reflected in participants' self-perceptions, and no significant correlation between Task Performance and Perceived Learning could be measured. Although educational studies report similar observations (e.g., Barzilai and Blau, 2014; Vinuales et al., 2019), this issue might hold important implications. Considering established IS theories, the participants' impression of the technology's usefulness affects usage intention (Davis, 1989). In this regard, users could unjustly reject an educational tool that does not affect perceived learning, although task performance improves, pointing towards untapped potentials of these technologies. In terms of motivation and empowerment, our findings did not provide evidence of a positive impact of the CA on the participants. Possible explanations may be related to issues with the CA's design. Literature suggests that long text messages and inadequate responses can lead to user boredom and frustration (Matsuura and Ishimura, 2017; Pérez et al., 2020). We found that some users experienced these issues, potentially impairing motivation. Additionally, UEQ results showed that the practical and hedonic quality of the CA could be improved. Compared to an industry benchmark (e.g., Hinderks et al., 2018), these scores indicate that it is in an early stage of development and needs further quality refinement. Although we considered and evaluated some initial criteria for ensuring general CA quality such as output formats, dialog control, or performance metrics (Lewandowski et al., 2023; Radziwill and Benton, 2017) in our initial design cycle, domain-independent quality attributes were not the focus.

**User Groups:** In line with our stakeholder analysis, we noted different archetypes of CA users. Some participants heavily relied on the tool, while others used it as a backup. This aligns with the two proposed user levels (beginner and advanced) and matches observation on different interaction modes in CA usage in the workplace setting (e.g., Gkinko and Elbanna, 2023). Additionally, our findings indicate a significant difference in the variance of motivation between the treatment groups, indicating that the experience of the CA users differed strongly from user to user. This finding could emphasize the existence of different learning or GUI preferences. This was also reflected in the qualitative responses where participants expressing less satisfaction referred to three main reasons: Either they found the CA to be somewhat circumstantial, preferred a more personal (i.e., human) support, or preferred other learning methods entirely.

### 9.5.2. Theoretical and Practical Contribution

With our research on designing a CA for data exploration in citizen science projects, we intend to advance the emerging practice of citizen science by guiding the designing of tools necessary for participation, thereby creating innovative and essential learning opportunities in data literacy. By going through and reporting the different steps of the DSR process, we provide a series of contributions for theorists and practitioners.

**Theoretical Contributions:** First, by connecting related literature on data literacy (e.g., D'Ignazio and Bhargava, 2016; Schüller et al., 2019), citizen science (e.g., Kloetzer et al., 2021; Tavanapour et al., 2019), and CAs (e.g., Okonkwo and Ade-Ibijola, 2021; Pérez et al., 2020), we provide valuable insights for future research projects at the intersection of these topics. Second, implementing a problem-centered design approach, we generate and evaluate design knowledge in the form of user needs and design principles on a CA assisting data exploration in the context of citizen science projects. Tailored to the inherent challenges of supporting citizen science projects and data exploration, the educational CA de-

sign presents a novel solution in this field. Thus, the contributed DSR knowledge can be classified as “Improvement” (Gregor and Hevner, 2013) and especially, DP1, DP2, and DP4 can be used as guidance for future design studies at the intersection of CAs, data literacy, and citizen science. In addition to descriptive principles, we also provide a design instantiation in the form of a prototypical CA (see Level 1, 2; Gregor and Hevner (2013)). In comparison to other research streams, giving insights in the design knowledge’s usefulness is a central component in DSR (Peffers et al., 2007). In an experimental study, we demonstrated the positive effect of the CA on task performance in data exploration, which is a promising sign for CA technology in citizen science projects. The experiments’ results, including the collected feedback and insights into user behavior, can be interesting starting points for further research.
**Practical Contributions:** Based on our theoretic findings and our artifact instantiation, our work provides several interesting implications for citizen science practitioners and system developers. While our evaluated design principles can inform the development of new and individualized citizen science support tools for the context of data exploration, the open-source code of our artifact implementation can be directly used and adapted by interested practitioners. As such, we provide citizen science practitioners with a simple and low-effort opportunity to support citizen participation beyond data collection in the proceeding analysis phase, thereby realizing learning opportunities. Likewise, for policymakers, we highlight how creative, non-formal teaching opportunities can be created and supported to address improving data literacy as a societal issue. Finally, beyond the citizen science use case, our CA design could be adopted by other stakeholders, such as educators or companies aiming to support their students or employees in improving their data exploration skills.

### 9.5.3. Limitations

As with any DSR study, the first cycle has limitations resulting from decisions made during the development of the artifact and the selected method for evaluation. Currently, the CA is limited in its outreach and applicability as it is only available in English and focuses exclusively on quantitative data, although it is intended to be dataset-independent. Regarding our approach to evaluation, we used an artificial use case instead of an actual citizen science project. This might have impacted participant motivation and perceptions of empowerment, as we envision such tools would enable actual citizen scientists to analyze data according to their inherent questions and ideas – rather than exogenous ones. Additionally, we observed participants’ ages to be lower than they usually are for citizen scientists (Mac Domhnaill et al., 2020). Moreover, participants were extrinsically motivated to participate in the experiment through monetary compensation, which would not be true for citizen scientists. Therefore, while the study provides exciting insights into the usefulness of CAs for data exploration with citizens, it needs further consideration of the circumstances in citizen science projects and additional evaluation cycles to make more generalizable statements about their usefulness in this context.

### 9.5.4. Future Work

This study's results and limitations give insights into possible future research streams. On the one hand, qualitative and quantitative feedback from the participants on the CA suggests a further refinement of its functionality and design. In this context, features that may enhance our artifact's hedonic quality should be tested. As a starting point, domain-independent quality attributes related to anthropomorphism, affect, or accessibility (Lewandowski et al., 2023; Radziwill and Benton, 2017; Seeger et al., 2021) could be further considered for the specific use case of our artifact. On the other hand, further analyzing the lack of correlation between perceived learning and task performance and ways to circumvent this are interesting starting points for future research. For instance, it has been shown that feedback positively impacts students' perceived learning (e.g., Chan and Ko, 2021; Eom et al., 2006). Thus, it could be interesting to investigate whether the indication of the actual performance of a participant after the analysis task would change the results of the perceived learning and the correlation between the variables in our experiment. Additionally, it could be interesting to evaluate whether the Dunning-Kruger effect could explain the differences in perceived learning and task performance between the groups. Kruger and Dunning (1999) showed that people who lack knowledge or skills are likelier to be unaware of this in self-assessment. Instead, they overestimate their skills compared to participants with more knowledge. Applied to our experiment, this could mean the amount of knowledge presented to the CA users could have negatively affected the user's assessment of aspects such as his or her feeling of empowerment or learning through the analysis activities, compared to a user not exposed to this knowledge. A third research direction could focus on evaluating the CA in a real-life citizen science environment, including assessing who can use it and who is excluded. This perspective would be critical as citizen science strives for inclusivity (National Academies of Sciences, Engineering and Education et al., 2018; Sorensen et al., 2019), and the under- or over-representation of particular groups can have negative consequences for project outcomes (Sorensen et al., 2019). Therefore, understanding who is excluded from our artifact and what alternatives could be created would be indispensable.

## 9.6. Conclusion

Inequalities in data access and literacy pose a risk to the individual and society as a whole. In this work, we therefore investigated the use case of citizen science as a means to empower citizens in accessing and working with data. We have presented results from the first cycle of a DSR project targeting the development of a CA to support data exploration in citizen science projects. Following the six steps for a problem-centered design process by Peffers et al. (2007), we have approached this challenge by structuring and elaborating on the problem space and its associated stakeholders, eliciting requirements and translating them into design principles for a solution artifact and finally presenting and testing a prototypical implementation of this artifact. The result of our first design cycle is a CA for data analysis activities offering flexible support on demand to multiple users in parallel. For inexperienced users, the tool provides seamless guidance through data analysis by providing dataset-independent knowledge and tips, allowing users to decide how deeply they want to dive into a particular topic. Advanced users can use a question-and-answer process to ask questions about data analysis freely, thus enriching only their knowledge with the CA and controlling the analysis process.

In its current state, the CA shows high potential for transferring required data literacy to citizens, enabling them to perform better in analytical tasks. Qualitative user feedback shows that multiple citizens perceive the tool's support as enjoyable and useful and point out potential application fields. Harnessing the advantages of its easy, resource-efficient provisioning, the usage of CAs in citizen science projects seems promising and could positively promote equitable access to data-driven knowledge. However, identified challenges, such as the participants' motivation, feeling of empowerment, and perceived learning effect, could not be solved adequately by the CA. This indicates that further research is necessary to refine the CA and its usability, which we intend to accomplish in a proceeding design cycle.

# Part V.

# Finale

**Abstract Part V**

*Part V, "Finale", concludes this dissertation with a discussion of the presented work and a concluding note. The various research endeavors of this dissertation have individually contributed multiple theoretical and practical contributions to the field of platform design in digital involvement. Collectively, they establish foundations for the cross-domain navigation of participatory practices and for dealing with the implications of datafication in its formats and platforms. As a result of this dissertation, this part proposes an overarching design and research framework of digital involvement in a datafied society. Furthermore, it discusses the limitations of the studies and the dissertation's collective approach. Together with the proposed research framework, they form a practical starting point for future research on digital involvement in a datafied society which is elaborated on in this part.*

# 10. Discussion and Conclusion

This chapter discusses the key contributions of this dissertation by first briefly answering the individual chapters' research questions, and second, integrating the components into a comprehensive design and research framework for digital involvement in a datafied society. In a third and fourth section, it addresses the main limitations of the dissertation and suggests directions for future research. Finally, the chapter closes with a concluding remark.

## 10.1. Contributions

This work examines the design of digital involvement in a datafied society, focusing on two key contributions to the field. First, recognizing the increasing convergence of public and private, economic and political sphere in datafied societies, it establishes theoretical and practical foundations to rethink traditional research streams in the field of civic participation in a joint consideration of DIP. The DIP taxonomy, along with its instantiations presented in Parts I and II of this dissertation, offer essential design artifacts to facilitate discussions on involvement practices, embrace their diversity, and guide practitioners effectively. Second, theoretical and practical foundations for dealing with the implications of datafication in formats and platforms of digital involvement are established. The studies and design artifacts presented in Parts III and IV of this dissertation provide insights into how digital involvement can be designed to inclusively communicate complex data and enable more people to access, understand, and use data. In the following, the individual research questions that inform these two overarching research contributions will be discussed.

The foundational part of this dissertation aims at defining digital involvement by answering the research question:

> **Research Question 1:** *What are the key characteristics to describe and distinguish DIP?*

Through a DSR-based taxonomy development process, this research question is answered by carving out 19 dimensions to describe key design characteristics of involvement practices. They consider the participation process, its individuals, and the digital infrastructure, enabling the description and differentiation of DIP in the field. The taxonomy provides a theoretical contribution by bringing together existing academic frameworks of participatory practices (e.g., Estellés-Arolas et al., 2015; Haklay, 2013a; Van Dijk, 2012) and organizing them into an overarching concept. In this way, it creates a theoretically grounded foundation for the emerging concept of DIP. In practice, the taxonomy enables the joint discussion and comparison of citizen science, e-participation, and crowdsourcing practices. This makes it easier for practitioners to engage in cross-domain exchange, share best practices, and approach the design and evaluation of participatory projects in a structured way.

Building upon this initial work, a consideration of how to further facilitate the navigation of the design of digital involvement for practitioners continues the DSR project. The cycle aims at answering the research question:

> **Research Question 2:** *How can we use the DIP taxonomy to capture design knowledge and make it available to domain practitioners?*

This research question is answered in the form of two evaluated design artifacts: the DIP web application and DIP archetypes. The web application consists of an interactive guide through the design knowledge captured in the DIP taxonomy, and a classification, as well as data collection tool. The DIP archetypes consist of eight design archetypes of participatory practices, uncovering design patterns between commercial, hybrid, and digital, non-commercial DIP projects. As a theoretical contribution, the research enhances the evaluative rigor of the DIP taxonomy by complementing the first cycle's qualitative evaluation with quantitative methods. The operationalization of the taxonomy enables a deeper understanding of their development over time. Practically, the web application facilitates practitioners' access and utilization of the DIP taxonomy, enabling the participatory creation of a DIP database for evidence-based research. The DIP archetypes offer a low-threshold anchor to break down the diversity of DIP and make differences between practices tangible. Combined, the two artifacts promote a continuous exchange between research and practice that is urgently needed in the rapidly developing field of DIP.

Finalizing Part II of this dissertation, the previous conceptual and practical works are put to use by applying the design artifacts to the realm of citizen science. A practitioners' study aims to answer the research questions:

> **Research Questions 3:**
>
> a. *What design characteristics do citizen science projects exhibit in an overarching taxonomy for participatory concepts and how are their preferences distributed?*
>
> b. *Given the diversity of citizen science practices, to what extent can distinct design clusters be identified?*
>
> c. *Regarding the multidisciplinary nature of citizen science, is there an association between a project's disciplinary focus and its assignment with specific design clusters?*

To answer these research questions, project design descriptions are collected from practitioners in Germany using the DIP web application. The research identifies over- and underrepresented design traits in citizen science compared to other participatory practices and distinguishes projects in 'Crowdsourcing Research' and 'Participatory Research'. These clusters represent two distinct streams of practice that vary across disciplines conducting citizen science projects. The study advances the theoretical discourse on citizen science by quantifying projects' design practices and how they are interconnected. In doing so, it supplements established typologies, which are primarily based on the degree of participation (e.g., Haklay, 2013a; Shirk et al., 2012). As a practical contribution, the study offers the German project landscape the opportunity to review self-imposed goals and principles while providing practitioners orientation for (re-)conceptualizing their projects. For this dissertation, the study shows how applying the design artifacts from Parts I and II can support self-reflection within a particular participatory practice while promoting perspectives beyond one's field.

Building on the unifying concept of digital involvement outlined in Parts I and II, the role of data in digital involvement and the design of tools simplifying complex participation scenarios are investigated. This part deals with theoretical foundations for considering the role of data and data literacy in society and explores the perceptions of citizens regarding the existence and importance of data (literacy) as well as the requirements for supporting tools. Following this, the part empirically explores how different data foundations and their representations impact digital involvement. Focusing on use cases in participatory urban planning, the following research questions are addressed:

> **Research Question 4:** *How does the usage of XR affect participants' literacy in the context of participatory urban planning, and what specific factors are crucial for successful XR participation formats?*

and

> **Research Question 5:** *How do properties of 3D visualizations affect their suitability for processes of e-participation in the context of urban planning?*

To answer this research questions, current experimental studies on XR in participatory urban planning are analyzed and documented according to their design, use, and evaluation of the support tool. The analysis demonstrates a lack of comparability and a gap in the understanding of the precise effects of individual factors on the suitability of data representations for inclusive participation. Consecutively, three factors are systematically evaluated in experimental studies. As a theoretical contribution, the lab study offers first insights into the effects of varying display types and support levels on a set of participants' literacies. Particularly noteworthy are the differences between observed effects on actual and perceived knowledge gains for VR data representations. These results evidence the importance of a differentiated tool evaluation for inclusive data communication. Furthermore, the online experiment robustly demonstrates how the variations in level of detail, realism, and resolution of 3D visualizations affect participants' motivation, trust, mental load, and perceived usefulness for participation. Insights argue for the importance of high-detailed, but especially high-quality, geospatial data in participatory urban planning. Thus, they provide a baseline to the theoretical discourse on e-participation and immersive systems, with regard to system design and evaluation, and the comparability of different study results. Practically, practitioners can utilize the research results to decide between different visualization approaches when communicating complex data given their objectives for the participatory urban development processes.

Given the central role of data and, thus, data-related skills for digital involvement and, more broadly, for participation in society, Part IV of this dissertation proceeds to explore the design of assistance and learning in digital involvement. It first aims to answer the research question:

> **Research Question 6:** *How can crowdsourcing platforms support the learning and skill development of crowdworkers?*

To answer this research question, five different teaching and learning approaches are identified in the crowdworking literature and mapped to their corresponding feature implementations and task types. The results of the literature review show that crowdwork allows flexibility in the combination of learning approach and task type. In addition, there is scope in the complexity of the implementation of learning functions and the integration of learning into the crowdwork task. A platform review and a user survey show how Kaggle, as an active crowdwork platform example, practically implements four of the five identified teaching and learning approaches for teaching data literacy. It further demonstrats that users perceive Kaggle as a valuable learning environment. As theoretical contributions, the chapter synthesizes the current research state on crowdworkers' learning, enabling the identification of research gaps and structuring further analysis of platforms and tools in this regard. It also provides structured knowledge on the ways learning is facilitated on Kaggle. Both insights can be used also by practitioners to obtain guidance on how to design and evaluate appropriate learning opportunities on their platforms or even optimize the usage of existing features for learning.

Expanding the research on operating involvement platforms, the following chapter addresses the research questions:

> **Research Questions 7:** *What multi-project citizen science platforms exist, and how do they support the conduction of citizen science projects? What are the potentials and challenges for promoting citizens' data literacy through digital citizen science?*

Through conducting a structured artifact review, 16 multi-project citizen science platforms are identified and systematically described based on aesthetics, usability, data standards, support, and communication features. Complementing these insights with qualitative interviews enables the derivation of three main positions toward current and future potentials for promoting data literacy on citizen science platforms. These comprise the narrow focus of current platforms on the participatory collection and analysis within the broad spectrum of data literacy; the advanced state of current platforms in terms of target group-oriented and user-friendly design; and the need for innovative support and communication mechanisms.

As a theoretical contribution to the current discourse on learning in citizen science, the chapter offers a new perspective by shifting the focus from learning opportunities in individual projects to the design of multi-project platforms. This perspective supports practitioners in choosing a suitable technical infrastructure for their needs and helps researchers identify structural gaps.
Ultimately, the final research chapter aims to address one such gap by answering the research question:

> **Research Question 8:** *How should a conversational agent be designed to support data exploration in citizen science applications?*

Using a DSR approach, five design principles for a CA are carved out. The principles entail design considerations on guiding the user through a data exploration while providing opportunities for free usage and personalization. They further consider handling of user mistakes and questions, and providing appropriate educating material. The principles are implemented as a functional CA artifact, demonstrating how to support citizens in both a guided and free mode during the data analysis process. The experimental evaluation demonstrates a positive effect on participants' task performance in scientific data analysis, opting for the effectiveness of four of the five design principles. Theoretically, the research contributes design knowledge based on the intersection of the literature on citizen science, data literacy, and CAs in the form of user stories and design principles. Furthermore, its instantiation in the form of the open-source chatbot artifact offers practical usage opportunities to citizen science practitioners. Concluding Part IV, the work demonstrates how creative, non-formal teaching opportunities can be developed within involvement projects to address the lack of data literacy as a societal issue.

In summary, this dissertation explores the design of digital involvement in a datafied society: Parts I and II establish the common research field of digital involvement and provid tools for its navigation, while Parts III and IV examine how data (literacy) can be mediated in digital involvement. In doing so, this dissertation intersects multiple sub-fields of IS research, including crowdsourcing, digital government, e-learning, CAs and XR, while integrating interdisciplinary knowledge in fields such as citizen science or data literacy. By addressing the individual research questions, the dissertation delivers multiple theoretical and practical contributions, including a systematic documentation of literature and platform landscapes, the acquisition of empirical findings, and the development and evaluation of new design artifacts.

## 10.2. A Design and Research Framework for Digital Involvement in a Datafied Society

In this section, the research contributions presented in this dissertation are combined and sorted into an overarching structure. To this end, a design and research framework for digital involvement in a datafied society is proposed. The framework integrates the foundations for defining and navigating DIP from Parts I and II (i.e., the DIP taxonomy) and the considerations of how to integrate data-related needs into the formats and tools of DIP from Parts III and IV (i.e., approaches and findings from the experimental, review and design studies) into a comprehensive, visual inquiry tool. In the following, the concept of the visual inquiry tool is first introduced, and then the design and research framework, and its intended utilization are presented.

**Conceptualization**: The concluding framework of this dissertation aims to support strategic research and dialogue on the planning, assessment, and enhancement of DIP. Therefore, the framework is conceptualized as a visual inquiry tool. Visual inquiry tools gained momentum since Osterwalders and Pigneur's approach to business model generation (Avdiji et al., 2020; Osterwalder and Pigneur, 2010). Their "Business Model Canvas" is a widely used visual inquiry tool for practitioners to plan, assess and enhance their business model (Osterwalder and Pigneur, 2010). It supports strategic planning and managerial decision-making by enhancing dialogue, improving communication, and exploring ideas while guiding through the consideration of key aspects of business models (Osterwalder and Pigneur, 2010). It is, thus, a suitable starting point for the conceptualization of the concluding framework of this dissertation.

Visual inquiry tools should encompass a conceptual model, instantiate it as a shared visualization, and enrich it by directions to use (Avdiji et al., 2020). For the conceptual model of the design and research framework, the insights on DIP discussed in previous chapters are consolidated and structured within a classical development process: 1) the identification of requirements, 2) the design and implementation of tools, and finally 3) the evaluation of the developed artifact. For the case of digital involvement, this translates to 1) the specification of the involvement project, 2) the implementation of the digital support tool, and finally 3) the process and tool evaluation. Transforming the conceptual model to a shared visualization, the components are visually arranged. Following best practices, the components of the conceptual model of the design and research framework are represented as empty design spaces and arranged in a way that they reflect the procedural associations between them (Avdiji et al., 2020). Finally, for directions to use, designers should ensure that users can explore the problem space, generate solutions and critically review them (Avdiji et al., 2020). The design and research framework includes guiding directions to supports this approach across multiple layers. In the outer layer, guiding the user through the structure of the classical development process inherently reflects exploration, implementation and review. In the inner layer of project specification, users are guided through a sequential progression of blocks, where closely related design dimensions are grouped together. Unlike the Business Model Canvas, which allows flexibility in starting points (Osterwalder and Pigneur, 2010), the archetype studies in Chapters 3 and 4 suggest that design decisions have varying levels of centrality, warranting a predefined sequence. To account for potential conflicts between decisions in later and earlier blocks, the design and research framework includes backward arrows to encourage iterative critical evaluation.

**Design and Research Framework:** The design and research framework for digital involvement in a datafied society integrates the main contributions of this dissertation structured along the specification of the participation and system design, and their evaluation (see Figure 10.1).

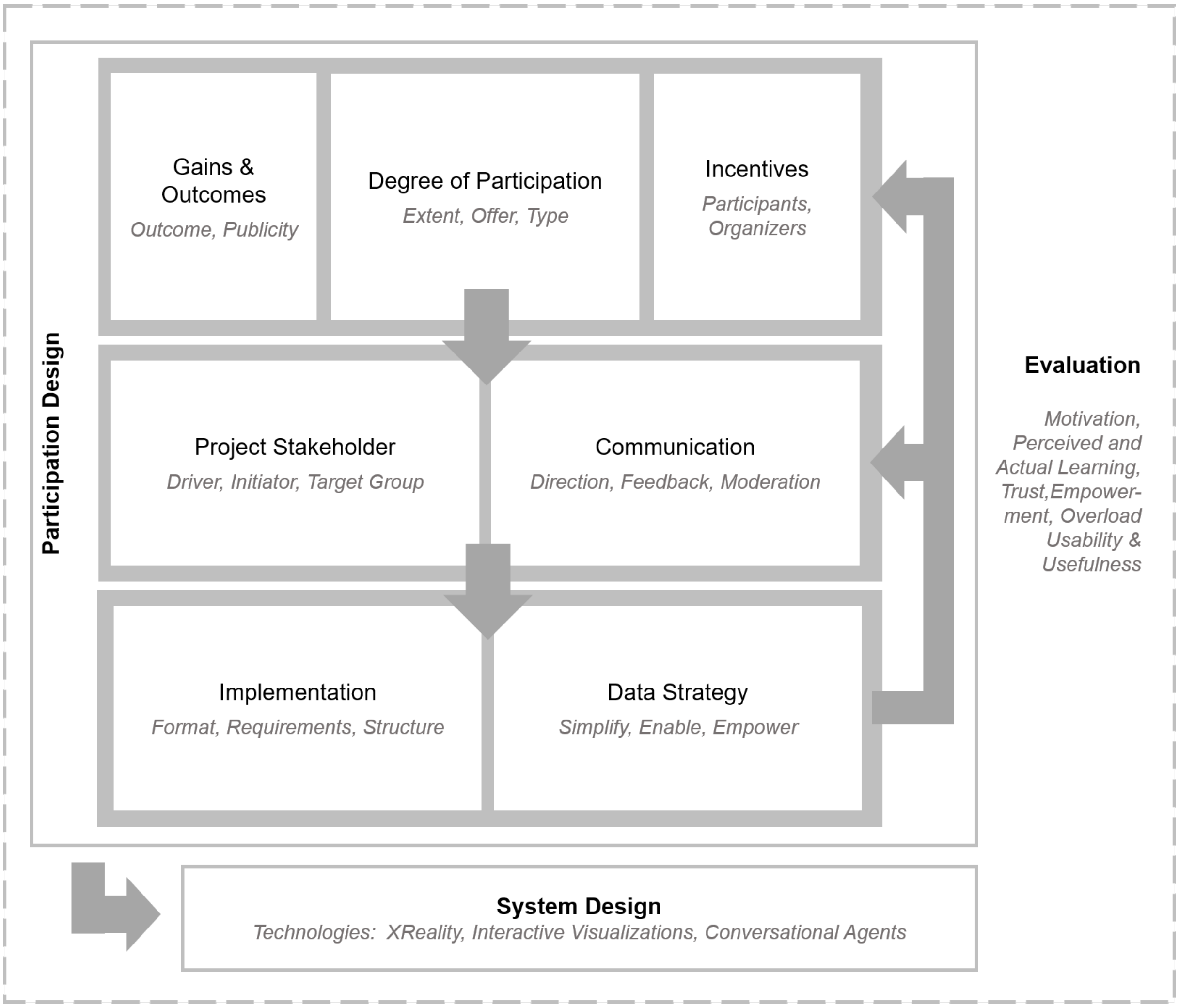


Figure 10.1.: Design and research framework for digital involvement in a datafied society.

In Parts I and II of this dissertation, the DIP taxonomy is developed, grouping key design dimensions of digital participation processes. As such, the taxonomy's individual dimension form the basis of involvement project specification in the framework. Additionally, the field "Data Strategy" is added as a direct result of the findings from Parts III and IV of this dissertation. As demonstrated in the respective chapters, the consideration of an appropriate data strategy in digital involvement that either targets the simplification of relevant data, enables the usage of complex data, or empowers the learning of participants is of central importance in a datafied society. Thus, the framework integrates this consideration in the design of the participation processes. The specification of the involvement project serves as input for developing appropriate digital infrastructures, as showcased for example in Chapters 6 and 9 of this dissertation. Digital infrastructures can comprise a range of tools, depending on the participation context and the technologies used. For instance, in this dissertation, digital platforms and mobile applications enriched with VR components, interactive visualizations, or CAs are considered. Finally, due to the intertwined nature of the participation process and the digital tool in DIP, the evaluation should evidently comprise both aspects. Given the differing effects of design decisions on various success measures that are observed, e.g., in the experiments in Chapters 6 and 9, the framework aims to encourage a nuanced evaluation. Thus, it highlights the range of success indicators that are evaluated in this dissertation, such as participants' motivation, empowerment, or perceived and actual learning gains.

Summarizing the results of this dissertation, the framework is intended to provide both practitioners and researchers with a first overview of the field. As a design framework, it follows the structure of a visual inquiry tool, proving orientation to practitioners planning and conducting DIP. Practitioners can use the framework analogously to the widely used Business Model Canvas, which supports easy adoption by economic and civil society actors. As a research framework, it structures key research areas of the field aiming to stimulate the academic community. The individual building blocks of the framework and their relationships each represent starting points for further exploring different aspects of digital involvement in a data-driven society. The framework also methodically reflects on the necessity for IS researchers to root system design in the individual objectives of the involvement project and critically evaluate a variety of research dimensions for the interaction of process and digital infrastructure. Being based on the findings of this dissertation, the contents of the proposed framework provide a solid foundation for both researchers and practitioners. However, its instantiation as visual inquiry tool should, as every design artifact, be demonstrated and evaluated in future work to ensure adoption by the target audiences.

## 10.3. Limitations

The individual studies presented in this dissertation, along with the work as a whole, have inherent limitations and shortcomings that should be considered when interpreting and applying the results. This section reflects on the key limitations of the individual studies, before extending to the broader context of this dissertation.

**Limitations of the DIP Taxonomy and its Instantiations**: Regarding the development and application of the DIP taxonomy and its instantiations in Chapters 2- 4, a key limitation lies in the exclusively descriptive nature of the design analysis. The studies carried out prevent conclusions about the effectiveness of individual or collective design decisions. Instead, they only provide insight into common practices that may have been chosen for various reasons.

Furthermore, the focus on the three DIP sub-areas, crowdsourcing, citizen science, and e-participation, is a limiting factor to the comprehensiveness of the taxonomy and its artifacts. The umbrella term digital involvement seeks to connect a variety of digital, participatory practices. However, multiple influential streams (e.g., Open Source, Open Science, or Open Innovation Practices) are yet to be integrated into the framework. Within the three evaluated sub-areas, additionally, only a finite number of projects are considered in the artifacts' application, either from an external perspective in Chapters 2 and 3 or from an internal perspective in Chapter 4. In the external evaluation conducted by researchers, consistency measures can be applied to the project characterizations. However, the descriptions rely solely on (potentially incomplete) public project information. In the internal evaluation conducted by project initiators, the obtained project descriptions are more precise; however the subjectivity of these descriptions must be assumed. As a consequence, both approaches limit the robustness of the cluster results.

**Limitations of Experimental Lab and Online Studies**: Regarding the experimental studies presented in Chapters 5 and 9 of the dissertation, factors such as sample size and real-life proximity are relevant constraints of this work. The VR laboratory experiment, for example, involved only a small number of participants and was originally intended as a preliminary study, limiting its ability to provide definitive findings. Due to changes within the research project, a larger study was conducted only for the 3D models factor. Similarly, the online experiment evaluating the CA focuses on qualitative insights, resulting in a smaller sample size and limiting the generalizability of the quantitative findings. All three studies employ fictional use cases and took place in experimental settings, which may have influenced results and participants' perceptions and behaviors. For example, the VR study relies on a homogeneous student panel and a low-complexity use case, which are both unrealistic in most urban planning scenarios. The sample's heterogeneity and the complexity of the use case were increased in the conducted online studies. However, in both the 3D visualization and CA study, the artificial nature of the use cases may have impacted key elements of the experiment, such as the participants' assessment of suitability for different levels of participation, or the motivation and empowerment in citizen science.

**Limitations of the Platform Reviews**: Finally, key limitations of the studies in Chapters 7 and 8 stem from the restricted, external access, and insights into the evaluated platforms. This applies both to the review of Kaggle and the review of the 16 citizen science platforms. For the reviews, researchers visited the platforms and created when possible user profiles, collecting information through the lens of a normal user. Platforms, however, can have functionalities for specific user groups or even hidden mechanisms, implying that a complete platform review can only be achieved with additional insights from the operators. The restriction to normal user accounts also limited the possibilities of contacting other users on the platform to participate in the Kaggle user survey. It was, therefore, not possible to obtain a representative sample which limits the expressiveness and transferability of insights into user experiences.

**Broader Limitations of the Dissertation**: The limitations of the individual studies also point to broader shortcomings within the context of this dissertation. First, one of the key motivations driving multiple endeavors of this dissertation was to ensure the relevance and applicability of research on DIP for practitioners. Therefore, a notable limitation lies in the fact that no involvement project was actively accompanied during its planning stages, nor was an evaluative field study conducted to assess real-world impacts. Although the reviews of projects and platforms and the conduct of expert interviews, lab and online experiments provide valuable insights, they remain inherently constrained by their controlled setting and reliance on external perspectives. Without the internal monitoring of ongoing projects or the integration of field-based evaluations, the findings of this dissertation can only offer a partial view of the dynamics at play in real-world settings.

A second, broader limitation is introduced in the fact that concepts or measurement instruments were not always consistently applied across the studies of this dissertation. This particularly concerns presented experimental and review studies in Parts III and IV. Since the evaluation of DIP formats and tools with regards to their ability for communicating complex data, skill-enabling participants or fostering learning is still an under-researched area, there is a lack of established concepts to demarcate dimensions of interest, and of research models to measure their successful fulfillment. Therefore, as a starting point, this work explores a variety of dimensions in the different research studies at the expense of a their closer interlinking and, thus, a strengthened coherence and comparability of the results between them.

Thirdly, the design of digital involvement in a datafied society is a comprehensive topic that spans several research streams addressed in this work. While multiple participatory approaches (e.g., crowdsourcing, citizen science, e-participation), technologies (e.g., digital platforms, XR, interactive visualizations, CAs), and use cases (e.g., urban planning, data science) are considered, not all facets of this complex topic can be fully addressed within the scope of this work. Alongside research endeavors that are initiated but require further investigation, certain areas of digital involvement are not examined in this dissertation. In addition, conducted studies examine primarily the Western European society. Therefore, the expansion of the local focus offers numerous additional opportunities to extend the research presented.

### 10.4. Future Work

The design of DIP in a datafied society remains an evolving and multifaceted research field. While the insights from the studies presented in this dissertation provide a foundational framework, summarized in the design and research framework, numerous avenues for further exploration emerge. Based on the findings and the limitations of this dissertation, the following key areas can be identified in particular:

**Exploring the Effectiveness of DIP Design**: The presented design and research framework, including instantiations of the taxonomy outlined in Chapters 3 and 4, present descriptive work of the status quo of involvement design. Expanding this work to considerations of effectiveness and ethical or policy recommendations would contribute a valuable next step in the guidance of practitioners of digital involvement. This may include future work on the individual components of the framework, such as incentives or feedback structures, but similarly on understanding the interplay between different components. Ultimately, a link between different DIP archetypes and their produced success would empower a targeted project design. A major challenge in following this research path is, however, the definition and measurement of success in projects that are as heterogeneous as DIP. This leads to the second key area for future research.

**Establishing Robust Evaluation Criteria**: A key aspect of the design and research framework is the evaluation component. This dissertation discusses a variety of dimensions of interest for the evaluation of digital involvement in a datafied society that were identified in the literature. Thereby, the studies provide an important understanding of the multifaceted impacts of design decisions on participants, and the complexity of successfully implementing DIP. However, as noted in the limitations, this explorative approach was chosen due to the absence of established concepts or research models and comes at the expense of consistency and comparability across the individual studies. This challenge was also identified in the examination of related work, (cf. Impacts of immersive systems in Chapter 6), which in turn complicates further research and development in DIP. To advance the emerging research field, therefore, the establishment of a coherent research model for the evaluation of DIP would be valuable. This dissertation creates an initial starting point for this future work, suggesting how different research dimensions can be measured and how they may or may not interact with each other. Nevertheless, further research is needed to develop and test an appropriate research model that is uniformly applicable in DIP. To this end, it could also be promising to explore the extent to which established research models from other areas can be transferred to DIP.

**Assessing Data Literacy Competencies to Optimize Data Strategies**: Parts III and IV of this dissertation argue for the societal relevance of data literacy and, consequently, the necessity of incorporating strategies for enabling learning into participative processes. While a considerable body of literature stresses a general lack of data literacy among citizens (e.g., Bhargava et al., 2016; Debruyne et al., 2021), representative studies on citizens' data literacy levels remain sparse, especially in Germany. Understanding the specific challenges faced by different societal groups would allow for optimization of data strategies within participation processes and, thus, better inform the development of supporting infrastructure. As such, future work should aim to fill this gap, systematically mapping competencies in society to identify precise deficiencies and tailor learning opportunities accordingly. This could foster increased inclusivity and empowerment in digital participation processes. In this regard, the results of the DaLi project's first German representative data literacy study, as mentioned in Chapter 5, will certainly be promising for future work in Germany.

**Engaging with Use Cases and Applications in the Field and Sustaining Continuous Channels Between Research and Practice**: Being a limitation of this work, also the literature review in Chapter 7 shows that research on participants' learning and support tool evaluations frequently focuses on experimental settings and artificial use cases. While controlled experiments support internal validity in research, for example by allowing individual effects to be isolated or specific components to be tested,

their findings have limited generalizability. Monitoring real DIP projects in the form of field studies would allow researchers to observe challenges and success in situ and ensure practical applicability, thereby, strengthening the external validity of DIP research. In this regard, limitations of this dissertation stem from a lack of suitable use cases or limited access to operating participation platforms. As such, establishing better partnerships between researchers and platform operators, as well as project initiators in the economy, politics, and civil society is an important prerequisite to further pursue this research avenue.
Such partnerships could also offer critical feedback loops that refine both theory and practice in dynamic research fields. However, to harness mutual exchange, research needs to keep open ways in which practical feedback can be continuously incorporated. A first step in this direction is taken with the DIP web application presented in Chapter 3, however, further research is necessary to investigate how continuous channels can be refined and sustained.

**Emerging Trends and Additions to the Framework**: The rapid pace at which digitization and datafication evolve and change our modes of working, communicating and informing ourselves makes the design of digital involvement a dynamic and demanding research field. Looking at the example of our work in Chapter 9, focusing on a CA for data exploration, the rapid development of Generative AI and the introduction of ChatGPT as a public support tool ten months after the presented experiment showcases how quickly, new technical opportunities develop, that can improve or threaten digital involvement. Due to its great success in supporting users to access complex knowledge, code, or analyze data (e.g., Shen et al., 2024; Zheng, 2023), Generative AI is certainly a promising technology for future applications of digital involvement, which should be explored. Simultaneously, it aggravated a digital phenomenon that has been ongoing for years, especially in online social networks: the spread of dis-, mis-, and malinformation (Weinhardt et al., 2024). Since the definition of digital involvement excludes settings of undefined participation contexts such as social media, this topic is largely neglected in the framework, together with other phenomena of societal polarization. However, they are likely to influence digital involvement practices in a datafied society in multiple ways. For example, societal polarization could affect the prerequisites of DIP projects in terms of motivated and trustful participants or the conduction and outcomes of DIP projects through the spread of disinformation. On the other hand, initiatives such as the "forum against fakes" or the "ichbinhier" movement show that digital involvement can be a promising approach in the fight against disinformation, hate speech and societal polarization. Therefore, Generative AI and disinformation are two examples of emerging trends that impact digital involvement but are yet missing in the design work on digital involvement in a datafied society.

### 10.5. Concluding Remarks

This dissertation deals with the design of digital involvement practices and tools in light of an ongoing datafication of our modern society. However, throughout the journey of this dissertation, numerous events occurred in the global political landscape that cast doubt on the future of democratic participation and the joint development of solutions to our social challenges. Military conflicts such as Russia's war of aggression against Ukraine or the resurgence of the Middle East conflict, but also the worldwide strengthening of right-wing-extremist parties and movements, and their entry into parliaments and governments, are dividing societies at national and international levels. In such times, it becomes increasingly difficult to believe that (digital) involvement projects can bring society together and create meaningful change — especially when looking at a digital space that is currently highly fragmented and controlled by political and economic actors.

Yet, it is precisely in these times that small initiatives may inspire and encourage us. Projects on mit:forschen! or Zooniverse are uplifting examples of volunteers dedicating time and efforts to preserving biodiversity or advancing health research; On platforms like Crowdify or Kickstarter, countless social projects and innovative products are funded by the collective efforts of users; Projects on Go Vocal, Adhocracy+, or Dialogzentrale demonstrate meaningful and inclusive societal discourse on local and regional matters. These and many more DIP projects exemplify social engagement, collaboration and integrative solutions — endeavors that are urgently needed and should be pursued in manifold ways in the face of major crises and societal challenges.

As a result, research on digital involvement remains crucial, especially in today's times, and the discipline of IS in particular should be involved in the discourse on our democracy (Weinhardt et al., 2024). This dissertation provides a scientific foundation for cross-domain research on processes and platforms of digital involvement, expanding the current research focus to include the context of datafication in society. Besides scientific objectives, throughout the course of this work, engagement in the practical discourse beyond the academic community was a concern. Involvement in economic, political and public events such as the *Hannover Messe*, structured dialogues of the *Federal Ministry for Digital and Transport (BMDV)* and the *Federal Ministry of Labour and Social Affairs (BMAS)*, *Bunte Nacht der Digitalisierung*, *Tage der Demokratie* or *Girls' Day* created awareness for the research topic and enabled the sharing of research insights. Through the conduct of multiple exchange formats with political, economic, and civil society actors, including three Round Tables on the works of this dissertation and two demonstrations in political visits from the *US Consulate General* in Frankfurt or the *Ministry of Science, Research and Arts* in Baden-Württemberg, it was aimed to bridge the gap between research and practice, gathering impulses and feedback for this work. As IS researchers in the field of digital democracy, it is essential to look beyond the confines of our discipline and academic research. Enabling participation in one's own work — like any involvement project — is complex and challenging to design. However, it brings valuable benefits, both for others and for oneself.

# A. Appendix to Chapter 10

| Category | Question | |
|---|---|---|
| Getting started | 1 | How many variables are in the dataset? |
| | 2 | What can you use the variable PassengerID for? |
| Cleanup | 3 | Perform a data cleaning. Which observations would you mark as conspicuous? |
| Analysis | 4 | Calculate how much a ticket typically cost for Titanic passengers. * |
| | 5 | Which figure did you calculate to get the typical ticket price? * |
| | 6 | What influence would you expect the variable Gender to have on the variable Passenger Class? ** |
| Knowledge Questions | 7 | What is a binary variable? |
| | 8 | What does the variance measure? |
| | 9 | What type of variable is the variable Passenger Class? |
| | 10 | You want to plot the variable Age. What plot type do you chose? |
| | 11 | What should you look out for when interpreting plots? |
| | 12 | What method could you chose to statistically test whether the variable Age (predictor) has an effect on the probability of survival? |

Note: All tasks have been free text questions, except for * and ** which have been selections or offered an optional cross table respectively.

Table A.1.: Experiment tasks.

| Construct | Item |
|---|---|
| Data Analysis Knowledge | I have knowledge in the field of data analysis. |
| Software Skills | I'm experienced in working with a spreadsheet software (e.g., Excel). |
| SMQ adapted from Glynn et al. (2011) | Learning Science is interesting. |
| | I'm curious about discoveries in science. |
| | I'm sure I can understand science. |
| | My career will involve science. |
| | Learning science will make my life more meaningful. |
| | Knowing science will give me a career advantage. |
| | Learning science will benefit my career. |
| Citizen Science Concept | I have heard about citizen science before. * |
| Dataset Experience | I have worked with datasets before. * |

Note: All items have been measured on 7-point Likert scales, except for *, which were binary variables.

Table A.2.: Prequestionnaire.

| Construct | Item |
|---|---|
| Open Questions | How did you go about processing the tasks in today's data analysis? What support did you use? * |
| | Where would you have liked more support in the data analysis process? How could a support have looked like? * |
| | Where could you imagine the usage of DataBot? For whom could it be interesting? * |
| Perceived Learning adapted from Alavi et al. (2002) | I became more interested in exploring scientific data. |
| | I gained a good understanding of how to familiarize myself with scientific data. |
| | I learned to identify central ideas in data analysis. |
| | I found today's analysis activity to be a good learning experience. |
| | Given a choice I would take part in a data exploration activity similar to today's session. |
| Motivation adapted from Center for Self-Determination Theory (2022) | I thought today's analysis activity was a boring activity. |
| | Working on the data exploration tasks did not hold my attention at all. |
| | I would describe today's analysis activity as very interesting. |
| | I enjoyed working on the data exploration tasks very much. |
| Empowerment adapted from Kim and Gupta (2014) | I have mastered the skills necessary for exploring scientific data. |
| | I am confident about my ability to familiarize myself with scientific data. |
| | I am self-assured about my capabilities to gain insights from scientific data. |
| | Based on my data analysis activity, my impact on what happens in a citizen science project is large. |
| | Based on my data analysis activities, I have significant influence over what happens in a citizen science project. |
| | Based on the chatbot usage, I have a great deal of control over what happens in a citizen science project. |
| | Being able to gain insights from scientific data is very important to me. |
| | My ability to explore scientific data is personally meaningful to me. |
| | Being able to familiarize myself with scientific data is meaningful to me. |
| | I have significant autonomy in determining how I gain insights from scientific data. |
| | I can decide on my own how to go about exploring scientific data. |
| | I have considerable opportunity for independence and freedom in familiarizing myself with scientific data. |
| UEQ adapted from Schrepp et al. (2017) | To me the bot was … |
| | ...obstructive/...supporting |
| | ...complicated/...simple |
| | ...confusing/...clear |
| | ...inefficient/...efficient |
| | ...boring/...exciting |
| | ...uninteresting/...interesting |
| | ...conventional/...innovative |
| | ...traditional/...novel |

Note: All items have been measured on 7-point Likert scales, except for *, which were free text answers. The CA has been introduced as "DataBot" to the participants.

Table A.3.: Main questionnaire.

| Author | Context | Contribution | Method | Sample Size |
|---|---|---|---|---|
| Janssen et al. (2020) | CAs (general) | Overview of design elements for CAs regarding the topics, intelligence, interaction and context | Taxonomy Development | Assessment of 103 chatbots from 23 domains |
| Rapp et al. (2021) | CAs (general) | Overview of prior research on text-based CAs: topics, user experience, technology acceptance, challenges | Structured literature review | 83 articles from 3 digital databases |
| Diederich et al. (2022) | CAs (general) | Overview of the status quo of CA research regarding user interaction, context, agent design, CA perceptions and outcomes | Structured literature review | Review of 262 studies |
| Kvale et al. (2021) | CAs for customer service | Derivation of determinants for user satisfaction | Analysis of customer satisfaction data from practice | 5.687 customer satisfaction reports from user-CA interactions |
| Gkinko and Elbanna (2023) | CAs for workplace applications | Derivation of key dimensions for the adoption of an AI CA and specification of user types | Taxonomy development: Case study in an organization | 46 interviewees |
| Stieglitz et al. (2022) | CAs for emergency management | Development of meta-requirements and design principles for CAs for emergency management | Semi-structured interviews | 16 interviewees |
| Tamayo-Moreno and Pérez-Marín (2016) | CAs for education | Adaptation of a pedagogic CA for early childhood education | Case study with field observation | 23 participants and 2 educators |
| Okonkwo and Ade-Ibijola (2021) | CAs for education | Overview of prior research on educational CAs: Potentials, challenges and research gaps | Structured literature review | 53 articles from 6 digital databases |

| | | | | |
|---|---|---|---|---|
| Pérez et al. (2020) | CAs for education | Overview of prior research on educational CAs: Types, technology usage, quality, learning experience | Structured literature review | 80 articles from 7 digital databases |
| Nguyen et al. (2019) | CAs for education | Design of an intelligent tutoring CA for solving mathematical function problems | Exemplary application demonstrations | - |
| Anh and Ngan (2021) | CAs for education | Design of a CA for teaching and learning Mathematics | Tool assessment in a field study | 75 participants from 4 high schools |
| Alaaeldin et al. (2021) | CAs for data-related work | Development of a CA to support decision-makers in the understanding of data analytics results and KPIs | DSR: qualitative interviews and test session | 5 participants consisting of 4 decision makers and 1 field expert |
| Narechania et al. (2021) | CAs for data-related work | Development of a toolkit, translating natural language requests into specifications for data visualization | Exemplary application demonstrations | - |
| Simud et al. (2020) | CAs for data-related work | Development of a CA supporting database queries in natural language | Test-set method for performance evaluation in a practical use case | 8.970 generated queries splitted in 80% training, 20% test data |
| Hoon et al. (2020) | CAs for data-related work | Development of a CA supporting database queries in natural language | Performance evaluation via a pilot case study | Data on reservations, customers in a hotel |
| Zhang et al. (2018) | CAs for data-related work | Development of a CA to discover suitable datasets or tools in data-intensive communities | Performance evaluation via perplexity metric | Collection of papers, tools and datasets for 2 domains |

| | | | | |
|---|---|---|---|---|
| Keyner et al. (2019) | CAs for open data | Development of a toolkit to automatically generate CAs for querying individual open data sources | - | - |
| Neumaier et al. (2017) | CAs for open data | Development of a CA to query open data sources | Usability study with test users | 7 participants |
| Holowka et al. (2021) | CAs for citizen science | Proposition of a CA for participatory research with chronically-ill patients | Assessment of currently available tools and participant survey | 137 applications, 98 participants |
| Isacco et al. (2018) | CAs for citizen science | Development of an online platform including a CA for gathering flood information | Tool demonstration in pilot applications | - |
| Lia et al. (2023) | CAs for citizen science | Development of a CA as recreation monitor for data collection with citizens | Tool assessment in a field study | 14 study sites, 1,618 visitors sent 12,550 messages |
| Tallyn et al. (2018) | CAs for citizen science | Development of a CA for participatory collection of ethnographic data | Tool assessment in a field study | 13 participant |
| Portela (2021) | CAs for citizen science | Evaluation of CA potentials for communicating with citizen scientists | Tool evaluation via an experimental study (Wizard-of-Oz method) | 13 participants |
| Tavanapour et al. (2019) | CAs for citizen science | Development of a CA for collecting citizens' ideas: Derivation of 6 design principles | DSR: semi-structured interviews, experimental study | 7 interview participants, 32 experiment participants |
| Athreya et al. (2018) | CAs for citizen science | Proposition of a CA for answering recurrent question within digital communities | - | 1400 users within six months (no evaluation) |

Table A.4.: Overview of key contributions in the literature on CAs.

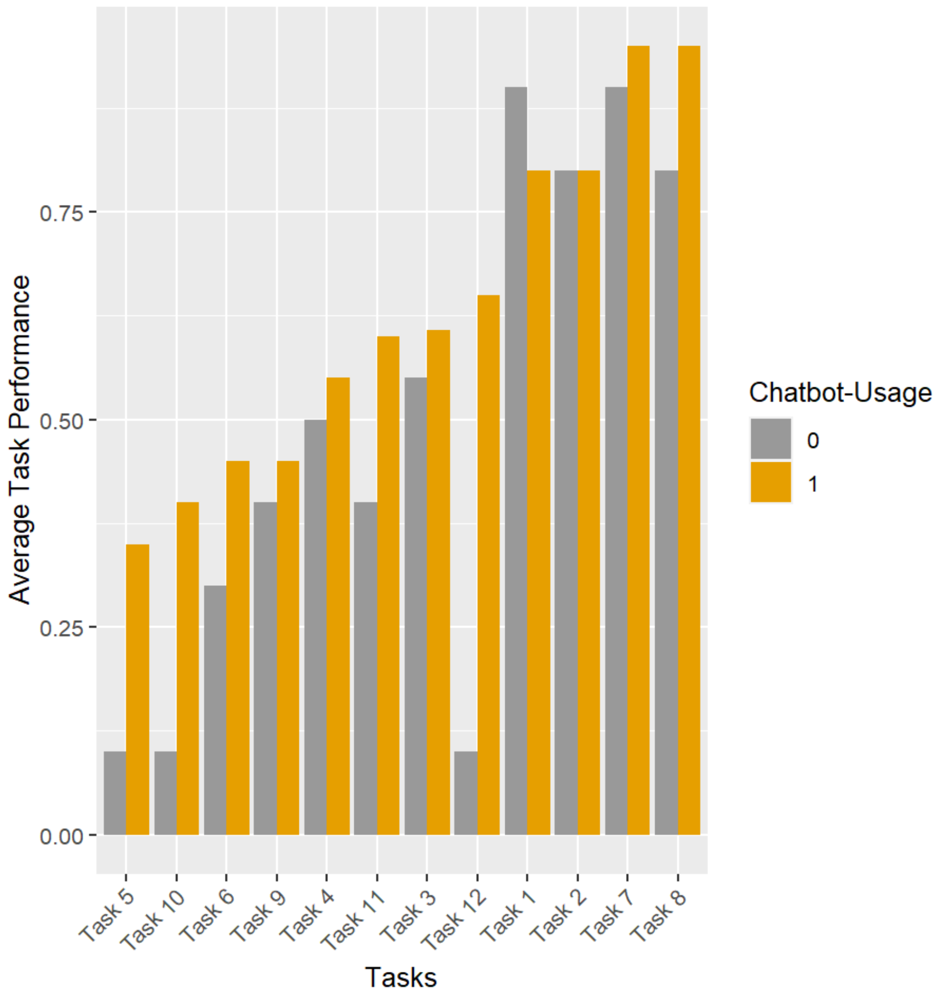


Figure A.1.: Average task performance per treatment and task.

# B. Bibliography


Aichholzer, G. and D. Allhutter, 2009: Public policies on eparticipation in austria. *International Conference on Electronic Participation*, Linz, Austria, 24–35.

Alaaeldin, R., E. Asfoura, G. Kassem, and M. S. Abdel-Haq, 2021: Developing chatbot system to support decision making based on big data analytics. *Journal of Management Information and Decision Sciences*, **24 (2)**, 1–15.

Alaimo, C. and J. Kallinikos, 2024: *Data Rules: Reinventing the Market Economy*. The MIT Press.

Alavi, M., G. M. Marakas, and Y. Yoo, 2002: A comparative study of distributed learning environments on learning outcomes. *Information Systems Research*, **13 (4)**, 404–415, URL `https://www.jstor.org/stable/23015721`.

Allianz Vielfältige Demokratie/ Bertelsmann Stiftung, 2018: Wegweiser breite Bürgerbeteiligung. URL `https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/wegweiser-breite-buergerbeteiligung`, [Accessed: Feb. 2025].

Altenried, M., 2020: The platform as factory: Crowdwork and the hidden labour behind artificial intelligence. *Capital & Class*, **44 (2)**, 145–158, URL `https://doi.org/10.1177/0309816819899410`.

Amado, M., C. Vitorino, E. Moura, and V. Silva, 2009: Public participation in sustainable urban planning. *International Journal of Human and Social Sciences*, **5**, 102–108.

Andersen, K. N., H. Z. Henriksen, C. Secher, and R. Medaglia, 2007: Costs of e-participation: The management challenges. *Transforming Government: People, Process and Policy*, **1 (1)**, 29–43, URL `http://dx.doi.org/10.1108/17506160710733689`.

Anh, N. N. and H. T. Ngan, 2021: Artificial intelligence in mathematics education: an empirical study of using chatbot in teaching and learning mathematics at vietnamese high schools. *5th ASIA PACIFIC International Modern Sciences Congress*, Sydney, Australia, 306–317.

Anya, O., 2015: Bridge the gap! What can work design in crowdwork learn from work design theories? *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Association for Computing Machinery, New York, NY, USA, 612–627, URL `https://doi.org/10.1145/2675133.2675227`.

Aristeidou, M. and C. Herodotou, 2020: Online citizen science: A systematic review of effects on learning and scientific literacy. *Citizen Science: Theory and Practice*, **5 (1)**, 1–12, URL http://oro.open.ac.uk/70067/.

Arnstein, S. R., 1969: A ladder of citizen participation. *Journal of the American Institute of Planners*, **35 (4)**, 216–224, URL http://www.tandfonline.com/doi/abs/10.1080/01944366908977225.

Athreya, R. G., A.-C. Ngonga Ngomo, and R. Usbeck, 2018: Enhancing community interactions with data-driven chatbots - the DBpedia chatbot. *The Web Conference 2018 - Companion of the World Wide Web Conference, WWW 2018*, Lyon, France, 143–146, URL https://dl.acm.org/doi/10.1145/3184558.3186964.

Avdiji, H., D. A. Elikan, S. Missonier, and Y. Pigneur, 2020: A design theory for visual inquiry tools. *Journal of the Association for Information Systems*, **21**, 695–734.

Baba, Y., T. Takase, K. Atarashi, S. Oyama, and H. Kashima, 2018: Data analysis competition platform for educational purposes: Lessons learned and future challenges. *Proceedings of the AAAI Conference on Artificial Intelligence*, **32 (1)**, 7787–7892, URL https://doi.org/10.1609/aaai.v32i1.11391.

Barnes, S.-A., A. Green, and M. De Hoyos, 2015: Crowdsourcing and work: Individual factors and circumstances influencing employability. *New Technology, Work and Employment*, **30 (1)**, 16–31, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ntwe.12043.

Barzilai, S. and I. Blau, 2014: Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, **70**, 65–79, URL https://www.sciencedirect.com/science/article/pii/S0360131513002224.

Baudry, J., Tancoigne, and B. J. Strasser, 2022: Turning crowds into communities: The collectives of online citizen science. *Social Studies of Science*, **52 (3)**, 399–424, URL https://doi.org/10.1177/03063127211058791.

Bautista-Puig, N., D. De Filippo, E. Mauleón, and E. Sanz-Casado, 2019: Scientific landscape of citizen science publications: Dynamics, content and presence in social media. *Publications*, **7 (1)**, 12.

Bayus, B. L., 2012: Crowdsourcing new product ideas over time: An analysis of dell's ideastorm community. *Management Science*, **59 (1)**, 226–244, URL https://papers.ssrn.com/abstract=1979557.

Becker, A., L. Ecker, I. Külpmann, K. Schwien, and P. Stobbe, 2023: Cooperative solidarity among crowdworkers? Social learning practices on a crowdtesting social media platform. *Convergence*, **30 (1)**, 428–449, URL https://doi.org/10.1177/13548565231183298.

Bedi, C., A. Kansal, and P. Mukheibir, 2023: A conceptual framework for the assessment of and the transition to liveable, sustainable and equitable cities. *Environmental Science & Policy*, **140**, 134–145, URL `https://www.sciencedirect.com/science/article/pii/S1462901122003665`.

Behrens, P., M. Calmbach, C. Schleer, W. Klingler, and T. Rathgeb, 2014: Mediennutzung und Medienkompetenz in jungen Lebenswelten. Repräsentative Onlinebefragung von 14- bis 29-Jährigen in Deutschland. *Media Perspektiven*, **4**, 195–218.

Bela, G., T. Peltola, J. C. Young, B. Balázs, I. Arpin, G. Pataki, J. Hauck, E. Kelemen, L. Kopperoinen, A. Van Herzele, H. Keune, S. Hecker, M. Suškevičs, H. E. Roy, P. Itkonen, M. Külvik, M. László, C. Basnou, J. Pino, and A. Bonn, 2016: Learning and the transformative potential of citizen science. *Conservation Biology*, **30 (5)**, 990–999, URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/cobi.12762`.

Bhargava, R., E. Deahl, E. Letouzé, A. Noonan, D. Sangokoya, and N. Shoup, 2015: Beyond data literacy: Reinventing community engagement and empowerment in the age of data. Data-Pop Alliance; White Paper Series, New York City, NY, USA, URL `https://datapopalliance.org/item/beyond-data-literacy`.

Bhargava, R., R. Kadouaki, E. Bhargava, G. Castro, and C. D'Ignazio, 2016: Data murals: Using the arts to build data literacy. *The Journal of Community Informatics*, **12 (3)**, 197–216, URL `https://doi.org/10.15353/joci.v12i3.3285`.

Bigham, J. P., K. Williams, N. Banerjee, and J. Zimmerman, 2017: Scopist: Building a skill ladder into crowd transcription. *Proceedings of the 14th International Web for All Conference*, Association for Computing Machinery, New York, NY, USA, 1–10, URL `https://doi.org/10.1145/3058555.3058562`.

Bigo, D., E. Isin, and E. Ruppert, 2019: *Data Politics: Worlds, Subjects, Rights*. Routledge.

Biljecki, F., H. Ledoux, and J. Stoter, 2016: An improved LOD specification for 3d building models. *Computers Environment and Urban Systems*, **59**, 25–37.

Billger, M., L. Thuvander, and B. S. Wästberg, 2017: In search of visualization challenges: The development and implementation of visualization tools for supporting dialogue in urban planning processes. *Environment and Planning B*, **44 (6)**, 1012–1035, URL `https://doi.org/10.1177/0265813516657341`.

Bittner, E., S. Oeste-Reiß, and J. M. Leimeister, 2019: Where is the bot in our team? Toward a taxonomy of design option combinations for conversational agents in collaborative work. *Proceedings of the 52nd Hawaii international conference on system sciences*, Grand Wailea, Hawaii, USA.

Bonney, R., T. B. Phillips, H. L. Ballard, and J. W. Enck, 2016: Can citizen science enhance public understanding of science? *Public Understanding of Science*, **25 (1)**, 2–16, URL `https://doi.org/10.1177/0963662515607406`.

Börsch-Supan, J., 2017: Coding & Charakter - Welche Kompetenzen betrachten die Deutschen als die wichtigsten für die digitale Zukunft? Eine repräsentative Befragung im Auftrag der Vodafone Stiftung. Vodafone Stiftung Deutschland GmbH, URL `https://www.vodafone-stiftung.de/coding-und-charakter/`.

Bowser, A., C. Cooper, A. De Sherbinin, A. Wiggins, P. Brenton, T.-R. Chuang, E. Faustman, M. Haklay, and M. Meloche, 2020: Still in need of norms: The state of the data in citizen science. *Citizen Science: Theory and Practice*, **5 (1)**, Artikel 18, URL `http://theoryandpractice.citizenscienceassociation.org/articles/10.5334/cstp.303/`.

Boyd, D. M. and N. B. Ellison, 2007: Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, **13 (1)**, 210–230, URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1083-6101.2007.00393.x`.

Brabham, D. C., 2008: Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence*, **14 (1)**, 75–90, URL `https://doi.org/10.1177/1354856507084420`.

Brenton, P., S. Von Gavel, E. Vogel, and M.-E. Lecoq, 2018: Technology infrastructure for citizen science. Hecker, S., M. Haklay, A. Bowser, Z. Makuch, J. Vogel, and A. Bonn, (Eds.), *Citizen Science: Innovation in Open Science, Society and Policy*, UCL Press, London, 63–80, URL `https://www.jstor.org/stable/j.ctv550cf2.12`.

Carmi, E., S. J. Yates, E. Lockley, and A. Pawluczuk, 2020: Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review*, **9 (2)**, 22, URL `http://dx.doi.org/10.14763/2020.2.1481`.

Center for Self-Determination Theory, 2022: Intrinsic motivation inventory (IMI). URL `https://selfdeterminationtheory.org/intrinsic-motivation-inventory/`, [Accessed: Feb. 2025].

Chan, S. C. H. and S. Ko, 2021: The dark side of personal response systems (PRSs): Boredom, feedback, perceived learning, learning satisfaction. *Journal of Education for Business*, **96 (4)**, 435–444, URL `http://dx.doi.org/10.1080/08832323.2020.1848769`.

Charness, G., U. Gneezy, and M. A. Kuhn, 2012: Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, **81 (1)**, 1–8, URL `https://www.sciencedirect.com/science/article/abs/pii/S0167268111002289`.

Chesbrough, H., 2006: chap. Open Innovation: A New Paradigm for Understanding Industrial Innovation. Chesbrough, H., W. Vanhaverbeke, and J. West, (Eds.), *Open Innovation: Researching a New Paradigm*, 1–12, Oxford University Press.

Chiang, C.-W., A. Kasunic, and S. Savage, 2018: Crowd coach: Peer coaching for crowd workers' skill growth. *Proceedings of the ACM on Human-Computer Interaction*, **2**, 1–17, URL https://doi.org/10.1145/3274306.

Choi, S.-S., S.-H. Cha, and C. C. Tappert, 2010: A survey of binary similarity and distance measures. *Journal of systemics, cybernetics and informatics*, **8 (1)**, 43–48.

Chow, W., 2019: A pedagogy that uses a kaggle competition for teaching machine learning: An experience sharing. *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, Yogyakarta, Indonesia, 1–5, URL https://ieeexplore.ieee.org/document/9226005.

Clarke, R., 2016: Big data, big risks. *Information Systems Journal*, **26 (1)**, 77–90, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/isj.12088.

Coetzee, D., S. Lim, A. Fox, B. Hartmann, and M. A. Hearst, 2015: Structuring interactions for large-scale synchronous peer learning. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Association for Computing Machinery, New York, NY, USA, 1139–1152, URL https://dl.acm.org/doi/10.1145/2675133.2675251.

Cohen, J., 1992: A power primer. *Psychological Bulletin*, **112 (1)**, 155–159.

Colombo, E., F. Mercorio, and M. Mezzanzanica, 2019: AI meets labor market: Exploring the link between automation and skills. *Information Economics and Policy*, **47**, 27–37, URL https://www.sciencedirect.com/science/article/pii/S0167624518301318.

Connor, H., 2024: John Graunt F.R.S. (1620-74): The founding father of human demography, epidemiology and vital statistics. *Journal of Medical Biography*, **32 (1)**, 57–69.

Culbertson, G., S. Shen, E. Andersen, and M. Jung, 2017: Have your cake and eat it too: Foreign language learning with a crowdsourced video captioning system. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, Association for Computing Machinery, New York, NY, USA, 286–296, URL https://doi.org/10.1145/2998181.2998268.

Davies, J. and R. Procter, 2020: Online platforms of public participation – a deliberative democracy or a delusion? Charalabidis, Y. and M. A. Cunha, (Eds.), *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2020)*, New York, NY, USA, 746–753.

Davis, F. D., 1989: Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, **13 (3)**, 319–340, URL https://www.jstor.org/stable/249008.

De Albuquerque, J. and A. Almeida, 2020: Modes of engagement: Reframing "sensing" and data generation in citizen science for empowering relationships. Davies, T. and A. Mah, (Eds.), *Toxic truths: Environmental justice and citizen science in a post-truth age*, Manchester University Press, 267–281.

De Leon Pereira, R., A. Tan, A. Bunt, and O. Tremblay-Savard, 2021: Increasing player engagement, retention and performance through the inclusion of educational content in a citizen science game. *The 16th International Conference on the Foundations of Digital Games (FDG 2021)*, ACM, Montreal QC Canada, 1–12, URL https://dl.acm.org/doi/10.1145/3472538.3472554.

Debruyne, C., A. Kearns, C. O'Neill, M. Colclough, L. Grehan, and D. O'Sullivan, 2021: DALIDA: Data literacy discussion workshops for adults. *Companion Publication of the 13th ACM Web Science Conference 2021*, Association for Computing Machinery, New York, NY, USA, 23–25.

Deng, J., W. Chai, J. Guo, Q. Huang, W. Hu, J.-N. Hwang, and G. Wang, 2023: CityGen: Infinite and controllable 3d city layout generation. URL http://arxiv.org/abs/2312.01508.

Deschênes, M., 2020: Recommender systems to support learners' agency in a learning context: A systematic review. *International Journal of Educational Technology in Higher Education*, **17 (1)**, 50, URL https://doi.org/10.1186/s41239-020-00219-w.

Deutscher Städtetag, wegewerk, 2017: 3D-Geodaten in der integrierten Stadtentwicklung: Deutscher Städtetag. wegewerk GmbH, URL https://www.staedtetag.de/publikationen/weitere-publikationen/3d-geodaten-integrierte-stadtentwicklung-2017.

Diederich, S., A. Brendel, S. Morana, and L. Kolbe, 2022: On the design of and interaction with conversational agents: An organizing and assessing review of human-computer interaction research. *Journal of the Association for Information Systems*, **23 (1)**, 96–138, URL https://aisel.aisnet.org/jais/vol23/iss1/9.

D'Ignazio, C., 2017: Creative data literacy: Bridging the gap between the data-haves and data-have nots. *Information Design Journal*, **23 (1)**, 6–18, URL http://www.jbe-platform.com/content/journals/10.1075/idj.23.1.03dig.

D'Ignazio, C. and R. Bhargava, 2016: DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics*, **12 (3)**, 83–107, URL https://openjournals.uwaterloo.ca/index.php/JoCI/article/view/3280.

Dontcheva, M., R. R. Morris, J. R. Brandt, and E. M. Gerber, 2014: Combining crowdsourcing and learning to improve engagement and performance. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 3379–3388, URL https://doi.org/10.1145/2556288.2557217.

Doroudi, S., E. Kamar, E. Brunskill, and E. Horvitz, 2016: Toward a learning science for complex crowdsourcing tasks. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 2623–2634, URL `https://doi.org/10.1145/2858036.2858268`.

Dow, S., A. Kulkarni, S. Klemmer, and B. Hartmann, 2012: Shepherding the crowd yields better work. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, Association for Computing Machinery, New York, NY, USA, 1013–1022, URL `https://doi.org/10.1145/2145204.2145355`.

Dresing, T. and T. Pehl, 2018: *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende*. 8th ed., Eigenverlag, Marburg.

Dwivedi, Y., L. Hughes, A. Baabdullah, S. Ribeiro-Navarrete, M. Giannakis, M. Al-Debei, D. Dennehy, B. Metri, D. Buhalis, and R. Felix, 2022: Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, **66**, URL `https://doi.org/10.1016/j.ijinfomgt.2022.102542`.

Eilola, S., K. Jaalama, P. Kangassalo, P. Nummi, A. Staffans, and N. Fagerholm, 2023: 3d visualisations for communicative urban and landscape planning: What systematic mapping of academic literature can tell us of their potential? *Landscape and Urban Planning*, **234**, 104 716, URL `https://www.sciencedirect.com/science/article/pii/S016920462300035X`.

Ekinci, E., S. I. Omurca, and N. Acun, 2018: A comparative study on machine learning techniques using titanic dataset. Antalya Turkey, 411–416.

Engel, J. and J. Döllner, 2012: Immersive visualization of virtual 3d city models and its applications in e-planning. *International Journal of E-Planning Research (IJEPR)*, **1 (4)**, 17–34.

Eom, S. B., H. J. Wen, and N. Ashill, 2006: The determinants of students' perceived learning outcomes and satisfaction in university online education: An empirical investigation. *Decision Sciences Journal of Innovative Education*, **4 (2)**, 215–235, URL `https://onlinelibrary.wiley.com/doi/10.1111/j.1540-4609.2006.00114.x`.

EOS/ Kantar, 2020: Was sind Daten wert? - EOS WhitePapers. EOS Holding GmbH.

Ericsson, K. A. and H. A. Simon, 1984: *Protocol analysis: Verbal reports as data*. The MIT Press.

Estellés-Arolas, E., R. Navarro-Giner, and F. G. Guevara, 2015: Crowdsourcing fundamentals: Definition and typology. *Advances in Crowdsourcing*, 33–48, URL `http://dx.doi.org/10.1007/978-3-319-18341-1_3`.

European Citizen Science Association, 2015: Ten principles of citizen science. URL https://ecsa.citizen-science.net/documents/, [Accessed: Feb. 2025].

European Comission, 2022: The digital economy and society index (desi). URL https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2022, [Accessed: Feb. 2025].

Fan, M., Y. Li, and K. N. Truong, 2020: Automatic detection of usability problem encounters in think-aloud sessions. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, **10 (2)**, 1–24, URL https://dl.acm.org/doi/10.1145/3385732.

Fantoni, G., R. Apreda, D. Gabelloni, and G. Montelisciani, 2012: You solve, i learn: A novel approach to e-learning in collaborative crowdsourcing. *2012 18th International ICE Conference on Engineering, Technology and Innovation*, IEEE, Munich, Germany, 1–10, URL https://ieeexplore.ieee.org/document/6297659.

Feenberg, A., 1999: *Questioning Technology*. Routledge.

Fegert, J., 2022: *Digital Citizen Participation – Involving Citizens Through Immersive Systems in Urban Planning*. Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.

Fegert, J., J. Pfeiffer, C. Peukert, A. Golubyeva, and C. Weinhardt, 2020: Combining e-participation with augmented and virtual reality: Insights from a design science research project. *Proceedings of the 41st International Conference on Information Systems, (ICIS 2022)*, AIS eLibrary (AISeL), Hyderabad, India, URL https://aisel.aisnet.org/icis2020/digitization_in_cities/digitization_in_cities/4.

Feldmann, N. and H. Gimpel, 2016: Financing projects through enterprise crowdfunding: Understanding the impact of proposal characteristics on funding success. *Proceedings of the 24th European Conference on Information Systems (ECIS 2016)*, Istanbul, Turkey.

Ferrari, A., 2013: *DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe*, Punie, Y. and B. N. Brečko, (Eds.). Publications Office of the European Union, Luxembourg.

Fischer, G., J. Lundin, and J. O. Lindberg, 2020: Rethinking and reinventing learning, education and collaboration in the digital age—from creating technologies to transforming cultures. *The International Journal of Information and Learning Technology*, **37 (5)**, 241–252, URL https://doi.org/10.1108/IJILT-04-2020-0051.

Flores, C. C. and D. A. Rezende, 2022: Crowdsourcing framework applied to strategic digital city projects. *Journal of Urban Management*, **11 (4)**, 467–478, URL https://www.sciencedirect.com/science/article/pii/S2226585622000723.

Fouda, E., 2020: chap. What Is Next? Fouda, E., (Ed.), *Learn Data Science Using SAS Studio: A Quick-Start Guide*, 211–219, Apress, Berkeley, CA, USA, URL https://doi.org/10.1007/978-1-4842-6237-5_9.

Fritz, S., L. See, T. Carlson, M. Haklay, J. L. Oliver, D. Fraisl, R. Mondardini, M. Brocklehurst, L. A. Shanley, S. Schade, U. Wehn, T. Abrate, J. Anstee, S. Arnold, M. Billot, J. Campbell, J. Espey, M. Gold, G. Hager, S. He, L. Hepburn, A. Hsu, D. Long, J. Masó, I. McCallum, M. Muniafu, I. Moorthy, M. Obersteiner, A. J. Parker, M. Weisspflug, and S. West, 2019: Citizen science and the united nations sustainable development goals. *Nature Sustainability*, **2 (10)**, 922–930, URL https://www.nature.com/articles/s41893-019-0390-3.

García, D. L., A. Fernandez, L. Giesen, M. Landi, D. Mackisack, M. Neven, and P. Wolf, 2023: Democracy technologies in europe—online participation, deliberation and voting. The Innovation in Politics Institute GmbH., URL https://democracy-technologies.org/report-2023/.

García, F. S., M. Pelacho, T. Woods, D. Fraisl, L. See, M. Haklay, and R. Arias, 2021: chap. Finding What You Need: A Guide to Citizen Science Guidelines. Vohland, K., A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, (Eds.), *The Science of Citizen Science*, 419–437, Springer International Publishing, Cham.

Gediga, G., K.-C. Hamborg, and I. Düntsch, 2002: chap. Evaluation of Software Systems. Kent, A. and J. G. Williams, (Eds.), *Encyclopedia of Library and Information Science*, Vol. 72, 127–153, Marc Dekker Inc., New York, NY, USA.

Gilardi, F., M. Alizadeh, and M. Kubli, 2023: ChatGPT outperforms crowd-workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, **120 (30)**, URL https://www.pnas.org/doi/epdf/10.1073/pnas.2305016120, 2303.15056[cs].

Girindran, R., D. S. Boyd, J. Rosser, D. Vijayan, G. Long, and D. Robinson, 2020: On the reliable generation of 3d city models from open data. *Urban Science*, **4 (4)**, 47, URL https://www.mdpi.com/2413-8851/4/4/47.

Gkinko, L. and A. Elbanna, 2023: The appropriation of conversational AI in the workplace: A taxonomy of AI chatbot users. *International Journal of Information Management*, **69**, 102 568, URL https://www.sciencedirect.com/science/article/pii/S0268401222001025.

Glynn, S. M., P. Brickman, N. Armstrong, and G. Taasoobshirazi, 2011: Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, **48 (10)**, 1159–1176, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/tea.20442.

Gnewuch, U. and A. Mädche, 2022: Toward a method for reviewing software artifacts from practice. *The Transdisciplinary Reach of Design Science Research*, Springer, Saint Petersburg, USA, 337–350, URL https://link.springer.com/chapter/10.1007/978-3-031-06516-3_25.

Golumbic, Y. N., B. Fishbain, and A. Baram-Tsabari, 2020: Science literacy in action: understanding scientific data presented in a citizen science platform by non-expert adults. *International Journal of Science Education, Part B*, **10 (3)**, 232–247, URL https://doi.org/10.1080/21548455.2020.1769877.

Gould, R., 2021: Toward data-scientific thinking. *Teaching Statistics*, **43**, 11–22, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/test.12267.

Gregor, S. and A. R. Hevner, 2013: Positioning and presenting design science research for maximum impact. *MIS Quarterly*, **37 (2)**, 337–356, URL https://doi.org/10.25300/MISQ/2013/37.2.01.

Gregor, S., C. L. Kruse, and S. Seidel, 2020: Research perspectives: The anatomy of a design principle. *Journal of the Association for Information Systems*, **21 (6)**, 1622–1652.

Grönlund, , 2009: ICT is not participation is not democracy–eparticipation development models revisited. A, M. and T. E, (Eds.), *International Conference on Electronic Participation*, Linz, Australia, 12–23.

Gupta, K., P. Sharma, and C. N. Bouza, 2018: Surviving the titanic tragedy: A sociological study using machine learning models. *Revista Suma de Negocios*, **9 (20)**, 86–92, URL https://papers.ssrn.com/abstract=3417433.

Haklay, M., 2013a: Citizen science and volunteered geographic information: Overview and typology of participation. Sui, D., S. Elwood, and M. Goodchild, (Eds.), *Crowdsourcing geographic knowledge*, Springer, Dordrecht, 105–122, URL https://link.springer.com/chapter/10.1007/978-94-007-4587-2_7.

Haklay, M., 2013b: Neogeography and the delusion of democratisation. *Environment and Planning A*, **45 (1)**, 55–69, URL http://dx.doi.org/10.1068/a45184.

Haklay, M., D. Dörler, F. Heigl, M. Manzoni, S. Hecker, and K. Vohland, 2021: chap. What Is Citizen Science? The Challenges of Definition. Vohland, K., A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, (Eds.), *The Science of Citizen Science*, 13–33, Springer International Publishing, Cham, URL https://doi.org/10.1007/978-3-030-58278-4_2.

Halmos, A., G. Misuraca, and G. Viscusi, 2019: From public value to social value of digital government: Co-creation and social innovation in european union initiatives. *Proceedings of the 52nd Hawaii International Conference on System Sciences 2019*, Grand Wailea, Hawaii, USA.

Hammon, L. and H. Hippner, 2012: Crowdsourcing. *Business Information Systems Engineering*, **4 (3)**, 1–4.

Hanley, H. W. A. and Z. Durumeric, 2023: Machine-made media: Monitoring the mobilization of machine-generated articles on misinformation and mainstream news websites. arXiv, URL `https://arxiv.org/abs/2305.09820v1`.

Hayek, U. W., 2011: Which is the appropriate 3d visualization type for participatory landscape planning workshops? A portfolio of their effectiveness. *Environment and Planning B: Urban Analytics and City Science*, **38 (5)**, 921–939, URL `https://doi.org/10.1068/b36113`.

Head, B. and J. Alford, 2013: Wicked problems: Implications for public policy and management. *Administration & Society*, **47 (6)**, 711–739, URL `https://journals.sagepub.com/doi/10.1177/0095399713481601`.

Hecker, S. and M. Taddicken, 2022: Deconstructing citizen science: a framework on communication and interaction using the concept of roles. *Journal of Science Communication*, **21 (1)**, A07.

Hedderich, M. A. and A. Oulasvirta, 2024: Explaining crowdworker behaviour through computational rationality. *Behaviour & Information Technology*, 1–22, URL `https://www.tandfonline.com/doi/abs/10.1080/0144929X.2024.2329616`.

Heigl, F., B. Kieslinger, K. T. Paul, J. Uhlik, D. Frigerio, and D. Dörler, 2020: Co-creating and implementing quality criteria for citizen science. *Citizen Science: Theory and Practice*, **5 (1)**, 23.

Herodotou, C., E. Scanlon, and M. Sharples, 2021: Methods of promoting learning and data quality in citizen and community science. *Frontiers in Climate*, **3**, URL `https://www.frontiersin.org/articles/10.3389/fclim.2021.614567`.

Hevner, A. R., 2007: A three cycle view of design science research. *Scandinavian Journal of Information Systems*, **19 (2)**, 87–92.

Hin, D., 2020: StackOverflow vs kaggle: A study of developer discussions about data science. arXiv, arXiv:2006.08334, URL `http://arxiv.org/abs/2006.08334`.

Hinderks, A., M. Schrepp, and J. Thomaschewski, 2018: A benchmark for the short version of the user experience questionnaire. *Proceedings of the 14th International Conference on Web Information Systems and Technologies (WEBIST 2018)*, SCITEPRESS- Science and Technology Publications, Seville, Spain, 373–377.

Hintz, A., L. Dencik, and K. Wahl-Jorgensen, 2018: *Digital Citizenship in a Datafied Society*. Polity, Cambridge, UK.

Hofmann, S., C. Madsen, and B. Distel, 2020: Developing an analytical framework for analyzing and comparing national e-government strategies. Viale Pereira, G., M. Janssen, H. Lee, I. Lindgren, M. P. Rodríguez Bolívar, H. J. Scholl, and A. Zuiderwijk, (Eds.), *Electronic Government*, Springer International Publishing, Cham, 15–28.

Holowka, E., S. Woods, A. Pahayahay, M. Roy, and N. Khalili-Mahani, 2021: Principles for designing an mHealth app for participatory research and management of chronic pain. Duffy, V. G., (Ed.), *Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management AI, Product and Service. Lecture Notes in Computer Science*, Springer, Cham, 50–67.

Hoon, G. K., L. J. Yong, and G. K. Yang, 2020: Interfacing chatbot with data retrieval and analytics queries for decision making. Abdul Majeed, A., J. A. Mat-Jizat, M. H. A. Hassan, Z. Taha, H. L. Choi, and J. Kim, (Eds.), *Proceedings of the 6th International Conference on Robot Intelligence Technology and Applications (RITA 2018)*, Springer Singapore, 385–394.

Horton, W., 2001: *Evaluating E-learning*. American Society for Training and Development.

Howe, J., 2006a: Crowdsourcing: A definition. URL `http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html`, [Accessed: Feb. 2025].

Howe, J., 2006b: The rise of crowdsourcing. *Wired*, **4 (2)**, URL `https://www.wired.com/2006/06/crowds/`.

Howe, J., 2008: *Crowdsourcing: How the Power of the Crowd is Driving the Future of Business*. Random House Business.

Huang, K., J. Zhou, and S. Chen, 2022: Being a solo endeavor or team worker in crowdsourcing contests? It is a long-term decision you need to make. *Proceedings of the ACM on Human-Computer Interaction*, **6**, 1–32, URL `https://dl.acm.org/doi/10.1145/3555595`.

Imottesjo, H. and J.-H. Kain, 2022: The urban CoCreation lab – an integrated platform for remote and simultaneous co llaborative urban planning and design through web-based desktop 3d modeling, head-mounted virtual reality and mobile augmented reality: Prototyping a minimum viable product and developing specifications for a minimum marketable product. *Applied Sciences*, **12 (2)**, URL `https://www.mdpi.com/2076-3417/12/2/797`.

International Association of Public Participation, 2007: IAP2 spectrum of public participation. URL `https://iap2.org.au/resources/spectrum/`, [Accessed: Feb. 2025].

Irwin, A. and Horlick-Jones, Tom, 1995: *Citizen Science: A study of People, Expertise and Sustainable Development*. 1st ed., Routledge, London, UK.

Isacco, S., P. Claps, S. Grasso, E. Ferrari, M. B. Guercio, R. E. Musumeci, G. E. Scarcella, P. Versace, and F. Laio, 2018: Floodbook: a social platform for flood hydrology. *Proceedings of the 13th International Conference on Hydroinformatics (HIC 2018)*, Palermo, Italy, Vol. 3, 941–949.

Jackson, C. B., C. Østerlund, K. Crowston, M. Harandi, and L. Trouille, 2020: Shifting forms of engagement: Volunteer learning in online citizen science. *Proceedings of the ACM on Human-Computer Interaction*, **4**, 1–19, URL `https://dl.acm.org/doi/10.1145/3392841`.

Janssen, A., J. Passlick, D. Rodríguez Cardona, and M. H. Breitner, 2020: Virtual assistance in any context: A taxonomy of design elements for domain-specific chatbots. *Business and Information Systems Engineering*, **62 (3)**, 211–225, URL `https://link.springer.com/article/10.1007/s12599-020-00644-1#citeas`.

Jayawickrama, T., S. Abdelaal, and R. Abadia, 2020: Using online learning environments to address digital literacy competencies of construction management graduates. *Proceedings of the 28th International Conference on Computers in Education, (ICCE 2020)*, Vol. 1, 670–679.

Jennett, C. and A. Cox, 2014: Eight guidelines for designing virtual citizen science projects. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, PKP Publishing Services Network, Pittsburgh, Pennsylvania, Vol. 2.

Jennett, C., L. Kloetzer, D. Schneider, I. Iacovides, A. Cox, M. Gold, B. Fuchs, A. Eveleigh, K. Methieu, Z. Ajani, and others, 2016: Motivations, learning and creativity in online citizen science. *Journal of Science Communication*, **15 (3)**.

Jensen, E. A., E. B. Kennedy, and E. Greenwood, 2021: Pandemic: public feeling more positive about science. *Nature*, **591 (7848)**, 34–34, URL `https://www.nature.com/articles/d41586-021-00542-w`.

Jones, M. G., G. Childers, T. Andre, E. N. Corin, and R. Hite, 2018: Citizen scientists and non-citizen scientist hobbyists: Motivation, benefits, and influences. *International Journal of Science Education. Part B: Communication and Public Engagement*, **8 (4)**, 287–306, URL `https://www.tandfonline.com/doi/full/10.1080/21548455.2018.1475780`.

Kaiser, R., 2014: *Qualitative Experteninterviews*. Springer Fachmedien Wiesbaden, Wiesbaden.

Kassambara, A., 2017: *Practical guide to cluster analysis in R: Unsupervised machine learning*, Vol. 1. Sthda.

Kelly, G., 2006: A survey of procedural techniques for city generation. *The ITB Journal*, **7 (2)**, Article 5, URL `https://doi.org/10.21427/D76M9P`.

Kelly, G., 2007: Citygen: An interactive system for procedural city generation. *Proceedings of the Fifth Annual International Conference in Computer Game Design and Technology (GDTW 2007)*, Liverpool, UK.

Kermish-Allen, R., K. Peterman, and C. Bevc, 2019: The utility of citizen science projects in k-5 schools: measures of community engagement and student impacts. *Cultural Studies of Science Education*, **14 (3)**, 627–641, URL `https://doi.org/10.1007/s11422-017-9830-4`.

Keyner, S., V. Savenkov, and S. Vakulenko, 2019: Open data chatbot. Hitzler, P., S. Kirrane, O. Hartig, V. de Boer, M.-E. Vidal, M. Maleshkova, S. Schlobach, K. Hammar, N. Lasierra, S. Stadtmüller, K. Hose, and R. Verborgh, (Eds.), *The Semantic Web: ESWC 2019 Satellite Events*, Springer, Cham, 111–115.

Kim, H.-W. and S. Gupta, 2014: A user empowerment approach to information systems infusion. *IEEE Transactions on Engineering Management*, **61 (4)**, 656–668, URL `http://dx.doi.org/10.1109/TEM.2014.2354693`.

Kirkpatrick, D., 1959: Techniques for evaluating training programs. *Journal of American Society of Training Directors*, **13**, 21–26.

Kittur, A., J. V. Nickerson, M. Bernstein, E. Gerber, A. Shaw, J. Zimmerman, M. Lease, and J. Horton, 2013: The future of crowd work. *Proceedings of the 2013 conference on Computer supported cooperative work (CSCW 2013)*, Association for Computing Machinery, New York, NY, USA, 1301–1318, URL `https://doi.org/10.1145/2441776.2441923`.

Kjelvik, M. K. and E. H. Schultheis, 2019: Getting messy with authentic data: Exploring the potential of using data from scientific research to support student data literacy. *CBE—Life Sciences Education*, **18 (2)**, 8, URL `https://www.lifescied.org/doi/full/10.1187/cbe.18-02-0023`.

Kloetzer, L., J. Lorke, J. Roche, Y. Golumbic, S. Winter, and A. Jõgeva, 2021: Learning in citizen science. Vohland, K., A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, (Eds.), *The Science of Citizen Science*, Springer International Publishing, Cham, 283–308.

Kloker, S., T. Straub, and C. Weinhardt, 2017: Designing a crowd forecasting tool to combine prediction markets and real-time delphi. *International Conference on Design Science Research in Information System and Technology*, Karlsruhe, Germany.

Kobayashi, M., H. Morita, M. Matsubara, N. Shimizu, and A. Morishima, 2021: Empirical study on effects of self-correction in crowdsourced microtasks. *Human Computation. A Transdisciplinary Journal*, **8 (1)**.

Koch, G., J. Füller, and S. Brunswicker, 2011: Online crowdsourcing in the public sector: How to design open government platforms. Ozok, A. A. and P. Zaphiris, (Eds.), *Online Communities and Social Computing*, Springer, Berlin, Heidelberg, 203–212.

Kranz, T. T., F. Teschner, and C. Weinhardt, 2014: Combining prediction markets and surveys: An experimental study. *Proceedings of the 22nd European Conference on Information Systems (ECIS 2014)*, Tel Aviv, Israel.

Krishnamurthy, R. and Y. Awazu, 2016: Liberating data for public value: The case of data.gov. *International Journal of Information Management*, **36 (4)**, 668–672, URL `https://www.sciencedirect.com/science/article/pii/S0268401216301098`.

Krueger, R. A. and M. A. Casey, 2014: *Focus Groups: A Practical Guide for Applied Research*. SAGE Publications, URL `https://us.sagepub.com/en-us/nam/focus-groups/book243860`.

Kruger, J. and D. Dunning, 1999: Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, **77 (6)**, 1121–34, URL `https://pubmed.ncbi.nlm.nih.gov/10626367/`.

Kuckartz, U. and S. Rädiker, 2022: *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. 5th ed., Beltz Juventa, URL `https://www.qualitativeinhaltsanalyse.de/`.

Kuek, S. C., C. Paradi-Guilford, T. Fayomi, S. Imaizumi, P. Ipeirotis, P. Pina, and M. Singh, 2015: The global opportunity in online outsourcing. The World Bank Group.

Kullenberg, C. and D. Kasperowski, 2016: What is citizen science? – A scientometric meta-analysis. *PLOS ONE*, **11 (1)**, e0147 152.

Kundisch, D., J. Muntermann, A. M. Oberländer, D. Rau, M. Röglinger, T. Schoormann, and D. Szopinski, 2021: An update for taxonomy designers. *Business & Information Systems Engineering*, **64 (4)**, 421–439, URL `https://doi.org/10.1007/s12599-021-00723-x`.

Kvale, K., E. Freddi, S. Hodnebrog, O. A. Sell, and A. Følstad, 2021: Understanding the user experience of customer service chatbots: What can we learn from customer satisfaction surveys? Følstad, A., T. Araujo, S. Papadopoulos, E. L.-C. Law, E. Luger, M. Goodwin, and P. B. Brandtzaeg, (Eds.), *Chatbot Research and Design*, Springer, Cham, Virtual Event, 205–218, URL `https://link.springer.com/chapter/10.1007/978-3-030-68288-0_14`.

Lafrance, F., S. Daniel, and S. Dragićević, 2019: Multidimensional web GIS approach for citizen participation on urban evolution. *ISPRS International Journal of Geo-Information*, **8 (6)**, 253, URL `https://www.mdpi.com/2220-9964/8/6/253`.

Landis, J. R. and G. G. Koch, 1977: The measurement of observer agreement for categorical data. *Biometrics*, **33 (1)**, 159–174, URL https://www.jstor.org/stable/2529310.

Le Blanc, D. and United Nations, 2020: E-participation: A quick overview of recent qualitative trends. UN Department of Economic and Social Affairs (DESA) Working Papers, URL https://www.un-ilibrary.org/content/papers/25206656/158.

Lee, D. J.-L., J. Lo, M. Kim, and E. Paulos, 2016: Crowdclass: Designing classification-based citizen science learning modules. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Austin, Texas, USA, Vol. 4, 109–118.

Lei, B., R. Stouffs, and F. Biljecki, 2023: Assessing and benchmarking 3d city models. *International Journal of Geographical Information Science*, **37**, 788–809.

Levy, M. and M. Germonprez, 2017: The potential for citizen science in information systems research. *Communications of the Association for Information Systems*, **40 (1)**, 22–39, URL https://aisel.aisnet.org/cais/vol40/iss1/2/.

Lewandowski, T., E. Kučević, S. Leible, M. Poser, and T. Böhmann, 2023: Enhancing conversational agents for successful operation: A multi-perspective evaluation approach for continuous improvement. *Electronic Markets*, **33 (1)**, URL https://doi.org/10.1007/s12525-023-00662-3.

Li, X., Y. Bai, and Y. Kang, 2022: Exploring the social influence of the kaggle virtual community on the m5 competition. *International Journal of Forecasting*, **38 (4)**, 1507–1518, URL https://www.sciencedirect.com/science/article/pii/S0169207021001643.

Lia, E. H., M. M. Derrien, E. M. White, S. G. Winder, and S. A. Wood, 2023: A text-messaging chatbot to support outdoor recreation monitoring through community science. *Digital Geography and Society*, **5 (4)**, 100 059, URL https://www.sciencedirect.com/science/article/pii/S2666378323000119.

Liu, H.-Y., D. Dörler, F. Heigl, and S. Grossberndt, 2021: chap. Citizen Science Platforms. Vohland, K., A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, (Eds.), *The Science of Citizen Science*, 439–459, Springer International Publishing, Cham, URL https://link.springer.com/chapter/10.1007/978-3-030-58278-4_22.

Logan, V., 2017: Information as a second language: Enabling data literacy for digital society. Gartner, URL https://www.gartner.com/en/documents/3602517.

Luna, S., M. Gold, A. Albert, L. Ceccaroni, B. Claramunt, O. Danylo, M. Haklay, R. Kottmann, C. Kyba, J. Piera, A. Radicchi, S. Schade, and U. Sturm, 2018: chap. Developing Mobile Applications for Environmental and Biodiversity Citizen Science: Considerations and Recommendations. *Multimedia*

*Tools and Applications for Environmental & Biodiversity Informatics*, 9–30, URL `https://link.springer.com/chapter/10.1007/978-3-319-76445-0_2`.

Mac Domhnaill, C., A. Nolan, and S. Lyons, 2020: The citizens in citizen science: Demographic, socioeconomic, and health characteristics of biodiversity recorders in Ireland. *Citizen Science: Theory and Practice*, **5 (1)**, 1–17, URL `https://theoryandpractice.citizenscienceassociation.org/articles/10.5334/cstp.283`.

Macintosh, A., 2004: Characterizing e-participation in policy-making. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, Big Island, Hawaii, USA, URL `https://ieeexplore.ieee.org/document/1265300`.

Macintosh, A., T. F. Gordon, and A. Renton, 2009: Providing argument support for e-participation. *Journal of Information Technology & Politics*, **6 (1)**, 43–59, URL `http://www.tandfonline.com/doi/abs/10.1080/19331680802662113`.

Makransky, G. and G. B. Petersen, 2021: The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, **33 (3)**, 937–958, URL `https://doi.org/10.1007/s10648-020-09586-2`.

Mamykina, L., T. N. Smyth, J. P. Dimond, and K. Z. Gajos, 2016: Learning from the crowd: Observational learning in crowdsourcing communities. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 2635–2644, URL `https://dl.acm.org/doi/10.1145/2858036.2858560`.

March, S. T. and G. F. Smith, 1995: Design and natural science research on information technology. *Decision Support Systems*, **15 (4)**, 251–266, URL `https://www.sciencedirect.com/science/article/pii/0167923694000412`.

Margaryan, A., 2016: Understanding crowdworkers' learning practices. *Internet, Policy and Politics 2016 Conference*, Glasgow Caledonian University, University of Oxford, UK.

Margaryan, A., 2019: Workplace learning in crowdwork: Comparing microworkers' and online freelancers' practices. *Journal of Workplace Learning*, **31 (4)**, 250–273, URL `https://doi.org/10.1108/JWL-10-2018-0126`.

Martin, D., B. V. Hanrahan, J. O'Neill, and N. Gupta, 2014: Being a turker. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 224–235, URL `http://dx.doi.org/10.1145/2531602.2531663`.

Matsubara, M., R. M. Borromeo, S. Amer-Yahia, and A. Morishima, 2021: Task assignment strategies for crowd worker ability improvement. *Proceedings of the ACM on Human-Computer Interaction*, **5**, Article 375, URL `https://dl.acm.org/doi/10.1145/3479519`.

Matsubara, M., M. Kobayashi, and A. Morishima, 2018: A learning effect by presenting machine prediction as a reference answer in self-correction. *2018 IEEE International Conference on Big Data (Big Data)*, IEEE Computer Society, Tsukuba Japan, 3522–3528.

Matsuura, S. and R. Ishimura, 2017: Chatbot and dialogue demonstration with a humanoid robot in the lecture class. Antona, M. and C. Stephanidis, (Eds.), *Universal Access in Human–Computer Interaction. Human and Technological Environments*, Springer International Publishing AG 2017, 233–246.

McIntosh, C., (Ed.) , 2013: *Cambridge Advanced Learner's Dictionary*, McIntosh, C., (Ed.). Cambridge University Press, Cambridge.

Metag, J. and G. Gurr, 2023: Too much information? A longitudinal analysis of information overload and avoidance of referendum information prior to voting day. *Journalism & Mass Communication Quarterly*, **100 (3)**, 646, URL `https://doi.org/10.1177/10776990221127380`.

Miller-Rushing, A., R. Primack, and R. Bonney, 2012: The history of public participation in ecological research. *Frontiers in Ecology and the Environment*, **10 (6)**, 285–290, URL `https://esajournals.onlinelibrary.wiley.com/doi/full/10.1890/110278`.

Moczek, N., S. Hecker, and S. L. Voigt-Heucke, 2021: The known unknowns: What citizen science projects in germany know about their volunteers—and what they don't know. *Sustainability*, **13 (20)**, 11 553, URL `https://www.mdpi.com/2071-1050/13/20/11553`.

Monzón Alvarado, C. M., A. Zamora Rendon, and A. Vázquez Pérez, 2020: Integrating public participation in knowledge generation processes: Evidence from citizen science initiatives in Mexico. *Environmental Science & Policy*, **114**, 230–241, URL `https://www.sciencedirect.com/science/article/pii/S1462901120300575`.

Musto, J. and A. Dahanayake, 2021: An approach to improve the quality of user-generated content of citizen science platforms. *ISPRS International Journal of Geo-Information*, **10 (7)**, 434, URL `https://www.mdpi.com/2220-9964/10/7/434`.

Nakayama, T., M. Matsubara, and A. Morishima, 2021: Crowd-worker skill improvement with AI co-learners. *Proceedings of the 9th International Conference on Human-Agent Interaction*, Association for Computing Machinery, New York, NY, USA, 316–322, URL `https://doi.org/10.1145/3472307.3484684`.

Narechania, A., A. Srinivasan, and J. Stasko, 2021: NL4dv: A toolkit for generating analytic specifications for data visualization from natural language queries. *IEEE Transactions on Visualization and Computer Graphics*, **27 (2)**, 369–379.

National Academies of Sciences, Engineering and Education, Division of Behavioral and Social Sciences and Education, Board on Science and Learning, and Committee on Designing Citizen Science to Support Science, 2018: *Designing for Learning*, Dibner, K. A. and R. Pandya, (Eds.). National Academies Press (US), Washington (DC), USA, URL https://www.ncbi.nlm.nih.gov/books/NBK535956/.

Neumaier, S., V. Savenkov, and S. Vakulenko, 2017: Talking open data. Blomqvist, E., K. Hose, H. Paulheim, A. Ławrynowicz, F. Ciravegna, and O. Hartig, (Eds.), *The Semantic Web: ESWC 2017 Satellite Events*, Springer, Vol. vol 10577, 132–136, URL https://link.springer.com/chapter/10.1007/978-3-319-70407-4_25#citeas.

Newman, G., D. Zimmerman, A. Crall, M. Laituri, J. Graham, and L. Stapel, 2010: User-friendly web mapping: Lessons from a citizen science website. *International Journal of Geographical Information Science*, **24 (12)**, 1851–1869, URL https://www.tandfonline.com/doi/full/10.1080/13658816.2010.490532.

Nguyen, H., J. Ahn, W. Young, and F. Campos, 2020: Where's the learning in education crowdsourcing? *Proceedings of the 7th ACM Conference on Learning @ Scale*, Association for Computing Machinery, Virtual Event, 305–308, URL https://dl.acm.org/doi/10.1145/3386527.3406734.

Nguyen, H. D., V. T. Pham, D. A. Tran, and T. T. Le, 2019: Intelligent tutoring chatbot for solving mathematical problems in high-school. *Proceedings of the 11th International Conference on Knowledge and Systems Engineering (KSE)*, Da Nang, Vietnam, 1–6.

Nickerson, R. C., U. Varshney, and J. Muntermann, 2013: A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, **22**, 336–359, URL https://link.springer.com/article/10.1057/ejis.2012.26.

Okonkwo, C. W. and A. Ade-Ibijola, 2021: Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, **2**.

Onyimbi, J., M. Koeva, and J. Flacke, 2017: Public participation using 3d city models: E-participation opportunities in kenya. *GIM International*, **31 (7)**, 29–31.

Osterwalder, A. and Y. Pigneur, 2010: *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. 1st ed., John Wiley & Sons, Hoboken, NJ, United States.

Paetsch, F., A. Eberlein, and F. Maurer, 2003: Requirements engineering and agile software development. *Proceedings of the 12th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE 2003)*, Linz, Austria, 308–313.

Palacin, V., M. Nelimarkka, P. Reynolds-Cuéllar, and C. Becker, 2020: The design of pseudo-participation. *Proceedings of the 16th Participatory Design Conference - Participation(s) Other-*

*wise*, Association for Computing Machinery, Manizales, Colombia, Vol. 2, 40–44, URL https://doi.org/10.1145/3384772.3385141.

Palacin, V., A. Zundel, V. Aquaro, and W. M. Kwok, 2021: Reframing e-participation for sustainable development. *14th International Conference on Theory and Practice of Electronic Governance*, ACM, Athens Greece, 172–180, URL https://dl.acm.org/doi/10.1145/3494193.3494218.

Palan, S. and C. Schitter, 2018: Prolific.ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, **17**, 22–27, URL https://www.sciencedirect.com/science/article/pii/S2214635017300989.

Pandey, S., O. G. McKinley, R. J. Crouser, and A. Ottley, 2023: Do you trust what you see? Toward a multidimensional measure of trust in visualization. *2023 IEEE Visualization and Visual Analytics (VIS)*, Melbourne, Australia, 26–30.

Panetta, K., 2021: A data and analytics leader's guide to data literacy. Gartner, URL https://www.gartner.com/smarterwithgartner/a-data-and-analytics-leaders-guide-to-data-literacy.

Parsons, J. and R. Lukyanenko, 2011: Rethinking data quality as an outcome of conceptual modeling choices. *Proceedings of the 16th International Conference on Information Quality (ICIQ-11)*, Computer Science, Adelaide, Australia, 244–258.

Pedersen, A. Y. and F. Caviglia, 2019: Data literacy as a compound competence. Antipova, T. and A. Rocha, (Eds.), *Digital Science*, Springer International Publishing, Budva, Vol. 850, 166–173.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, 2007: A design science research methodology for information systems research. *Journal of Management Information Systems*, **24 (3)**, 45–77, URL https://doi.org/10.2753/MIS0742-1222240302.

Peninsula, T., 2009: Much of internet yet to come founders. *The Peninsula - Online*, URL https://cds.cern.ch/record/1224215.

Pettibone, L., K. Vohland, and D. Ziegler, 2017: Understanding the (inter)disciplinary and institutional diversity of citizen science: A survey of current practice in germany and austria. *PLoS ONE*, **12 (6)**, 1–16.

Peukert, C., J. Pfeiffer, M. Meißner, T. Pfeiffer, and C. Weinhardt, 2019: Acceptance of imagined versus experienced virtual reality shopping environments: Insights from two experiments. *Proceedings of the 27th European Conference on Information Systems (ECIS 2019)*, Stockholm & Uppsala, Sweden, 153.

Phillips, T., N. Porticella, M. Constas, and R. Bonney, 2018: A framework for articulating and measuring individual learning outcomes from participation in citizen science. *Citizen Science: Theory and Practice*, **3 (2)**, 3, URL http://theoryandpractice.citizenscienceassociation.org/article/10.5334/cstp.126/.

Pirannejad, A., M. Janssen, and J. Rezaei, 2019: Towards a balanced e-participation index: Integrating government and society perspectives. *Government Information Quarterly*, **36 (4)**, 101 404, URL https://www.sciencedirect.com/science/article/pii/S0740624X17302836.

Polak, J. and D. Cook, 2021: A study on student performance, engagement, and experience with Kaggle InClass data challenges. *Journal of Statistics and Data Science Education*, **29 (1)**, 63–70, URL https://doi.org/10.1080/10691898.2021.1892554.

Portela, M., 2021: Interfacing participation in citizen science projects with conversational agents. *Human Computation*, **8 (2)**, 33–53.

Pristl, A.-C. and M. Billert, 2022: Citizen participation in increasingly digitalized governmental environments – a systematic literature review. *Proceedings of the 30th European Conference of Information Systems (ECIS 2022)*, Timisoara, Romania, URL https://aisel.aisnet.org/ecis2022_rp/28.

Prpić, J., P. P. Shukla, J. H. Kietzmann, McCarthy, and P. Ian, 2015a: How to work a crowd: Developing crowd capital through crowdsourcing. *Business Horizons*, **58 (1)**, 77–85.

Prpić, J., A. Taeihagh, and J. Melton, 2015b: The fundamentals of policy crowdsourcing. *Policy Internet*, **7 (3)**, 340–361.

Pullinger, J., 2021: Misuse of statistics: Time to speak out. *Statistical Journal of the IAOS*, **37 (1)**, 79–84, URL https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji210783.

Pérez, J. Q., T. Daradoumis, and J. M. M. Puig, 2020: Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, **28 (6)**, 1549–1565.

Radchenko, I. and O. Maksimenkova, 2016: Principles of citizen science in open educational projects based on open data. *Proceedings of the 12th Central and Eastern European Software Engineering Conference in Russia*, ACM, Moscow Russia, 1–5, URL https://dl.acm.org/doi/10.1145/3022211.3022216.

Radermacher, W., 2021: Literacy in statistics for the public discourse. *Statistical Journal of the IAOS*, **37 (3)**, 747–752.

Radermacher, W., 2024: Official statistics: Language for public discourse. Data & Policy Blog, Medium, URL https://medium.com/data-policy/

official-statistics-language-for-public-discourse-part-1-294d31e56b6e, [Accessed: Feb. 2025].

Radziwill, N. M. and M. C. Benton, 2017: Evaluating quality of chatbots and intelligent conversational agents. arXiv, 1704.04579, URL http://arxiv.org/abs/1704.04579.

Rapp, A., L. Curti, and A. Boldi, 2021: The human side of human-chatbot interaction: A systematic literature review of ten dates of research on text-based chatbots. *International Journal of Human-Computer Studies*, **151**, 102 630, URL https://www.sciencedirect.com/science/article/abs/pii/S1071581921000483.

Rauschnabel, P. A., R. Felix, C. Hinsch, H. Shahab, and F. Alt, 2022: What is XR? Towards a framework for augmented and virtual reality. *Computers in Human Behavior*, **133**, 107 289, URL https://www.sciencedirect.com/science/article/pii/S074756322200111X.

Reed, J., W. Rodriguez, and A. Rickhoff, 2012: A framework for defining and describing key design features of virtual citizen science projects. *Proceedings of the 2012 iConference*, Association for Computing Machinery, New York, NY, USA, 623–625, URL https://doi.org/10.1145/2132176.2132314.

Reinwald, F., M. Berger, C. Stoik, M. Platzer, and D. Damyanovic, 2014: Augmented reality at the service of participatory urban planning and community informatics – A case study from vienna. *The Journal of Community Informatics*, **10 (3)**.

Ridsdale, C., J. Rothwell, M. Smit, M. Bliemel, D. Irvine, D. Kelley, S. Matwin, B. Wuetherick, and H. Ali-Hassan, 2015: *Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report*. Dalhousie University, URL http://hdl.handle.net/10222/64578.

Roberts, T., P. Lowry, and P. Sweeney, 2006: An evaluation of the impact of social presence through group size and the use of collaborative software on group member "voice" in face-to-face and computer-mediated task groups. *IEEE Transactions on Professional Communication*, **49 (1)**, 28–43, URL https://ieeexplore.ieee.org/abstract/document/1599552.

Rogers, Y., 1992: Coordinating computer-mediated work. *Computer Supported Cooperative Work (CSCW)*, **1 (4)**, 295–315, URL https://doi.org/10.1007/BF00754332.

Romesburg, H. C., 2004: *Cluster Analysis for Researchers*. Lulu Press, Morrisville, URL https://digitalcommons.usu.edu/envs_facpub/42/.

Roumpani, F., 2022: Procedural cities as active simulators for planning. *Urban Planning*, **7 (2)**, 321–329, URL https://doi.org/10.17645/up.v7i2.5209.

Rousseeuw, P. J., 1987: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, **20**, 53–65, URL `https://www.sciencedirect.com/science/article/pii/0377042787901257`.

Royo, S. and A. Yetano, 2015: "Crowdsourcing" as a tool for e-participation: two experiences regarding CO2 emissions at municipal level. *Electronic Commerce Research*, **15 (3)**, 323–348, URL `https://doi.org/10.1007/s10660-015-9183-6`.

Ryan, R. M., 1982: Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, **43 (3)**, 450–461, URL `http://dx.doi.org/10.1037/0022-3514.43.3.450`.

Rzeszewski, M. and M. Orylski, 2021: Usability of WebXR visualizations in urban planning. *ISPRS International Journal of Geo-Information*, **10 (11)**, 721, URL `https://www.mdpi.com/2220-9964/10/11/721`.

Sanford, C. and J. Rose, 2007: Characterizing eParticipation. *International Journal of Information Management*, **27 (6)**, 406–421, URL `http://dx.doi.org/10.1016/j.ijinfomgt.2007.08.002`.

Santamaría-Philco, A., J. H. Canós Cerdá, and M. C. Penadés Gramaje, 2019: Advances in e-participation: A perspective of last dates. *IEEE Access*, **7**, 155 894–155 916.

Saraswat, S., 1999: Structured content analysis of corporate web-sites: A methodological perspective. *AMCIS 1999 Proceedings 73*, 206–208, URL `https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1430&context=amcis1999`.

Sasmito, G. W., L. O. M. Zulfiqar, and M. Nishom, 2019: Usability testing based on system usability scale and net promoter score. *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Yogyakarta, Indonesia, 540–545.

Schader, M., D. Geiger, M. Rosemann, and E. Fielt, 2012: Crowdsourcing information systems - Definition, typology, and design. *Proceedings of the 33rd International Conference on Information Systems (ICIS 2012)*, Orlando, Florida, USA, Vol. 4.

Schilling, R., M. Haki, and S. Aier, 2017: Introducing archetype theory to information systems research: A literature review and call for future research. *Wirtschaftsinformatik 2017 Proceedings*, URL `https://aisel.aisnet.org/wi2017/track05/paper/10`.

Schleicher, K. and C. Schmidt, 2020: Citizen science in Germany as research and sustainability education: Analysis of the main forms and foci and its relation to the sustainable development goals. *Sustainability*, **12 (15)**, 6044.

Schmidt, F. A., 2017: *Digital Labour Markets in the Platform Economy: Mapping the Political Challenges of Crowd Work and Gig Work*. Friedrich-Ebert-Stiftung, Bonn, Germany.

Schmidthuber, L., D. Hilgers, and K. Randhawa, 2022: Public crowdsourcing: Analyzing the role of government feedback on civic digital platforms. *Public Administration*, **100 (4)**, 960–977, URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/padm.12811`.

Schrepp, M., A. Hinderks, and J. Thomaschewski, 2017: Design and evaluation of a short version of the user experience questionnaire (UEQ-s). *International Journal of Interactive Multimedia and Artificial Intelligence*, **4 (6)**, 103–108, URL `http://dx.doi.org/10.9781/ijimai.2017.09.001`.

Schrom-Feiertag, H., M. Stubenschrott, G. Regal, T. Matyus, and S. Seer, 2020: An interactive and responsive virtual reality environment for participatory urban planning. *Proceedings of the 11th Annual Symposium on Simulation for Architecture and Urban Design (SimAUD 2020)*, Virtual Event Austria, 1–7.

Schubert, D., 2015: chap. Stadtplanung – Wandlungen einer Disziplin und zukünftige Herausforderungen. *Stadt und Gesellschaft im Fokus aktueller Stadtforschung. Konzepte-Herausforderungen-Perspektiven*, 121–176, Springer, URL `https://link.springer.com/chapter/10.1007/978-3-658-07384-8_5`.

Schüller, K., P. Busch, and C. Hindinger, 2019: *Future Skills: a Framework for Data Literacy. Kompetenzrahmen und Forschungsbericht*, Vol. 47. Hochschulforum Digitalisierung, Berlin, URL `https://hochschulforumdigitalisierung.de/wp-content/uploads/2023/09/HFD_AP_Nr_47_DALI_Kompetenzrahmen_WEB.pdf`.

Seeger, A.-M., J. Pfeiffer, and A. Heinzl, 2021: Texting with humanlike conversational agents: Designing for anthropomorphism. *Journal of the Association for Information Systems*, **22 (4)**, 931–967, URL `https://aisel.aisnet.org/jais/vol22/iss4/8/`.

Shah, H. R. and L. R. Martinez, 2016: Current approaches in implementing citizen science in the classroom. *Journal of Microbiology & Biology Education*, **17 (1)**, 17–22, URL `https://journals.asm.org/doi/full/10.1128/jmbe.v17i1.1032`.

Shanley, L. A., A. Parker, S. Schade, and A. Bonn, 2019: Policy perspectives on citizen science and crowdsourcing. *Citizen Science: Theory and Practice*, **4 (1)**, URL `https://theoryandpractice.citizenscienceassociation.org/articles/10.5334/cstp.293`.

Sharma, S. and S. Pandey, 2013: Revisiting requirements elicitation techniques. *International Journal of Computer Applications*, **75 (12)**, 35–39, URL `https://www.ijcaonline.org/archives/volume75/number12/13166-0889/`.

Shen, Y., X. Ai, A. G. Soosai Raj, R. J. Leo John, and M. Syamkumar, 2024: Implications of chatgpt for data science education. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education*, Association for Computing Machinery, New York, NY, USA, 1230–1236, URL https://doi.org/10.1145/3626252.3630874.

Shirk, J. L., H. L. Ballard, C. C. Wilderman, T. Phillips, A. Wiggins, R. Jordan, E. McCallie, M. Minarchek, B. V. Lewenstein, K. M. E, and R. Bonney, 2012: Public participation in scientific research: A framework for deliberate design. *Ecology and Society*, **17**, 29–48.

Shirk, J. L. and R. Bonney, 2018: chap. Scientific impacts and innovations of citizen science. *Citizen Science: Innovation in Open Science, Society and Policy*, 41–51, UCL Press, London, URL https://www.jstor.org/stable/j.ctv550cf2.10.

Simonofski, A., A. Zuiderwijk, A. Clarinval, and W. Hammedi, 2022: Tailoring open government data portals for lay citizens: A gamification theory approach. *International Journal of Information Management*, **65**, 102 511, URL https://www.sciencedirect.com/science/article/pii/S0268401222000421.

Simpson, R., K. R. Page, and D. De Roure, 2014: Zooniverse: Observing the world's largest citizen science platform. *Proceedings of the 23rd International Conference on World Wide Web*, Association for Computing Machinery, New York, NY, USA, 1049–1054, URL https://doi.org/10.1145/2567948.2579215.

Simud, T., S. Ruengittinun, N. Surasvadi, N. Sanglerdsinlapachai, and A. Plangprasopchok, 2020: A conversational agent for database query: A use case for thai people map and analytics platform. *2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing*, Bangkok, Thailand, 1–6.

Singla, A., I. Bogunovic, G. Bartók, A. Karbasi, and A. Krause, 2014: Near-optimally teaching the crowd to classify. arXiv, 1402.2092, URL http://arxiv.org/abs/1402.2092.

Skarlatidou, A., A. Hamilton, M. Vitos, and M. Haklay, 2019: What do volunteers want from citizen science technologies? A systematic literature review and best practice guidelines. *Journal of Science Communication*, **18 (1)**, A02, URL https://jcom.sissa.it/archive/18/01/JCOM_1801_2019_A02.

Socientize Consortium, 2014: Green paper on citizen science for europe: Towards a society of empowered citizens and enhanced research. European Union, URL https://eu-citizen.science/resource/9.

Somech, A., 2002: Explicating the complexity of participative management: An investigation of multiple dimensions. *Educational Administration Quarterly*, **38 (3)**, 341–371, URL https://www.academia.edu/34269022/Explicating_the_Complexity_of_Participative_Management_An_Investigation_of_Multiple_Dimensions.

Sorensen, A. E., R. C. Jordan, S. L. LaDeau, D. Biehler, S. Wilson, J.-H. Pitas, and P. T. Leisnham, 2019: Reflecting on efforts to design an inclusive citizen science project in West Baltimore. *Citizen Science: Theory and Practice*, **4 (1)**.

Soßdorf, A., C. Stein, I. Bezzaoui, and J. Fegert, 2024: Literacies against fake news: Examining the role of data literacy and critical media literacy to counteract disinformation. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, **59**, 55–76.

Spann, M. and B. Skiera, 2003: Internet-based virtual stock markets for business forecasting. *Management Science*, **49 (10)**, 1310—-1326.

Spasiano, A., S. Grimaldi, A. M. Braccini, and F. Nardi, 2021: Towards a transdisciplinary theoretical framework of citizen science: Results from a meta-review analysis. *Sustainability*, **13 (14)**, 7904, URL https://www.mdpi.com/2071-1050/13/14/7904.

Spiers, H., A. Swanson, L. Fortson, B. Simmons, L. Trouille, S. Blickhan, and C. Lintott, 2019: Everyone counts? Design considerations in online citizen science. *Journal of Science Communication*, **18 (1)**, 32, URL https://jcom.sissa.it/article/pubid/JCOM_1801_2019_A04/.

Splichal, S., 2022: *Datafication of Public Opinion and the Public Sphere*. Anthem Press, London, UK.

Srnicek, N., 2016: *Platform Capitalism*. Polity, Cambridge, UK.

Stein, C. and J. Fegert, 2024: Bridging realities: Exploring enablement factors for XR participatory urban planning. *Proceedings of the 32nd European Conference on Information Systems (ECIS 2024)*, Paphos, Cyprus, URL https://aisel.aisnet.org/ecis2024/track17_greenis/track17_greenis/31.

Stein, C., J. D. Fegert, A. Wittmer, and C. Weinhardt, 2023a: Digital participation for data literate citizens – a qualitative analysis of the design of multi-project citizen science platforms. *IADIS International Journal on Computer Science and Information Systems*, **18 (1)**, 1–17.

Stein, C., M. Müller, and J. Fegert, 2024a: Designing (for) change: A taxonomy-based approach to project design. *ARPHA Proceedings*, **6**, 73–77, URL https://doi.org/10.3897/ap..e126585.

Stein, C., A. Sossdorf, O. K. Kirikci, and J. Fegert, 2024b: Learning while earning? A literature review and case study on learning opportunities in crowdwork. *Wirtschaftsinformatik 2024 Proceedings*, URL https://aisel.aisnet.org/wi2024/22.

Stein, C., T. Straub, J. Jachimowicz, and J. Fegert, 2023b: Same same but different - Towards a taxonomy for digital involvement projects. *Proceedings of the 31st European Conference on Information Systems (ECIS 2023)*, AIS eLibrary (AISeL), Kristiansand, Norway, URL https://aisel.aisnet.org/ecis2023_rp/258.

Stein, C., T. Teubner, and S. Morana, 2024c: Designing a conversational agent for supporting data exploration in citizen science. *Electronic Markets*, **34 (1)**, 23, URL https://doi.org/10.1007/s12525-024-00705-3.

Stein, C., A. Wittmer, L. Buß, and J. Fegert, 2025a: The devil is in the details? Investigating 3d visualization types for e-participation in urban planning. *Proceedings of the 58th Annual Hawaii International Conference on System Sciences*, Big Island, Hawaii, Forthcoming.

Stein, C., A. Wittmer, J. Fegert, and C. Weinhardt, 2023c: Citizen science as a service? A review of multi-project citizen science platforms. *Proceedings of the ES 2023 + ML 2023*, Lisbon, Portugal, 36–55.

Stein, C., A. Wittmer, C. Weinhardt, and J. Fegert, 2025b: From (design) theory to (participation) practice: Leveraging a taxonomy for digital involvement projects. *Proceedings of the 58th Annual Hawaii International Conference on System Sciences*, Big Island, Hawaii, Forthcoming.

Steinbach, M., N. Wilker, and S. Schöttle, 2020: E-participation on the local level – A census survey approach for researching its implementation. *Journal of Information Technology & Politics*, **17 (1)**, 12–32, URL https://doi.org/10.1080/19331681.2019.1676361.

Stieglitz, S., L. Hofeditz, F. Brünker, C. Ehnis, M. Mirbabaie, and B. Ross, 2022: Design principles for conversational agents to support emergency management agencies. *International Journal of Information Management*, **63**, 102 469.

Straub, T., H. Gimpel, F. Teschner, and C. Weinhardt, 2015: How (not) to incent crowd workers. *Business & Information Systems Engineering*, **57 (3)**, 167–179, URL https://doi.org/10.1007/s12599-015-0384-2.

Sturm, U., S. Schade, L. Ceccaroni, M. Gold, C. Kyba, B. Claramunt, M. Haklay, D. Kasperowski, A. Albert, J. Piera, J. Brier, C. Kullenberg, and S. Luna, 2018: Defining principles for mobile apps and platforms development in citizen science. *Research Ideas and Outcomes*, **4 (3)**, 13, URL https://zenodo.org/record/1150156/export/hx.

Suzuki, R., N. Salehi, M. S. Lam, J. C. Marroquin, and M. S. Bernstein, 2016: Atelier: Repurposing expert crowdsourcing tasks as micro-internships. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, arXiv, San Jose, California, USA, 2645–2656, URL http://arxiv.org/abs/1602.06634, 1602.06634[cs].

Szopinski, D., T. Schoormann, and D. Kundisch, 2019: Because your taxonomy is worth it: Towards a framework for taxonomy evaluation. *Proceedings of the 27th European Conference on Information Systems (ECIS 2019)*, Stockholm-Uppsala, Sweden, 2–20.

Tallyn, E., H. Fried, R. Gianni, A. Isard, and C. Speed, 2018: The ethnobot: Gathering ethnographies in the age of IoT. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018)*, Montreal, Canada, 1–13, URL `https://dl.acm.org/doi/10.1145/3173574.3174178`.

Tamayo-Moreno, S. and D. Pérez-Marín, 2016: Adapting the design and the use methodology of a pedagogical conversational agent of secondary education to childhood education. *2016 International Symposium on Computers in Education (SIIE)*, IEEE Computer Society, Salamanca, Spain, 1–6.

Tan, Y.-R., A. Agrawal, M. P. Matsoso, R. Katz, S. L. M. Davis, A. S. Winkler, A. Huber, A. Joshi, A. El-Mohandes, B. Mellado, C. A. Mubaira, F. C. Canlas, G. Asiki, H. Khosa, J. V. Lazarus, M. Choisy, M. Recamonde-Mendoza, O. Keiser, P. Okwen, R. English, S. Stinckwich, S. Kiwuwa-Muyingo, T. Kutadza, T. Sethi, T. Mathaha, V. K. Nguyen, A. Gill, and P. Yap, 2022: A call for citizen science in pandemic preparedness and response: Beyond data collection. *BMJ Global Health*, **7 (6)**, e009 389, URL `https://doi.org/10.1136/bmjgh-2022-009389`.

Tauchert, C., P. Buxmann, and J. Lambinus, 2020: Crowdsourcing data science: A qualitative analysis of organizations' usage of Kaggle competitions. *Proceedings of the 53rd Annual Hawaii International Conference on System Sciences*, Maui, Hawaii, USA, 10, URL `http://hdl.handle.net/10125/63768`.

Tauginienė, L., E. Butkevičienė, K. Vohland, B. Heinisch, M. Daskolia, M. Suškevičs, M. Portela, B. Balázs, and B. Prūse, 2020: Citizen science in the social sciences and humanities: the power of interdisciplinarity. *Palgrave Communications*, **6 (1)**, 89.

Tavakol, M. and R. Dennick, 2011: Making sense of cronbach's alpha. *International Journal of Medical Education*, **2**, 53–55, URL `https://doi.org/10.5116/ijme.4dfb.8dfd`.

Tavanapour, N., M. Poser, and E. A. C. Bittner, 2019: Supporting the idea generation process in citizen participation - Toward an interactive system with a conversational agent as facilitator. *Proceedings of the 27th European Conference on Information Systems (ECIS 2019)*, Stockholm and Uppsala, Sweden, URL `https://aisel.aisnet.org/ecis2019_rp/70/`.

Tech at Meta, 2021: Connect 2021: Our vision for the metaverse. URL `https://tech.facebook.com/reality-labs/2021/10/connect-2021-our-vision-for-the-metaverse/`, [Accessed: Feb. 2025].

Terwiesch, C. and Y. Xu, 2008: Innovation contests, open innovation, and multiagent problem solving. *Management Science*, **54 (9)**, 1529–1543.

Teschner, F., R. R. A. Mazarakis, and C. Weinhardt, 2011: Participation, feedback incentives in a competitive forecasting community. Beath, C., M. D. Myers, and K. K. Wei, (Eds.), *Proceedings of the 32nd International Conference on Information Systems (ICIS 2011)*, Shanghai, China.

Tong, J., 2024: *Data Journalism and the COVID-19 Disruption*, Tong, J., (Ed.). Routledge & CRC Press.

Trzaskowski, J., 2022: Data-driven value extraction and human well-being under EU law. *Electron Markets*, **32 (2)**, 447–458, URL `https://doi.org/10.1007/s12525-022-00528-0`.

Turner, J. A., 1984: Computer mediated work: The interplay between technology and structured jobs. *Communications of the ACM*, **27 (12)**, 1210–1217, URL `https://doi.org/10.1145/2135.2139`.

Twidale, M. B., C. Blake, and J. P. Gant, 2013: Towards a data literate citizenry. *iConference 2013 Proceedings*, 247–257, URL `https://hdl.handle.net/2142/38385`.

United Nations, 2018: World urbanization prospects: The 2018 revision. Population Division of the United Nations Department of Economic and Social Affairs (UN DESA), New York, NY, USA.

United Nations, 2022: United Nations e-government survey 2022: The future of digital government. United Nations, Department of Economic and Social Affairs (UN DESA), URL `https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2022`.

Van der Goot, M. J., L. Hafkamp, and Z. Dankfort, 2021: Customer service chatbots: A qualitative interview study into the communication journey of customers. Følstad, A., T. Araujo, S. Papadopoulos, E. L.-C. Law, E. Luger, M. Goodwin, and P. B. Brandtzaeg, (Eds.), *Chatbot Research and Design. Conversations 2020. Lecture Notes in Computer Science*, Springer, Cham, Virtual Event, Vol. 12604, 190–204, URL `https://link.springer.com/chapter/10.1007/978-3-030-68288-0_13`.

Van Dijck, J., 2014: Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, **12 (2)**, 197–208, URL `https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/datafication`.

Van Dijk, J. A., 2012: Digital democracy: Vision and reality. *Innovation and the Public Sector*, **19**, 49–62, URL `http://dx.doi.org/10.3233/978-1-61499-137-3-49`.

Van Leeuwen, J. P., K. Hermans, A. Jylhä, A. J. Quanjer, and H. Nijman, 2018: Effectiveness of virtual reality in participatory urban planning: A case study. *Proceedings of the 4th Media Architecture Biennale Conference*, ACM, Beijing China, 128–136, URL `https://dl.acm.org/doi/10.1145/3284389.3284491`.

Venable, J., J. Pries-Heje, and R. Baskerville, 2016: FEDS: a framework for evaluation in design science research. *European Journal of Information Systems*, **25 (1)**, 77–89, URL `https://doi.org/10.1057/ejis.2014.36`.

Venkatesh, V., M. G. Morris, G. B. Davis, and F. D. Davis, 2003: User acceptance of information technology: Toward a unified view. *MIS Quarterly*, **27 (3)**, 425–478, URL `https://www.jstor.org/stable/30036540`.

Videira Lopes, C. and C. Lindstrom, 2012: Virtual cities in urban planning: The Uppsala case study. *Journal of theoretical and applied electronic commerce research*, **7 (3)**, 88–100.

Vinuales, G., S. R. Magnotta, E. Steffes, and G. Kulkarni, 2019: Description and evaluation of an innovative segmentation, targeting, and positioning activity using student perceived learning and actual student learning. *Marketing Education Review*, **29 (1)**, 24–36, URL `https://doi.org/10.1080/10528008.2018.1493932`.

Vohland, K., C. Göbel, B. Balázs, E. Butkevičienė, Daskolia, B. Duží, S. Hecker, M. Manzoni, , and S. Schade, 2021a: chap. Citizen Science in Europe. Vohland, K., A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, (Eds.), *The Science of Citizen Science*, 33–53, Springer International Publishing, Cham.

Vohland, K., A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, 2021b: chap. Editorial: The Science of Citizen Science Evolves. Vohland, K., A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht, (Eds.), *The Science of Citizen Science*, 1–12, Springer International Publishing, Cham, URL `https://doi.org/10.1007/978-3-030-58278-4_1`.

Vom Brocke, J., A. Hevner, and A. Mädche, 2020: chap. Introduction to Design Science Research. vom Brocke, J., A. Hevner, and A. Maedche, (Eds.), *Design Science Research. Cases*, 1–13, Springer International Publishing, Cham, URL `https://doi.org/10.1007/978-3-030-46781-4_1`.

Wald, D. M., J. Longo, and A. Dobell, 2016: Design principles for engaging and retaining virtual citizen scientists. *Conservation Biology*, **30 (3)**, 562–570, URL `https://www.jstor.org/stable/24760984`.

Wang, A. Y., D. Wang, J. Drozdal, X. Liu, S. Park, S. Oney, and C. Brooks, 2021: What makes a well-documented notebook? A case study of data scientists' documentation practices in Kaggle. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 1–7, URL `https://doi.org/10.1145/3411763.3451617`.

Wang, N.-C., D. Hicks, and K. Luther, 2018: Exploring trade-offs between learning and productivity in crowdsourced history. *Proceedings of the ACM on Human-Computer Interaction*, **2**, Article 178, URL https://doi.org/10.1145/3274447.

Watson, J. and R. Callingham, 2004: Statistical literacy: From idiosyncratic to critical thinking abstract. Burrill, G. and M. Camden, (Eds.), *Curricular Development in Statistics Education International Association for Statistical Education (IASE) Roundtable*, International Statistical Institute Voorburg, Lund, Sweden, 116–162.

Webster, J. and R. T. Watson, 2002: Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, **26 (2)**, 11, URL https://www.jstor.org/stable/4132319.

Weinhardt, C., J. Fegert, O. Hinz, and W. M. P. Van der Aalst, 2024: Digital democracy: A wake-up call. *Business  Information Systems Engineering*, **66 (2)**, 127–134.

Weinhardt, C., S. Kloker, O. Hinz, and W. M. P. Van der Aalst, 2020: Citizen science in information systems research. *Business & Information Systems Engineering*, **62 (4)**, 273–277.

Wiggins, A. and J. Wilbanks, 2019: The rise of citizen science in health and biomedical research. *The American Journal of Bioethics*, **19 (8)**, 3–14.

Wikipedia, 2020: Category:internet-based activism. URL https://en.wikipedia.org/w/index.php?title=Category:Internet-based_activism&oldid=980867875, [Accessed: Jun. 2024].

Wikipedia, 2024: List of crowdsourcing projects. URL https://en.wikipedia.org/w/index.php?title=List_of_crowdsourcing_projects&oldid=1218020924, [Accessed: Jun. 2024].

Wirtz, B. W., J. C. Weyerer, M. Becker, and W. M. Müller, 2022: Open government data: A systematic literature review of empirical research. *Electronic Markets*, **32 (4)**, 2381–2404, URL https://doi.org/10.1007/s12525-022-00582-8.

Wise, A. F., 2020: Educating data scientists and data literate citizens for a new generation of data. *Journal of the Learning Sciences*, **29 (1)**, 165–181, URL https://doi.org/10.1080/10508406.2019.1705678.

Wissen, U., O. Schroth, E. Lange, and W. A. Schmid, 2008: Approaches to integrating indicators into 3d landscape visualisations and their benefits for participative planning situations. *Journal of Environmental Management*, **89 (3)**, 184–196, URL https://doi.org/10.1016/j.jenvman.2007.01.062.

Wolf, M., H. Söbke, and F. Wehking, 2020: chap. Mixed Reality Media-Enabled Public Participation in Urban Planning. Progress in IS, Jung, T., M. C. tom Dieck, and P. A. Rauschnabel, (Eds.), *Augmented Reality and Virtual Reality: Changing Realities in a Dynamic World*, 125–138, Springer International Publishing, Cham, URL https://doi.org/10.1007/978-3-030-37869-1_11.

Wormer, H., 2020: German media and coronavirus: Exceptional communication—or just a catalyst for existing tendencies? *Health and Science Controversies in the Digital World: News, Mis/Disinformation and Public Engagement*, **8 (2)**, 467–470, URL `https://doi.org/10.17645/mac.v8i2.3242`.

Wu, H., Z. He, and J. Gong, 2010: A virtual globe-based 3d visualization and interactive framework for public participation in urban planning processes. *Computers, Environment and Urban Systems*, **34**, 291–298.

Yadav, P. and J. Darlington, 2016: Design guidelines for the user-centred collaborative citizen science platforms. *Human Computation*, **3 (1)**, 205–211.

Yates, S. J., E. Carmi, E. Lockley, B. Wessels, and A. Pawluczuk, 2021: Understanding citizens data literacies research report. University of Liverpool, URL `https://www.nuffieldfoundation.org/wp-content/uploads/2019/11/Understanding-citizens-data-literacies.pdf`.

Ye, J. and M. Jensen, 2022: Effects of introducing an online community in a crowdsourcing contest platform. *Information Systems Journal*, **32 (6)**, 1203–1230, URL `https://doi.org/10.1111/isj.12397`.

Zagalsky, A., J. Feliciano, M.-A. Storey, Y. Zhao, and W. Wang, 2015: The emergence of GitHub as a collaborative platform for education. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Association for Computing Machinery, New York, NY, USA, 1906–1917, URL `https://dl.acm.org/doi/10.1145/2675133.2675284`.

Zarifhonarvar, A., 2023: Economics of ChatGPT: a labor market view on the occupational impact of artificial intelligence. *Journal of Electronic Business & Digital Economics*, **3 (2)**, 17, URL `https://doi.org/10.1108/JEBDE-10-2023-0021`.

Zhang, Y., P. Calyam, T. Joshi, S. Nair, and D. Xu, 2018: Domain-specific topic model for knowledge discovery through conversational agents in data intensive scientific communities. *Proceedings of 2018 IEEE International Conference on Big Data, Big Data 2018*, IEEE, Seattle, USA, 4886–4895, URL `https://doi.org/10.1109/BigData.2018.8622309`.

Zhao, Y. and Q. Zhu, 2014: Evaluation on crowdsourcing research: Current status and future direction. *Information Systems Frontiers*, **16 (3)**, 417–434, URL `https://doi.org/10.1007/s10796-012-9350-4`.

Zheng, Y., 2023: Chatgpt for teaching and learning: An experience from data science education. *Proceedings of the 24th Annual Conference on Information Technology Education*, Association for Computing Machinery, New York, NY, USA, 66–72, URL `https://doi.org/10.1145/3585059.3611431`.

Zhu, H., S. P. Dow, R. E. Kraut, and A. Kittur, 2014: Reviewing versus doing: Learning and performance in crowd assessment. *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, Association for Computing Machinery, New York, NY, USA, 1445–1455, URL `https://doi.org/10.1145/2531602.2531718`.

Zlabinger, M., M. Sabou, S. Hofstätter, M. Sertkan, and A. Hanbury, 2020: DEXA: Supporting non-expert annotators with dynamic examples from experts. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Association for Computing Machinery, New York, NY, USA, 2109–2112, URL `https://doi.org/10.1145/3397271.3401334`.

# C. List of Abbreviations

**AI** Artificial Intelligence

**AR** Augmented Reality

**CA** Conversational Agent

**DIP** Digital Involvement Project(s)

**DSR** Design Science Research

**HMD** Head-Mounted Display

**ICT** Information and Communication Technology

**IS** Information Systems

**LOD** Level of Detail

**ML** Machine Learning

**SD** Standard Deviation

**SMQ** Science Motivation Questionnaire

**UEQ** User Experience Questionnaire

**VR** Virtual Reality

**XR** X-Reality

# D. List of Figures

# E. List of Tables