

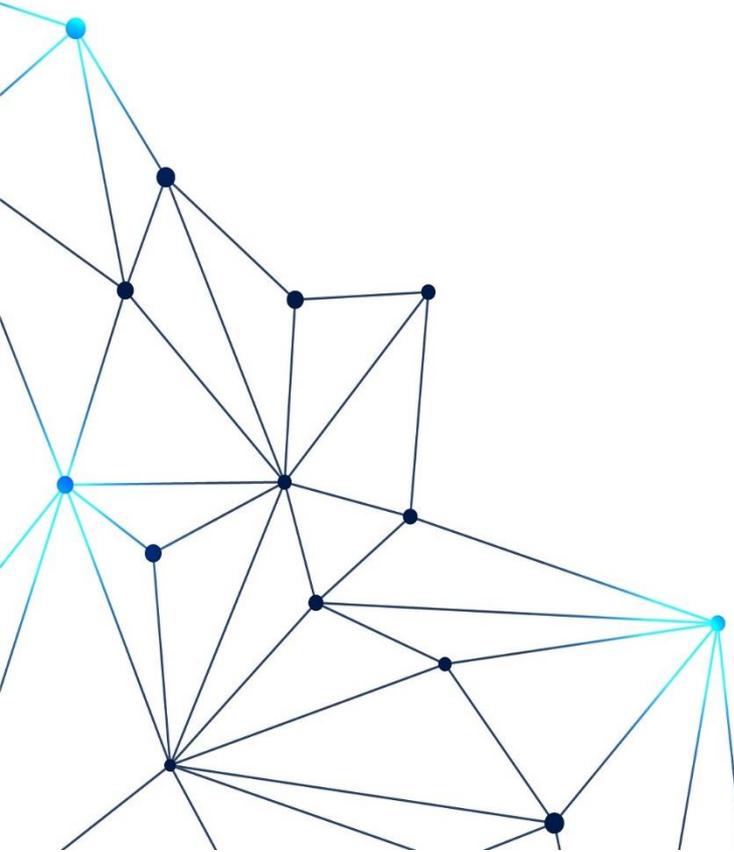
Master Thesis

MSc in Business Administration and Information Systems (e-Business)

Copenhagen Business School

The Conundrum of Open Government Data Utilization

A Danish commercial user-perspective



Student: Pernille Kvamsdal

Student number: 106970

Supervisor: Rony Medaglia

Hand-in Date: 15.11.2018

Characters: 133.650

Pages: 59

Acknowledgments

I would like to express my gratitude to the following people:

To the case organization who opened their doors to me and the individuals who kindly let me interview them. I am deeply thankful for the opportunity.

To my supervisor, Rony Medaglia, who always has the right advice at the right time. Thank you for the role you have played as a teacher and supervisor throughout my master's degree. I am grateful for everything you have taught me.

To my friend Nikolina, who helped me brainstorm when I was stuck and always were available for a chat about the process. I am grateful for your help in getting over the hurdles faced during the study and the thesis process.

To my friend Maria, who helped edit and proof this work. I am grateful for your thoughtful insights and sharp eye which helped elevate my writing and get over the finish line.

Finally, a heartfelt thank you to my boyfriend Eric and my dear family for their support and encouragement throughout my studies.

Abstract

The expected impacts of open government data (OGD) are not realized, and the missing user-perspective is a bottleneck in understanding why. This research takes a step in bridging the gap by investigating a Danish commercial user-perspective of how the facilitating conditions influence the perceived use-value of OGD. The study takes an abductive approach to theory development and adopts a single case study of a Danish organization.

Interviews with employees and observational accounts were used to explore the user-perspective. The theoretical lens of framing was applied to analyze data. This study identified four archetypical organizational users, which differed in relation to how they work with OGD: business developers, data visualizers, software developers, and data scientists. The user archetypes presented a diversity in how they framed the influence of the facilitating conditions - data-openness and data-quality - on the perceived use-value of OGD. This study contributes to current research within this field by providing empirical insights into the missing user-perspective, and further contributes to practice by outlining a set of recommendations for public managers who design OGD initiatives.

Keywords

Open Government Data, Private-sector, Data Utilization, Use-value, User-perspective, Framing, Denmark

Content

- 1 Introduction 1
- 2 Background 4
 - 2.1 Facilitating Use of OGD..... 4
 - 2.2 Who are the OGD Users? 5
 - 2.3 The Value-perspective..... 6
- 3 Theoretical Lens 8
- 4 Methodology 10
 - 4.1 Research Philosophy 10
 - 4.2 Research Strategy and Setting 11
 - 4.3 Data Collection..... 13
 - 4.4 Data Analysis 16
 - 4.5 Reliability and Validity 17
 - 4.6 Ethical Concerns..... 18
- 5 Findings 20
 - 5.1 User Archetypes 20
 - 5.2 Use-conditions..... 21
 - 5.3 The Use-conditions’ Influence on the Perceived Use-value..... 23
 - 5.4 Literature Corroboration..... 36
- 6 Discussion 43
 - 6.1 Summary of Findings 43

6.2	Contribution to Research.....	44
6.3	Contributions to Practice.....	47
6.4	Limitations.....	50
7	Conclusion.....	52
8	References.....	54
9	Appendices.....	i
9.1	Appendix I: Interview Guide.....	i
9.2	Appendix II: Data categorization.....	ii

1 Introduction

Open government data (OGD) is increasingly published across the globe. “Government data” is produced or commissioned by government or government-controlled entities¹ and collected for the use in the public-sectors administrative processes (T. Jetzek, Avital, & Bjorn-Andersen, 2014). To be “open” entails data being free to use, modify, and share by anyone for any purpose (European Commission, 2015; Open Knowledge Foundation, 2018). The popularity of the term OGD followed the publication of the Open Government Directive of the Obama administration in 2009 (Attard, Orlandi, Scerri, & Auer, 2015). In Europe, initiatives related to a similar term, public-sector information, can be traced back to the 1980’s (K. Janssen, 2011). With almost a decade of scholarly efforts, substantial contributions have been made to OGD as a research field.

In Denmark, the debate of generating value from OGD has been high on the public agenda with national OGD initiatives such as the basic data program, and municipal efforts. A dialogue event called “From Data to Value” was held in May 2018 in collaboration between the Open Data DK association, and the National Union of Municipalities². The event brought together public and private stakeholders for knowledge-sharing and debate about how the public-sector could facilitate commercial organizations value generation from OGD. Significant funds are allocated to opening datasets and to make it easier for commercial organizations to utilize OGD. The goal of the public-sector representatives seemed to be collecting input for the prioritization of the funds from the private-sector representatives³.

The primary motivation for publishing OGD is the potential impacts achieved through its use (European Commission, 2015; Safarov, Meijer, & Grimmelikhuisen, 2017). Existing research discuss the potential impacts as values, benefits and effects (Cranefield, Robertson, & Oliver, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a). However, the expected impacts span widely and include transparency, accountability, civic engagement, economic value, social value and service innovation (Kassen, 2013; Worthy, 2015). The

¹ <https://opengovernmentdata.org/>

² <http://www.opendata.dk/om/kalender/fra-data-til-vaerdi>

³ Observation from event “Fra Data Til Værdi”, Authors elaboration of observational notes.

impacts of OGD are mostly assumed (Herala, Vanhala, & Krri, 2016) and impacts have been found to be “more complex, more unpredictable, and more political than advocates claim” (Worthy, 2015, p. 788).

A “chicken and the egg” paradox is existing in the OGD field. To generate value from open data, investments for making open data more useful are necessary. However, the public-sector stakeholders “are not willing to expend additional capital on open data unless they are able to perceive evidence of value” (T. H. Jetzek, 2015, p. 10). Furthermore, the public-sectors motivation for publishing OGD decreases when they cannot see how their work benefits the users (SEISMONAUT, 2017).

OGD is described as “a shared resource that offers value beyond what is captured from the original intended use” (T. Jetzek, Avital, & Bjorn-Andersen, 2014, p. 102). OGD does not have any value on its own (Attard, Orlandi, & Auer, 2016a; Müller & Jensen, 2017) and the expected impacts can only be realized through the use of the released data by users such as citizens, researchers or businesses (Magalhaes & Roseira, 2017; Zeti, Bartalis, Dobre, & Pitu, 2017). A requisite for enabling users of OGD entails understanding under which conditions the users would adopt OGD (Anneke Zuiderwijk, Janssen, & Dwivedi, 2015).

Despite the existing calls for adopting a user-perspective to bridge the gap between the expected and actual impacts of OGD (M. Janssen, Charalabidis, & Zuiderwijk, 2012; T. Jetzek, Avital, & Bjorn-Andersen, 2014), research focusing on the users is scarce. Little is still know about who the users are and which use-conditions are necessary for facilitating the use of OGD (Safarov et al., 2017). To address the gap, the research adopts a user-perspective to investigate how organizational use of OGD can be facilitated. The study examines what facilitating conditions influence the perceived use-value of OGD. The present study will address the gap through the following research questions:

How does the use-conditions for open government data influence the perceived use-value for different organizational users?

- Who are the organizational users of OGD?
- What use-conditions are relevant to the organizational users?
- How does the relevant use-conditions influence the perceived use-value of OGD for the users?

This research aims at answering this by investigating the case of a Danish commercial organization which uses OGD as a core resource. It will examine who the organizational users are and how the users perceive use-conditions to influence the use-value of OGD. This context is a useful starting point for investigating OGD users and how to facilitate their use of OGD, because the setting ensures a lower degree of influence

from external complexities such as variety in goals, differences in objectives, variations in the type of data utilized as well as cultural or institutional variations.

In practice, having increased sensitivity to how certain conditions influence the users can facilitate the design of OGD initiatives and prioritizations of aspects which support the reuse of OGD to stimulating the OGD market. Scholarly research can benefit from more clarity about the private-sector perspective, and the boundaries between public and private stakeholders (T. Jetzek, Avital, & Bjørn-Andersen, 2014a). Focusing on the organizational users is a first step towards identifying appropriate measures of the benefits related to the direct users.

The remainder of this paper is structured as follows; Below section presents the background and motivates the study. Section 3 presents the theoretical lens applied in this study. Section 4 then outlines the research philosophy and methodological choices of the study. Section 5 describes the results from the case study, and the corroborating evidence found in existing literature. Section 6 provides a discussion of the results, outlines the significance of the findings for research, and for practice. Finally, section 7 concludes this study and offers suggestions for future research.

2 Background

The background includes three sections that describe existing research and motivates the present study. The next section presents existing research about facilitating the use of OGD. Then part 2.2 presents the of the current user-perspectives. Lastly, part 2.3 discusses the concept of value, the value-perspectives in existing literature and motivates the value-perspective used in this research.

2.1 Facilitating Use of OGD

Facilitating use of OGD is central to the ability to realize any benefits or values. However, the link between the conditions to use (use-conditions) and the value is hardly understood. Two groups of papers were found to address this link.

One group of papers draw a unidirectional link between use-conditions, utilization, and value-outcomes. The group includes research on the assessments and evaluation of OGD initiatives (Chu & Chiang, 2014; Purwanto, Janssen, & Zuiderwijk-van Eijk, 2017), infrastructures (Alexopoulos, Loukis, & Charalabidis, 2016) ecosystems (Abella, Ortiz-De-Urbina-Criado, & De-Pablos-Heredero, 2017; F. Welle Donker & van Loenen, 2017), or the data value chain (Attard, Orlandi, & Auer, 2016b). Furthermore, some links are drawn between use-conditions and utilization in research on the adoption and use of OGD (Saxena & Janssen, 2017; Anneke Zuiderwijk et al., 2015), and by investigating critical factors for publication and use (A. Zuiderwijk, Susha, Charalabidis, Parycek, & Janssen, 2015). Lastly, some research includes these links by focusing on the process of value creation (Attard, Orlandi, & Auer, 2017; T. Jetzek, 2016). Only one paper includes the user as a central point of evaluation (F. Welle Donker & van Loenen, 2017).

The second group of papers move beyond the unidirectional link between use-conditions and value-outcomes. Some papers investigate the perceived value and identify what value is created for various stakeholders (Cranefield et al., 2014; Lakomaa & Kallberg, 2013). Other research unfolds the value creation mechanisms which are transforming use-conditions into value (T. Jetzek, Avital, & Bjørn-Andersen, 2013, 2014b), one study empirically investigates a value creation mechanism within a private organization (T. Jetzek, Avital, & Bjorn-Andersen, 2014).

The first group consists mainly of conceptual or quantitative research and does not include an in-depth exploration of use-conditions, utilization or values. Hence, causal assumptions are made between use-conditions and value outcomes, while little understanding exists as to why the causalities are significant. The second group of papers does have some more elaborate ideas of how use-conditions influence the

value creation from data. However, the articles primarily focus on the value outcome which is a result of use, rather than focusing on the value data in itself holds in its use. As such, the research focuses mainly on the links to some value-outcome.

Overall, the two groups give the impression that the causal links between use-conditions and the impacts have been pre-emptively quantified without considerations to the value data in itself holds in its use. Since the two groups contain the central assumption that value can be achieved by utilizing data, it is curious that use-value is not given much consideration. Hence, the present research's focus on the use-situation is beneficial, as it can provide insights into the value that data has when it is used.

2.2 Who are the OGD Users?

Research within OGD has increasingly included the user-perspective. For example, as part of assessment frameworks (F. M. Welle Donker, 2016), to understand organizational capabilities (Eckartz, van den Broek, & Ooms, 2016), in exploring OGD business models (Magalhaes & Roseira, 2017), or in investigating OGD value creation mechanisms (T. Jetzek, Avital, & Bjorn-Andersen, 2014). Nevertheless, little knowledge exist about who the OGD users are, and most research makes estimates about users (Safarov et al., 2017).

In general, users are those who utilize OGD to create value beyond the original intended use (T. Jetzek, Avital, & Bjorn-Andersen, 2014). A recent literature review divides OGD users into two categories; direct users and indirect users (Safarov et al., 2017). *Direct users* make use of data for themselves, for example by creating applications or enriched datasets. This group includes individual citizens, developers, businesses, researchers, NGOs and journalists (Safarov et al., 2017). *Indirect users* utilize data processed by intermediaries and can be considered the beneficiaries of the applications or enriched datasets created by the direct users. This group includes citizens (the crowd), NGOs, and businesses (Safarov et al., 2017).

Direct and indirect users are only one way of distinguishing between users in existing literature. Other classifications are derived by synthesizing literature (Gonzalez-Zapata & Heeks, 2015), or by investigating pre-acknowledged stakeholder groups such as entrepreneurs (Lakomaa & Kallberg, 2013). Only one identified paper aimed at classifying different users (Smith & Sandberg, 2018). This paper divides the user-types of an OGD platform holistically based on a mix of “motivations, objectives, pre-conditions and approaches as well as in experiences of innovation barriers” (Smith & Sandberg, 2018, p. 257). The users are sorted into three groups; *employees*, who use data as part of their employment, *entrepreneurs*, who use data for pursuing innovations and *hobbyists*, who use data for hobby projects and fun (Smith & Sandberg, 2018).

The user-groups are distinguishable in multiple ways; if considering *unit of analysis*, users are found both on the individual level of analysis (Lakomaa & Kallberg, 2013; Saxena & Janssen, 2017) and organizational level of analysis (T. Jetzek, Avital, & Bjorn-Andersen, 2014; F. M. Welle Donker, 2016). If considering *legal nature*, they range from civil users, to commercial organizations, non-governmental organizations, and even researchers (Abella et al., 2017). If considering their *network role*, it includes users such as data producer, data publisher, data facilitator, data broker, or data consumer (Attard et al., 2017). Overall, many ways of differentiating users are present and all have certain appropriate applications.

Likewise, the existing research investigates a variety of users. Some articles use a broad range of potential users including students (Alexopoulos et al., 2016; Saxena & Janssen, 2017) or professional users (Anneke Zuiderwijk, Janssen, & Sussha, 2016), researchers, citizens, and civil servants (Anneke Zuiderwijk et al., 2015). A few articles focus on a particular group in isolation, such as entrepreneurs (Lakomaa & Kallberg, 2013; Sussha, Grönlund, & Janssen, 2015a), or private organizations (T. Jetzek, Avital, & Bjorn-Andersen, 2014; Magalhaes & Roseira, 2017). Overall, these insights indicate that broad user groups are applied when the users are investigated in existing research.

Understanding how to facilitate use depends on the use of OGD, *by the users*. Thus, the many ways of distinguishing users and the application of broadly defined user groups can lead to a great deal of complexity and confusion. First, it becomes difficult to compare the findings across literature, as key assumptions behind classifications are never made explicit. Second, broad categories assume uniformity across user groups, without considering that the contextual reality of the users might influence how the utilization should be facilitated. Overall, the broad classifications make it difficult to grasp where boundaries of different stakeholders lie.

Previous research have called for acknowledging the diversity of open data perspectives and focus research on one area (Anneke Zuiderwijk et al., 2015). This research focuses on the perspective of a direct user of OGD, more specifically the users within a commercial organization. Zooming in on the commercial users is a beneficial starting point for understanding how to facilitate the use of OGD as it limits the diversity of perspectives by only investigating one contextual situation.

2.3 The Value-perspective

The literature on the concept ‘value’ is subject to problems of definition as value refers to different phenomena (Bowman & Ambrosini, 2000). The concept of *use-value* and *exchange value* has been applied to clarify the different phenomena (Bowman & Ambrosini, 2000). *Use-value* is defined as “the specific quality of a new job, task, product, or service as perceived by users in relation to their needs” (Lepak,

Smith, & Taylor, 2007, p. 181). *Exchange value* is defined as “the monetary amount realized at a single point in time when the exchange of the good takes place” (Bowman & Ambrosini, 2000, p. 3). Exchange value is considered to have no meaning to individuals beyond the result of work (Repo, 1986). Public-sector initiatives for opening data imply that data should be disseminated free of charge⁴. Hence, opening government data can be thought of as removing the economic exchange value of data. While exchange- and use-value were defined initially on an organizational level, they are applicable at multiple levels of analyses including, societal, organizational and individual (Lepak et al., 2007).

The term ‘value’ comes in multiple variations in OGD literature. It includes economic and social value (Eckartz et al., 2016), public value (Pereira, Macadar, Luciano, & Testa, 2017), transparency and accountability (Viscusi, Spahiu, Maurino, & Batini, 2014; Worthy, 2015), participation and innovation (Anneke Zuiderwijk et al., 2015) and sustainable value which has an environmental dimension as well (T. H. Jetzek, 2015). Most of the existing values in OGD literature can be considered societal level values. It implies that research has focused mainly on a societal level of analysis as opposed to the organizational or individual level of analysis. This observation is in line with the previously described lack of focusing on the user-perspective as research and practice are driven mainly to serve high-level assessments. Instead, perceived use-value for the direct users seem to relate to the organizational or individual level of analysis.

Creating value relies on the relative value that is realized by the user, who is the focus of value creation (Lepak et al., 2007). Hence, a central aspect for value creation is the extent to which the perceived use-value for the target users are realized. The perceived use-value relies on perceptions of the usefulness of the product by the user (Bowman & Ambrosini, 2000). OGD use is defined as “the activity that a person or an organization conducts to view, understand, analyze, visualize or in other ways use a dataset that has been provided to the public by a governmental organization” (Anneke Zuiderwijk et al., 2015, p. 429). Following this definition, the present study defines the use-value of OGD as the usefulness perceived by users, as they view, understand, analyze, visualize, or in other ways use OGD.

The above sections discussed the gaps in existing research motivating this study. It included insights into existing research of facilitating the use of OGD, the user-perspectives applied, and the concept of value. The following section presents the theoretical lens applied to investigate the gap.

⁴ <https://opengovernmentdata.org/>

3 Theoretical Lens

This section presents the theoretical lens of framing applied in the present research. This study adopts a similar perspective of framing as applied by Sun & Medaglia, where framing is used as a lens to investigate the perceived challenges of public sector AI adoption by different stakeholder groups (2018). Frames are used by individuals to organize what they see in everyday life (Borah, 2011). Framing is considered a general blueprint which aims at investigating how groups construct and negotiate meaning around a general phenomenon (Sun & Medaglia, 2018, p. 3).

The concept of technological frames has been applied in research on Information Technologies. Framing is defined as “a collectively constructed set of assumptions, knowledge and expectations regarding a technology and its uses and applications in organizations” (Cornelissen & Werner, 2014, p. 185). Framing is applied in the research on adoption and implementation of technology. Within this line of research, technological frames have been used to investigate the consequences of frames on effective adoption and use of technology (Davidson, 2006), and how utilization was framed differently within two occupational groups (Mazmanian, 2013).

Studies applying *framing* mostly develop unique frames (Borah, 2011). This aspect is a weakness of framing as it increases the difficulty of generalizing from studies utilizing this approach (Davidson & Pai, 2004). Furthermore little is known about how frames are constructed (Borah, 2011), and the notion of changes occurring in frames over time is considered a methodological challenge (Orlikowski & Gash, 1994). Hence, the context-bound nature of frames is criticized for not enabling generalization, as it is difficult to scale for multiple cases or survey research (Davidson & Pai, 2004). However, the premise of framing is intimately connected with the qualitative application, as it relies on the shared perceptions of a group. Hence, to establish frames, one must engage with the subjective experience of the individuals. Framing has the following strength: it enables the discovery of insights into the relational information around scenarios that “link situations with actions and outcomes, or causes with effects.” (Cornelissen & Werner, 2014, p. 192). Framing is hence suited to establish links between cause and effect qualitatively. Framing does not provide any constructs or propositions which can be applied to create categories of cause and effect. As identified in this research, insights from existing literature led to apply use-conditions as causal elements, and emerging findings led to discovering the use-value as an effect from use-conditions. Thus, these constructs became central to the cause-effect links enabled by the lens of framing.

Research on frames includes a focus on congruence or incongruence in the frames (Davidson & Pai, 2004; Orlikowski & Gash, 1994). Thus, framing assumes the existence of multiple realities and hence, provides

a lens that can enable discovering subtle differences between perspectives. In this study, framing then serves as a sensitizing device (Klein & Myers, 1999) for understanding how organizational users frames the link between conditions and the use-value of OGD.

To the best of the author's knowledge, no previous research within OGD has applied the concept of framing as a lens to investigate differences in user perceptions. Still, framing is considered beneficial to this research's empirical investigation of organizational users, as it affords the opportunity to differentiate between different users' framing of the links between use-conditions and perceived use-value. The approach is suiting for the situation faced in this study, where the quantitative links established between use-conditions (cause) and the perceived use-value (effects) of OGD lack qualification. Hence, framing can offer new insights into how use-conditions influence the perceived use-value and enable the discovery of in-depth knowledge which existing research is missing.

4 Methodology

The methodology section discusses the methodological choices taken to answer the research question: “*How does the use-conditions for open government data influence the perceived use-value for different organizational users?*”. First, the research philosophy is presented. Secondly, section 4.2 outlines the research strategy and the research setting of this study. Section 4.3 presents the data collection, and section 4.4 describes the data analyses approach. Section 4.5 describes the measures taken to ensure the reliability and validity of the findings. Finally, section 4.6 outlines the ethical considerations of the present research.

4.1 Research Philosophy

This research explores the organizational use of OGD to understand how use-conditions influence organizational users. The focus is on the understanding of a pertinent challenge faced in practice, namely that public-sector managers know little about how to facilitate the use of OGD. Adopting a user-focus assumes the world consists of multiple realities. The closeness to practical challenges as well as the assumption of multiple realities makes it appropriate to adopt a pragmatic philosophy as the foundation of this study (Saunders, Lewis, & Thornhill, 2016). Pragmatic epistemology considers that reality cannot be impartially observed since observers are guided by their interests, expectation, and theory⁵. In pragmatism, reality is a consequence of ideas and propositions are valid if they enable action (Saunders et al., 2016).

Upon starting this research, it was considered essential to understand the subjective reality of the users. Furthermore, a great deal of existing research within OGD explores the links between conditions, utilization, and value-outcomes. The immediate connection between the users as the desired subject of investigation and the existing research was not apparent. Nevertheless, existing research might still contain evidence relevant to solving the practical problem of facilitating use. Why it was considered necessary not to disregard previous research. Hence, an abductive approach to theory development was selected. Abduction is a mode of inquiry facilitating the production of novel generalizations which are embedded in the experienced reality (Tavory & Timmermans, 2014). The abductive approach is grounded in pragmatism and entails iteratively analyzing data, contrasting it to theory, and eventually arrive at a theoretical generality (Saunders et al., 2016). Abduction involves generalizing from interactions between the specific

⁵ <https://www.iep.utm.edu/pragmati/>

and the general (Saunders et al., 2016). Abduction is value bound, which implies the researcher's knowledge of the research field and the observed phenomenon is a critical aspect in observing the "surprising facts," while general knowledge of theoretical directions enable the development of theoretical generalizations (Saunders et al., 2016). Hence, using an abductive approach risks excluding insights or theoretical directions simply because the researcher does not realize their significance or have little knowledge of some theories. However, abduction is beneficial when the research field is not new, and there already exist some research, but where it is difficult to produce a framework for deductive inquiry (Saunders et al., 2016). Abductive inquiry facilitates a better connection between the knowledge of existing research in the field and the subjective insights from the respondents. Thus, an abductive approach does not strive to stay detached from previous research efforts which an inductive approach would be prone to. Nor does the abductive approach attempt to deductively test a hypothesis of relationships on a small sample when a quantitative approach could be more suitable to test general applicability of relationships. Instead, abduction attempts to encompass the strengths of both. Thus, abduction is beneficial as it can enhance both the generality and subsequent applicability of the findings when conducting a single case study.

4.2 Research Strategy and Setting

A case study is beneficial when investigating 'how' and 'why' questions, and to examine a phenomenon "in depth within its real-world context" (Yin, 2018, p. 15). The case study methodology is appropriate for this study as it focuses on the facilitation of use by looking at the context of data use. Case studies are criticized for not being rigorous enough and for not facilitating generalizations (Yin, 2018). The rigor concern arises when the researcher does not follow a systematic procedure. The generalization concern occurs when it is doubtful if case study findings can be quantified and applied in other case-contexts.

In order to address the previous concerns of rigor and generalization, the present research adopts the Structured-Pragmatic-Situational (SPS) approach to case studies (Pan & Tan, 2011). The method is an explorative case study which aligns well with the abductive theory development. The SPS approach finds it impossible to have a clean theoretical slate and considers that existing research is instead value-adding (Pan & Tan, 2011). Furthermore, the SPS approach advocates iteration between observations, analysis, and theory with the aim of conducting an exploratory case study contributing to generating new theory (Pan & Tan, 2011).

The SPS approach consists of eight actionable steps which facilitate rigor in the execution of the case study. The steps are divided into two cycles; the *framing cycle* and the *augmenting cycle*. Each of the cycles is meant to be iterated through as the theoretical lens evolves. The *framing cycle* starts with the

conceptualization of the phenomenon and then moves into initial data collection and analyses by selecting a reference theory and extending this lens with the collected data. Once confidence in the emerging model is reached, the augmenting cycle is entered. The *augmenting cycle* starts with confirmation and validation of the data in the emergent model, then fills the developing model with at least two pieces of evidence and continually tests for Theory-Data-Model alignment.

There is a lack of cases in OGD research focusing on businesses with a sustainable income (Cranefield et al., 2014). The lack indicates that the sustainable OGD business is a setting which is difficult to explore. Following the “planned opportunism approach” (Pan & Tan, 2011), the author gained access to a sustainable Danish OGD organization. It was decided to investigate this organization as a single case study, as this is beneficial when researching a unique case that few have examined before (Saunders et al., 2016). The OGD field has mainly been investigated from a public-sector perspective. The present study examines the private-sector perspective which provides an opportunity to determine if there is an alternative explanation to the problem of facilitating use, compared to the existing research field (Yin, 2018). Finally, as opposed to most potential cases in the Danish context, the case organization had enough employees to make it possible to have an empirically strong case with a variety of voices (Pan & Tan, 2011).

4.2.1 Research Setting

The research setting for this study was the case of a Danish organization founded in 2013 when the first open government data in Denmark was released. The case organization specializes in geodata solutions. The organization utilizes OGD such as geographical data and the business registry. The case organization aims to put data into play, which is achieved in two ways. First, by utilizing OGD as inspiration for or as a central part of their solutions. The case organization manipulates, analyzes, enriches and visualizes the OGD sources. Second, by facilitating the utilization of OGD, they enable the reuse of data for their customers and their respective end-users through either software solutions, plug-ins or courses in tools that can work with the data. The customer base is mainly public or private organizations. Private organizations include a broad array of industries such as construction, shipping, finance, and retail. Public customers include both municipalities, government agencies, and other governmental organizations. The largest part of their customer base is currently the public segment. The organization employs eighteen people and is self-sustaining. Thus, they do not depend on any subsidies to keep afloat, as their business activities can sustain their costs.

4.3 Data Collection

This section consists of three parts. First, the data collection done during the process of conceptualizing the phenomenon is outlined. The next section describes the data collection within the case organization. Lastly, the data collection which served as confirmation and validation of the interviews within the case organization is described.

4.3.1 Conceptualizing the Phenomenon

The conceptualization of the phenomenon started before having gained access to the case organization. It was not possible to determine a pertinent research question before knowing more about what was going on in the field (Yin, 2018, p. 4). First, a broad inquiry of non-technical literature, policy documents, and existing research was conducted on open data in Denmark. These sources provided insights into the context and issues pertinent to the Danish field. Furthermore, informant interviews with a public-sector and a private-sector representative were conducted, focusing on exploring key challenges, ongoing processes, recent events and what they wanted to know. Lastly, the researcher collected observations from a dialogue event called “From Data to Value”⁶ between public and private stakeholders.

Access to the case organization was granted end of May 2018. Following this, background information on the company was collected by reviewing their web-page, and a range of news articles related to the case organization. The background information collected enabled a better understanding of the case organization’s context. The contextual understanding facilitated the discovery of a plausible research question within the case organization, which addressed the related concern of facilitating the use of OGD in practice.

Overall the sources presented above, overall enabled a higher sensitivity to the subsequent data collection within the case organization and served as part of the basis for formulating questions for the initial interviews (Pan & Tan, 2011). As appropriate, these sources were used to validate and ensure contextual understanding of the case-study findings.

⁶ <http://www.opendata.dk/om/kalender/fra-data-til-vaerdi> - 24th of May 2018.

4.3.2 Data Collection within the Case Organization

Data collection within the case organization spanned from July to October 2018. The data collection consisted both of observational data and interviews with the employees. The researcher iterated between data collection and analysis throughout the data collection phase (Pan & Tan, 2011).

The researcher was present on a weekly basis throughout most of the data collection phase. The presence of the researcher served two purposes. First, it established a level of repour with the case organization's employees. Second, observations from six meetings provided insights into the distribution of roles among the employees and a better understanding of the tasks and projects which employees reflected upon during the interviews. During the meetings, notes were taken by hand and subsequently written and expanded on in a digital observational note.

A total of twelve interviews were conducted during the research period, which is a total of two-thirds of the case organization's employees. Since the present study focusses on the Danish context, it was prioritized to avoid possible biases arising from institutional differences between different nationalities and countries. Hence, only Danish full-time employees were interviewed, which resulted the exclusion of potential respondents from the interview pool. As the refinement of the user archetypes appeared, each group was represented by at least two interviews.

The initial interviews were of a more explorative nature which provided an overview of the organizational processes while also informing the basis of the analytical categories. The questions posed in these interviews were created on the basis of the gathered contextual knowledge. Later interviews were conducted as semi-structured interviews which followed an interview guide (see appendix 9.1). The semi-structured interview format provides better comparability among interview-accounts. The probing of semi-structured interviews relies on attention and sensitivity to the emergent meaning and situational cues within the interviews (Kvale & Brinkman, 2009). The interview format allowed for probing into relationships linking to earlier interviews, as well as new insights provided by the interviewees. The semi-structured interview was an appropriate method for the later interviews since they served to further refine the theoretical model and propositions established through the previous interview accounts. The accuracy of how the author understood the interview accounts was ensured by asking for clarification of the interpretation throughout the interviews. All interviews took place face-to-face in a meeting room at the case organizations location. That way, the respondents experienced a familiar and safe environment with minimal disruption of their workday. The interviews lasted between 50 minutes and 1 hour 20 minutes.

Upon initiating the interviews, the respondents received information about the purpose of the interview and how their data would be treated, including that they would be anonymized in the report. Since the thesis would be written in English, the respondents was given the opportunity to adjust and comment on the translations of the analytical quotes used in this research paper.

Two follow-up interviews of approximately 30 minutes were held with the earliest interviewees. The first interviews were broader in scope, so follow-up interviews were used to complement and provide more depth to these interviews. Furthermore, the follow-up interviews served to receive some clarification about emerging themes. Follow-up questions were sent by e-mail to five respondents after entering the augmenting cycle. The follow-up was used for clarification of aspects related to the different dimensions of the use-conditions from the interviews and to see if there was some evidence that could further substantiate the findings.

Table 1 shows the code of the respondents, their user archetype, the duration of the interview and whether they received follow-up questions.

Code	User Archetype	Interview duration	Follow-up
R1	Business Developer	51 minutes	Interview
R2	Data Visualizer	57 minutes	Interview
R3	Software Developer	49 minutes	E-mail
R4	Software Developer	48 minutes	No
R5	Business Developer	1 h 19 minutes	E-mail
R6	Data Scientist	1 h 01 minute	E-mail
R7	Data Scientist	54 minutes	No
R8	Data Scientist	1 hour 2 minutes	E-mail
R9	Data Visualizer	53 minutes	No
R10	Business Developer	55 minutes	No
R11	Software Developer	1 hour 3 minutes	No
R12	Data Scientist	1 hour 1 minute	E-mail

Table 1: Interviewees

4.3.3 Data Confirmation and Validation

The SPS approach advocates the collection of additional data to ensure an objective account if conflicting interpretations are present (Pan & Tan, 2011). As this research is a single-case study of the organizational

users, the researcher chose to conduct open-ended interviews with three public-sector representatives within the OGD collection and publication. The interviews with the public-sector representatives were conducted to avoid false understandings of the publication process, arising from e.g., frustration or preferences for certain ways of publishing data. So, the interviews with the public-sector representatives served as verification of the organizational user's accounts. The public-sector representatives were identified by the author's specification of the relevant points of the publication process and the case organization facilitated the relevant contacts.

4.4 Data Analysis

The data analysis section describes the iterative analytical approach taken in this study. First, a description of the analytical tools is provided. Second, an explanation of the construction of the theoretical lens and how it was extended.

4.4.1 Analytical Tools

The interviews were conducted in Danish to ensure the richest possible answers and then transcribed. Follow-up information collected was added to the transcription file. The empirical material was analyzed through a well-structured inductive approach which applied memo writing (Corbin & Strauss, 2015). Memo writing was mainly used early on to identify theoretical themes and directions. It was beneficial for articulating the analytical ideas and directions which arose from the early interviews and facilitated in identifying relevant directions for selecting a theoretical reference theory. Transcription files were imported to NVivo Version 11 and coded through a process of open coding. When the analytical reference theory was selected data were coded in accordance with the theory (Pan & Tan, 2011).

4.4.2 Constructing and Extending the Theoretical Lens

The analysis of the data started from the collection of the first interviews. A broad array of theories was read to identify a candidate theory and the data was tested against multiple theories throughout the analytical process.

The SPS approach entails selecting a reference theory to guide the theoretical generalizations. The present research's reference theory developed throughout the duration of the study. The first theoretical reference theory applied, facilitated the identification of use-conditions and the actual use situation faced by the respondents. As the data collection proceeded, it was clear that respondents discussed the value of OGD in relation to projects they were working on. The way value was discussed led to discovering the concept

of use-value. The interview data was then reorganized based on this concept, leading to new insights into the perceived use-value as well as the use-conditions.

Further interviews and observations led to the insight that there were four different user types describing the use-conditions and use-value differently. This observation made the emerging theoretical explanation seem “detached from reality” (Pan & Tan, 2011), as the theoretical explanation could not account for the observed differences. The differences in perspectives led to applying the theoretical lens *framing*.

The accumulated data was then systematically explored for perceptual variation with the theoretical lens *framing* applied as a sensitizing device (Klein & Myers, 1999). Framing could account for the apparent differences between the user groups, as well as link the use-conditions and use-value. It enabled solidifying the user archetypes and considered the way user archetypes framed the influence of use-conditions on the use-value to establish a link between these. An example of interview coding as sorted by the use-conditions can be found in Appendix 9.2.

Insights from direct observations within the case organization, interviews with public-sector representatives and an open data dialogue event complemented the analysis. The observations were collected as descriptive notes and served as verification of the interview accounts. Furthermore, the observations provided contextual background which facilitated the understanding of the interview accounts’ relevance.

The augmenting cycle was entered when the point of theoretical confidence was reached, that is when only incremental learning took place, and the theoretical propositions were considered to explain the empirical reality (Pan & Tan, 2011).

The use-conditions were triangulated by existing literature which served to identify corroborating or contrasting evidence between literature and interview-accounts. All user archetypes were represented within the case organization by at least three individuals. Selective coding was applied, which entails filling the occurring table with at least two sources of data for every category (Pan & Tan, 2011). That meant every impact identified within a user archetype had to be supported by at least two individuals, or an individual and an observation note.

4.5 Reliability and Validity

As this study is conducted by a single researcher, internal reliability can be an issue (Saunders et al., 2016). Reliability has instead been sought through application of the SPS approach. Furthermore, by keeping a

reflective research diary, possible influences of the researchers own values have become apparent throughout the research process. The reflective diary was reviewed throughout the research process to identify changes in directions and to reflect on the overall research design. Lastly, refining theoretical categories through memo writing has provided some early clarity and stability in the constructs and dimensions used for analyzing data (Saunders et al., 2016).

The twelve interviews supported by observations were used to provide a variety of voices (Pan & Tan, 2011). Dominating voices were avoided by ensuring the triangulation of all propositions included in this study by two to three independent sources of data (Saunders et al., 2016). In some cases, emerging propositions were excluded, as it was not possible or difficult to support it with collected data or observations.

The SPS approach advocates utilizing Theory-Data-Model alignment throughout the augmenting cycles. Theory-data alignment was applied by checking for alternative explanations for the data in existing theory and ensuring that e.g. the dimensions of the use-conditions was the best explanation of the empirical data collected. Different variations of the dimensions were tried against the data to ensure theory-data alignment. Data-model alignment was conducted by doing a test of inclusion and exclusion (Pan & Tan, 2011) to ensure data from interviews corroborated the constructs of the model. Empirical evidence was checked for its influence on the model. The inclusion of data was tested to see if the constructs could explain it or created any changes in the model. Excluded data was checked to see if it contradicted the fundamental argument. When overlaps among the frames were found, they were discussed as “shared frames.”. Theory-Model alignment was applied by triangulating existing literature with the use-conditions identified. The corroboration with existing literature made it possible to claim logical replication of the findings as the model resonates with previous theory (Pan & Tan, 2011).

4.6 Ethical Concerns

A central ethical concern for the case organization in this research was how business sensitive information was treated. Upon agreement with the contact person, a non-disclosure agreement was signed to provide the case organization the confidence that business-critical information would not be shared. Furthermore, per the agreement, the raw transcription files from interviews would not be available, as these would include information about customers and ongoing business developments which were to be kept private. Finally, no business sensitive information could be included in the report. As the case organization is small, it was decided to not add the name of the case organization in the report.

Above sections outlined the methodological considerations, including the research philosophy, the research strategy, and research setting. Furthermore, the data collection, data analyzes, and the considerations about reliability, validity, and ethical concerns. The next section presents the findings which emerged from applying the strategies and methods described above.

5 Findings

The findings section presents the results of the study. In the following section, the identified user archetypes are presented. Section 5.2 provides a definition and explanation of the relevant use-conditions and their respective dimensions. Section 5.3 outlines the findings related to how users frame the influence of use-conditions on the perceived use-value of OGD. Lastly, section 5.4 presents the corroborating evidence from existing research.

5.1 User Archetypes

Four archetypical users were identified within the organizational setting: business developers, data visualizers, software developers, and data scientists. The user archetypes are distinguishable by how close they work with OGD in their role within the case organization. Closeness to data indicates how likely a user is to process and manipulate the data. Hence, the further away a user is from the data, the more conceptual their work with data becomes.

Business Developers		Business developers work with data on a conceptual level, and rarely use data directly. Their main task involves identifying new customers and business opportunities as well as coordinating the delivery of OGD solutions.
Data Scientists		Data scientists work hands-on with data. They mostly work with raw data. Their main tasks involve collecting, processing, manipulating, and analyzing the data to make it useful for a specific OGD solution.

Software Developers		Software developers work with integrating data into a software solution. They use both pre-processed and raw data. Their main tasks involve developing the OGD solutions, ensuring it integrates and displays the data included.
Data visualizers		Data visualizers work with visual aspects of data. They primarily use data which is pre-processed for the specific purpose. Their main tasks involve designing the visual specifications of data and creating the visual aspects of an OGD solution.

Table 2: User Archetypes

The user archetypes have individual but complementary perspectives, which in combination facilitates the successful delivery of the OGD solutions. Some respondents have responsibilities which span different roles. However, respondents were assigned a main user archetype by assessing their self-reported responsibility, the emphasis they put on certain aspects of data, and the way they discussed data as a factor in their day-to-day work.

5.2 Use-conditions

When the user archetypes discussed how OGD was beneficial in their work, they related to a set of conditions which influenced their experienced use-value. This led to identifying two primary conditions through the analysis: data-openness and data-quality. Both conditions have a set of dimensions which were similarly identified during the data analysis process. Below sections provide a definition of the conditions and dimensions.

5.2.1 Data-openness

Data-openness involved both the availability and accessibility of the data. Data-openness was the most discussed condition among the respondents. Hence, it was a central condition for achieving the overall organizational goal of developing new solutions and innovating with data.

Data-openness included three dimensions: legal, economic, and technical.

1. The *legal* dimension focuses on questions of the appropriate use of licenses and legislation for commercial use. It involves risk concerns related to continuous OGD delivery.
2. The *economic* dimension focuses on questions of costs associated with the data procurement, including the cost associated with re-distribution of data.
3. The *technical* dimension focuses on technical accessibility to OGD and how access is provided.

These dimensions were discussed at varying degrees, with the legal dimension being the least expressed overall and the technical dimension being the most expressed. How close a user archetype worked with data influenced the way they discussed the different dimensions. For example, the closer to data user archetypes worked, the more likely they were to emphasize the technical dimension relative to the legal. Hence, business developers focused more on the implications of the legal dimensions than the remaining groups.

5.2.2 Data-quality

Data-quality includes assumptions and perceptions around the influence of data-quality on achieving the goals across the case organization (e.g., developing new solutions and innovate with data) and on the application of data in the work of the user archetypes. This dimension was the second most discussed condition in the interviews. Four dimensions were identified: sustainability, comprehensibility, accuracy, and functionality:

1. The *sustainability* dimension focuses on the perceptions around data-delivery, whether the update-rate is reliable and if data is delivered consistently. It involves the question of whether data is trustworthy.
2. The *comprehensibility* dimension focuses on the presence of metadata and specifications. It includes documentation to understand the data model, technical quality, and how data is measured. Comprehensibility does not mean that the various quality aspects are factual; rather it concerns itself with the information being documented in a manner where users can understand it.
3. The *accuracy* dimension focuses on the standardization of the data formats as well as on objectivity. With objectivity is meant how well data represents the reality it is supposed to. Furthermore, it includes the question of whether data is accurate to its specification.
4. The *functionality* dimension focuses on perceptions of the capabilities associated with datasets. The capabilities include interoperability between datasets, the amount of information within datasets, and the completeness of data, i.e. if data is of national or municipal scope.

The user archetypes emphasized different dimensions. The business developers focused mainly on comprehensibility. The data visualizers emphasized comprehensibility and sustainability. The software developers were primarily concerned with accuracy, while the data scientists focused on all dimensions while emphasizing functionality.

The above sections defined the user archetypes and use-conditions. The following section describes the findings which occurred from investigating how the use-conditions influence the perceived use-value of OGD.

5.3 The Use-conditions' Influence on the Perceived Use-value

This section describes the findings of how the user archetypes framed the influence of the use-conditions on the perceived use-value. First, the data-openness condition is described followed by the data-quality condition. Lastly, a summary of the differences in framing within each identified dimension is presented. Table 3 on page 29 contains an overview of the ways use-conditions influence the different user archetypes within the case organization.

5.3.1 Data-openness

Shared Frames of the Four Archetypes

When the Danish basic data program enabled the release of OGD in Denmark, it facilitated the start of the case organization. The central impact of the basic data program related to the three dimensions of data-openness; legal, economic, and technical. However, the release of all the relevant datasets is not complete yet⁷, and there are some datasets which are legally free which still lies behind a payment wall. All user archetypes considered the economic dimension of data-openness to be important. They saw the payments for data to be a barrier, even if the data was legally free. The payment wall meant they only used the dataset when hired by a customer to build a solution that needed this data. Hence, overall the economic cost of procuring data was considered a barrier to the case organization's ability to innovate.

⁷ Time schedule for release of datasets run until 2021. http://grunddata.dk/wp-content/uploads/2017/06/Tidsplan_GDP.pdf

Business Developers

Business developers had the highest focus on the legal dimension compared to the remaining user archetypes and they were the only group to consider the administrative burdens related to the economic dimension. They only discussed the technical dimension in passing related to the general problems they knew about, for example the issues faced by the data-scientists.

The economic and legal dimensions are typically considered upon initiating new business opportunities. The importance given to the legal and economic dimension by business developers might be explained by their main responsibilities. Since business developers work with identifying and searching for new business opportunities, they are faced with these questions when considering new opportunities. The business developers are the primary individuals responsible for new opportunities, and as such the legal and economic concerns introduce a higher degree of administrative burden on business developers' work than the remaining archetypes.

Licensing could influence the pricing and delivery scheme necessary for different solutions. Furthermore, the business developers mentioned various types of restricting licenses which would hinder the use of data and underlined the need for licensing considerations to be facilitating in the commercial data-use. One respondent stated; "licensing is important for data, since you with few words in a license, in fact, can restrict its use" [R5]. The business developers expressed an administrative concern related to keeping track of what data was used within their solutions, as data sometimes could be subject to attribution requirements from the data licenses. The business developers expressed a degree of uncertainty regarding when data would be combined or processed to a degree where it would no longer need to be attributed because it could be considered a new type of data. A similar concern in relation to the attribution of data was how big a part the dataset should be of a solution to be attributed. For example, was attribution necessary if the only data field used was the road names? Despite the significance of the licensing aspect, it was not perceived as a problem due to the public-sector's adoption of standardized OGD licenses in the Danish context.

Some solutions created by the case organization would use data which is legally, but not economically open. In these cases, the business developers perceived a lower degree of business potential because the case organization could only develop solutions for customers who had the appropriate rights to the data.

The legal and economic dimensions mattered in terms of ensuring contingency of the business operations. These considerations influenced the perceived risk associated with solutions development. For example, a business developer expressed concern about the sustainability of the distribution platform for OGD, and

highlighted the importance of having legislation in place to support this: “What happens if the funding of the basic data registers or the data distributor vanishes? In my opinion there is an antidote which is to ensure that the registers, as well as the data distributor, are regulated by law to as high a degree as possible” [R10].

The technical dimension was not central to the business developers’ reflections on data-openness, and during the interviews, the discussion of the technical dimension related only to data’s impact on their internal business processes such as finding and procuring the data. The small focus on the technical dimension might be because their work with data is primarily conceptual, and they do not face the troubles in applying the data within solutions. However, during a meeting, a business developer responded to the perceived complexity of integrating a dataset released as a service, “the more difficult, the better” [meeting note]. The comment suggests that business developers consider the complexity of the technical openness a business advantage, rather than an obstacle.

Data Visualizers

Data visualizers did not discuss the legal dimension and the economic dimension was mentioned only when identifying data-costs as a barrier. They considered the technical dimension of openness in terms of the accessibility to data-services.

The lacking focus on the legal and economic dimension for the data visualizers seems to be related to their focus on the visual aspects of data. In the work of data visualizers, the most crucial aspect is to create the visual design and usability for the end-users. Hence, data has mostly been processed or made available in a way where they can be sure that they can use it, and any economic aspects have been dealt with.

Data visualizers discussed the indirect influence of openness on their work. The central influences experienced was related to the impression that more services and technical tools for manipulating and displaying data were developed when data was opened. Thus, the benefits of data-openness were experienced more indirectly. One data visualizer mentioned; “it [data-openness] just does, that instead of being locked into this one pencil, you suddenly have a palette of color pencils next to which you can use” [R9]. Hence, the technological advances resulting from opening data had increased the toolbox at their disposal. The new tools influenced how they could manipulate and work with the graphical display of data. It enabled a more straightforward and diverse way of working with data. The data visualizers are good at conveying information, and they can see what type of data is missing to increase the solution’s usefulness to end-users. This understanding is visible in their role as mediators between the technical solutions and the end-users.

The technical dimension was discussed in relations to having access to appropriate services. The services which data visualizers utilized could be either government provided or internally developed data services. Overall, data visualizers used data services to deliver certain visual aspects that could be incorporated into the presentation of data to the end-user. A benefit of having data delivered through services was expressed by one data visualizer as; “the services we use, from my perspective, can be called directly. Before I would perhaps need to involve someone, who could create something [a solution] which could call further [re-direct the call] due to some limitations” [R2]. The statement indicates that having services available streamlines their technical data work, as they in some cases do not need to have additional adjustments made for the services they draw on.

Software Developers

Software developers only focused on the legal dimension in passing. Instead, software developers focused pragmatically on the economic and technical aspects of openness in terms of their work.

Software developers considered that they could develop solutions on top of data, if there were not any need to pay for it: “If they [data] are not open or accessible, well, then you will never start making a solution” [R3]. Even when data was legally open, if data was covered by a distribution cost it hindered their use. The payment barrier was related to the lack of opportunity to understand the data and if data could be applied in solutions. Hence, data-openness is a central way for them to gain knowledge of data and its capabilities. A respondent mentioned that removal of economic cost led to “another opportunity to provide suggestions for solutions to some problems because you have access and can see them [data]” [R11].

Software developers considered the technical dimension in relation to the context of the solution they were building. While it was beneficial to have access through services, this was only the case if the service could help them with the need they had in a context. Software developers considered that services could hinder the more general use-value of data, as it required first to get familiar with new technology. One respondent described it as follows “Now with the data distributor, a lot of nice web services [are created], but it becomes more difficult to use for regular people [... because] they need to get to know a lot of new technology, the web services, which might not be standardized” [R4]. The existing services were not applicable to their solution when the information delivered by the service differed from what was necessary for end-users. In these cases, it was necessary to find workarounds for their solution. While workarounds were possible, it was perceived as more beneficial to have access to the raw data instead, so that the case-organization could create the appropriate services in-house.

Data Scientists

The data scientists did not consider the legal dimensions as important. The economic dimension was important to data scientists as it had a direct influence on their perceived use-value of data. The data scientists mainly discussed the technical dimension.

The value for the data scientists is strongly related to the ability to work hands-on with the data and investigate the connections and nuances of the dataset outside their initially intended function. The data scientists' understanding of the connections between datasets is dependent mainly on the availability of data. Having data opened enables data scientists to test out solutions and new ideas against the data rapidly, which has a central impact on the use-value of data. Furthermore, data scientists ensured that data was collected for application within the case organizations solutions in local databases.

The lack of focus on the legal dimension might be due to this aspect not being a central responsibility of the data scientists. Instead, it was the responsibility of the business developers. Data scientists were instead impacted by the technical hindrances.

The economic aspects were a direct influence on their use-value, as the economic cost of procurement was considered a barrier for their exploration of the datasets. Hence, if there is a cost barrier, they have difficulty judging data's use-value. As one data scientist put it; "It [the cost barrier] just did, that nobody makes solutions on it [data] at all. Then they [data] can rot [...] in some place with some IT company, but nobody will use them. At least not private organizations." [R7]. Furthermore, having an economic barrier might put up a technical hindrance in integrating data to their local databases.

Another aspect relating to economic openness was having free access to tools which could be used to manipulate the data. Appropriate tools can provide the ability to see and process the data, which is necessary in order for the data to have any use-value. In the case of geographical data, which is a central type of OGD for the case organization, the existence of an open source tool ensures a low entry fee to the processing in a market where specialist tools typically have been covered by expensive licensing. As remarked by a data scientist "But they [data] are not as free, if you have to pay 50.000 to get started with a program [to use data]" [R8].

The data scientists primarily focused on the technical dimension of data-openness, as this impacts their ability to manipulate and analyze data. To make analyses for new services and to innovate with data, it is necessary to have access to the raw data. For most of their analytical and experimental tasks, data scientists preferred access to the full dataset instead of having access through web-services. The problem with data

delivered as services was that it complicated their ability to access the raw data. Services were considered to lock data inside a “black box” as they only deliver parts of a complete dataset. As one respondent put it; “when you try to figure out who your users are and display it all [the datasets] as services, then all the users you have not considered are screwed” [R6].

Furthermore, service access can be challenging to integrate with the tools they use to manipulate data. Hence the requirement of new plug-ins for the data to be accessible and useful within the tools lowers the immediate use-value of the data. Related to the less analytical tasks data scientists did, they acknowledged services to facilitate access to data for specific purposes quickly. However, they wanted the option to automate the data collection process: “It is nice to be able to download [data] without having to, e.g., access a website and put in your nemID or something else that cannot be automated” [R12]. While download access appeared more desirable for accessibility, an ability to directly read the database and have an internally updated copy could be preferred. Particularly as database access could combine the real-time benefits of services with the data which they could then use for service innovation.

Condition	Business Developer	Data Visualizers	Software Developer	Data Analysts
Openness	(Licensing) can hinder or complicate the use within the commercial setting.	Openness leads to the development of tools and services. Results in new and diverse ways of using data.	Increase their knowledge of data and its applicability.	Leads to an increase and sustaining of their knowledge of data and its ability to be combined.
	(Economic) the presence of costs creates administrative overhead.	(technical) access through services increases use-value.	Openness eases the ability to troubleshooting applications.	(technical) access to the complete datasets enables experimentation.
	(Technical) complex access increase the use-value, by enabling business opportunities.		(technical) Both service and download can hinder or promote the use-value.	(technical) service-access creates complexity in the data collection which lowers the use-value.
Quality	(Comprehensibility) metadata facilitates the assessment of business opportunities.	(Sustainability) Trustworthiness of data matters to the extent it matters to the customer.	(Sustainability) Is important to the users of the application. Thus it should be considered.	(Sustainability) is nice to have, but not need to have.
	(Functionality) National scope creates business opportunities.	(Comprehensibility) impacts their ability to guide the design of a solution for their customers.	(Accuracy) ease of integrating data into their software.	(Comprehensibility) facilitate the establishment of datasets accuracy.
		(Comprehensibility) leads to understanding data's visual application opportunities.		(Accuracy) eases the maintenance of local data copy
				(Functionality) impacts their ability to experiment with new combinations.

Table 3: use-conditions influence on user archetypes perceived use-value

5.3.2 Data-quality

Shared Frames of the Four Archetypes

The expression “get data into play” is used as a way of describing a core belief shared across the user archetypes. “Get data into play” includes the assumption that data-quality is a consequence of using data. As explained by a software developer; “what happens when data is put into play is that its flaws are displayed” [R3]. These flaws are detected by the case organization when they utilize data and subsequently reported to the public agency responsible for correcting the dataset.

Data-quality is bound to the context of its creation, and the specifications are developed for a specific purpose. A business developer specialized in the public-sector mentioned: “a new type of usage [by the public-sector] requires a redefinition of the purpose and specification of the dataset [which] introduce new measures for data-quality.” [R10]. New specifications can thus create a different standard for data-quality. Since data is created for a pre-specified purpose within the public-sector, the data-quality standard is defined by them, and dependent on them.

Data-quality furthermore depends on the context of data’s application. For the case organization, data-quality depends on how “fit for purpose” data is. From this follows that data-quality is not always a hindrance for the case organizations data-use. Even if the case organization has an unstandardized dataset, it can be beneficial to use within a solution. As a respondent pointed out: “It is just an example that some half- or incomplete data, [where] people can almost write what they want. But it can still make sense to present them for a user” [R11]. The main concern for the case organization is whether they know when the specifications and actual data-quality does not add up.

Business Developers

The business developers did not discuss sustainability and accuracy. Instead, business developers considered the comprehensibility dimension a central point for data-quality. The functionality dimension was limitedly discussed.

The lack of focus on the sustainability dimension is surprising since data-delivery and reliability of update-rate seem to be critical aspects of the proper functioning of the business. This might be explained by the focus of the business developer’s role within the organization. Business developers are not directly dealing with the problems arising from data-delivery and reliability, but primarily focusing on what are possible avenues of development given their knowledge of the data. Similarly, it might be the explanation for why the accuracy of data was not a central focus for business developers.

Business developers focus on ensuring a good inflow of business to the case organization, so, the use-value of OGD is highly related to their ability to deliver value to their customers. Value delivery requires that business developers can accurately compare the value of data to possible business avenues. The business developers consider that business opportunities are found when the case organization takes the risk of exploring the datasets applicability to different problems, which can happen regardless of the overall data-quality. They do not consider data-quality to be a hindrance to the case organization as a whole, because of the technical capabilities within the case organization. As stated by a business developer: “I can't think of any examples, where poor data-quality actually has been a showstopper.” [R10].

The business developer's knowledge of data-quality is essential to their work since they draw on this when judging new opportunities. Hence, the comprehensibility dimension is important as business developers perceived use-value depends on the knowledge collected about the datasets. The knowledge they have about the data-quality enables them to assess the contextual requirements of business opportunities better.

The need to understand data was, furthermore, reflected in having “domain knowledge,” which was the closest business developers came to discuss the sustainability and accuracy dimensions. Domain knowledge is the insights they have about the different aspects of data-quality. The business developers primarily draw on this when they are exploring opportunities. One respondent mentioned: “If we, for example, have a national dataset where you register recycling stations, then we are the ones who know, that it only happened in Tårnby and Holstebro municipalities” [R5].

The business developers consider the functionality to be an important aspect related to e.g. data scope as they consider data should be nationally scoped to be an enabler for exploring broadly relevant business opportunities; “data scope is directly related to the business potential” [R1 - note]. When data was nationally scoped, they could develop solutions for a broader market, whereas locally scoped data had limitations in their ability to create interesting solutions for a broader market segment. Locally scoped data was only considered beneficial when the customer was a municipality who needed a solution which applied this specific data.

Data Visualizers

The data visualizers only discussed the sustainability and comprehensibility dimensions. They did not discuss the accuracy or functionality dimension of data-quality.

The sustainability dimension was important for the data visualizers since their focus on data-quality is mainly related to ensuring the usability of solutions to the end-users. The perceived use-value of OGD for

data visualizer thus depends on the task at hand, and the context that is faced by the end-users of the solution. The customer's requirements and the end-users need then become the measuring stick for the data-quality. As such, if their customers care about the data being trust-worthy, then it is essential to the data visualizers. A data visualizer explained this while referring to public customers: "It [data-quality] is relevant to the degree the customer thinks it is relevant. It depends on the projects, but overall that if you cannot trust it [data] as it is, then you don't dare to use any of it" [R9]. It seems then that the quality requirements may be changing for the data visualizers. While it was clear that sustainability mattered when considering the public customers, overall the most important aspect for data visualizers was similar to the business developers, that is; whether they understood the data-quality.

The comprehensibility dimensions were important to the data visualizers since the understanding of data-quality enables them to understand how data is useful to meet the requirements and end-users needs. Knowledge enables data visualizers to provide appropriate advice to the customers who need to use the visual design they are creating. In data visualizers technical work, data-quality matters to the extent it can be combined with their solutions, and live up to the customers' demands: "The data-quality is not so important to me, of course, it is important for the user and the product we deliver, but for me [...] I just need some data [...] and know that I can combine it" [R2].

The data visualizers did not discuss the accuracy or functionality of data. It is a bit surprising that data visualizers did not explicitly discuss these two dimensions, but it might be because they did not face the problem in their work. A possible explanation is that they focus on the requirements of the customers and end-users. Thus, data visualizers do not experience any direct impact of non-standardized data or a lack of interoperability of datasets. However, if aspects related to standardization and interoperability would be a concern for the customer or end-user, they might also be for the data visualizer.

Software Developers

The software developers limitedly discussed the sustainability and comprehensibility dimension. Instead, they primarily focused on the accuracy dimension, and to a lower extent the functionality dimension.

Software developers did not put a central emphasis on the sustainability dimension, and they only briefly discussed update rates. They did acknowledge that problems with the trustworthiness of data could cause problems for the end-users of the solutions if data were not trustworthy. A software developer said: "It is bad for the ones who need to use the solution, that data is erroneous, but it is never the suppliers [the case organizations] mistake" [R3]. The reason for the lack of focus here might be that the case organization has processes in place to create a local server copy of the data which their solutions need, which is then applied

by the software developers. Hence, they do not feel the impact of sustainability directly unless they are also helping with the development of workarounds.

The reason they did not focus on comprehensibility relates to their perception of the data-openness. For software developers, data-openness was the aspect which allowed them to comprehend the data, so they did not necessarily consider the metadata explicitly.

Instead, the accuracy dimension was an important aspect for the software developers. Software developers perceived use-value of OGD to depend on the ability of data to be integrated seamlessly into solutions without having to create several workarounds. Accuracy ensures data format is consistent for the IT systems to handle the data: “with geographical data [...] you need to have certain specifications met before data is valid, and [existing] data doesn’t always comply. That means IT systems break down” [R4]. So, if the formats are accurate and consistent, they can apply data. However, accuracy does not need to be a hindrance, firstly because software developers can create workarounds which handle the data imperfections. Secondly, in some contexts, if the software developers know that data is not perfect, then they can still use it if data adds value to the solution. A respondent mentioned, “It is often important to be aware of the quality. That you do not think it [data] is complete if it is not. Then you can often live with data that is not perfect” [R11].

Software developers discuss the functionality dimension as relating to the need for data being uniform across the country. Uniformity made for better solutions and eased the ability to integrate data into solutions. For software developers to integrate data into a solution, it was beneficial if the datasets were interoperable. Interoperability would lower the need for workarounds.

Data Scientists

The data scientists discussed all four data-quality dimensions, which might be because data scientists work hands-on with data, and thus are aware of all aspects of data-quality to a larger extent than the remaining user archetypes. Data scientists discuss functionality the most.

Data scientists do not perceive data-quality as something highly fixed, because they can further process, combine, and improve data, so it is suitable for other purposes than the one for which it is initially intended. One data scientist mentioned; “Basically, it is the same data we are redistributing [...] in a completely different format, and much more useful.” [R12]. Yet, they do consider all the data-quality aspects to have some degree of importance.

The sustainability of data matters as an indicator of trustworthiness since data is from an authoritative source: “They actually come from an authoritative source [...] there is not a crown stamped on, but something similar. There is a somewhat higher trust in the [data] since these are being used to govern society.” [R12]. However, even if data was authoritative, it was not as important as data being functional. Hence, authoritative data was considered “nice to have.”

Comprehensibility is essential when dealing with new datasets. Metadata and data models are also relevant as software developers need to know what various codes within the data might stand for, which is not always clear from the raw data.

The accuracy is typically not on the metadata level. Hence their understanding typically comes from the data-openness combined with the comprehensibility dimension. Software developers use technical specifications to assess the accuracy of the data. Hence, software developers get an idea of the expected data-quality by looking at the metadata and compare expectation with the OGD to deduce how accurate data is. Accuracy mattered to data scientists as it could facilitate the collection of the data copies. Local copies ensured less dependence on services from the public-sector as they always had a local copy which they could rely on.

For data scientists, functionality is an essential influence on the use-value of data. They like to have access to more data as data is a central aspect in their analytical work. For example, a respondent stated: “sometimes it [the analysis] is about pulling more data in than necessary, and then you have to see what sticks” [R8]. The interoperability of datasets matters because, if there are only a logical connection and not a “physical” connection of two data points, then data can become erroneous.

The amount of data points within a dataset is important, that is, how many different real-life objects are present and linked to the dataset. The information within a dataset matters as it impacts the possibilities for using the data for various purposes. Furthermore, nationally scoped data is essential in order to enhance their analyses. The functionality dimension is particularly visible when they compare OGD such as Geo Denmark data⁸, with open street map data⁹, which is open data updated in a crowd-based manner. For some purposes, data scientists need to use open street map data instead of the OGD equivalent (Geo Denmark data) as its functionality is higher. As mentioned by a data scientist; “Let’s say data was perfect in both

⁸ OGD which the public agency Geo Danmark is responsible for.

⁹ A geographical dataset collected through crowd participation worldwide.

cases, then you would still be able to do more with ‘open street map’ than what you can with Geo Denmark data. Simply because there is more in the [open street map] model than the [Geo Denmark model]” [R6]. Functionality had the highest influence on the data scientists use-value. The strong influence might be due to interoperability making it easier for data scientists to conduct data analyses or calculations. Furthermore, it facilitates testing new ways of combining data and creating services which can subsequently be integrated into software and visualized.

5.3.3 Summary

Data-openness

The four user archetypes put weight on different aspects of the data-openness condition. Business developers primarily focused on the legal and economic dimension, albeit recognizing some technical challenges. The three other archetypes primarily focus on the technical dimension, although they did consider the economic dimension to some extent. The data visualizers primarily considered the more holistic influence of openness as leading to developments in the tools and services which they used and preferred access to data through services. Data-openness led software developers to understand the use-value of data relative to the solutions they were developing. Data-openness made it more accessible to troubleshoot applications for software developers, and they could use access to both services and raw data. For the data scientists, data-openness meant they could explore how to combine data and create new interesting analyses. They primarily wanted access to raw data and considered the economic barrier to be a significant issue.

Data-quality

The four user archetypes framed the influence of the data-quality dimension differently. Both business developers and data visualizers focused centrally on the comprehensibility dimension. However, the business developers aim was to understand data’s applicability related to business opportunities, whereas, data visualizers aim to understand and link the requirements of a solution and the end-users needs to the possible datasets available. The software developers were mainly concerned with the accuracy dimension as data’s accuracy could have direct influences on their ability to create solutions which did not break down. Lastly, the data scientists focused on all the dimension in lieu of their role working hands-on with the data. The data scientists did focus mainly on the functionality as an essential part in their ability to create analyses, new calculation methods, and experiment with data in general.

Having outlined how use-conditions influence the user archetypes differently, the following section presents the evidence from existing research which relates to the use-conditions analyzed above.

5.4 Literature Corroboration

The literature corroboration was done to ensure triangulation of the evidence with existing research. From the existing literature, the corroborating evidence ranged from being a few sentences in the findings to the inclusion of a dimension in part or encompassing the overall condition. Existing research identifies the use-conditions but typically did not identify their influences on users. Nevertheless, certain aspects still overlap with existing research. An overview of the corroborating data for the use-conditions and dimensions can be found in Table 4 on page 42. The following section outlines the corroborating evidence in relation to this study.

5.4.1 Data-openness

Some existing research applied the concept of data-openness in their work (T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a). This condition was, furthermore, described in existing literature as; attainable (F. Welle Donker & van Loenen, 2017) or availability (T. Jetzek, 2016; Safarov et al., 2017).

Research utilizing data-openness has conceptualized it as an enabling factor for value generation mechanisms. The concept includes two of the dimensions found in this research under openness, namely the legal, and technical accessibility (T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a). The concept did not consider the economic aspect as a significant dimension. In comparison, all three dimensions were present under the construct “attainable” (F. Welle Donker & van Loenen, 2017).

In some literature, data-openness has been found to enable innovation (T. Jetzek, Avital, & Bjorn-Andersen, 2014; Safarov et al., 2017). The present research found evidence of this influence as the case organization aimed to innovate and build new services on top of data. Across all user archetypes, it was clear that data-openness was a necessary condition for data to have any use-value. Closed data was only of use in the context where data was created.

Legal Dimension

The legal dimension was highly present in existing literature. The treatment of this dimension included focusing on the legislation and policies in place (Attard et al., 2016b; Safarov et al., 2017; A. Zuiderwijk

et al., 2015) as well as the necessity of having open copyrights or licenses (Abella et al., 2017; Cranefield et al., 2014; T. Jetzek, 2016; Safarov et al., 2017; A. Zuiderwijk et al., 2015).

The literature strongly supported the legal dimension. However, this research found that the legal dimension was primarily a concern among the business developers.

Economic Dimension

The economic dimension was only considered necessary in a few articles. Some articles identified charges to be hindering innovators and start-ups from having access (Lakomaa & Kallberg, 2013; F. Welle Donker & van Loenen, 2017), other research considered affordability an affordance of liquid open data, and outlines the complex challenge it was for the public sector to make data freely available in the Danish context (T. Jetzek, 2016).

Some existing research discussed the economic dimension, but not in a way that directly corroborated with the present study. For example, one study discussed economics related to the public-sector financing data publication (Attard et al., 2016b). Another study found the economic dimension not to be a concern for their case organization, yet openness did lower the complexity in the administration of including external information (T. Jetzek, Avital, & Bjorn-Andersen, 2014).

Some articles discussed the users relative to economic openness. Some studies found that fees mattered to innovators and start-ups (Lakomaa & Kallberg, 2013; F. M. Welle Donker, 2016), while other research considered fees to not matter for professional users (F. Welle Donker & van Loenen, 2017). What is meant with professional users compared to innovators and start-ups is not clear, but it might relate to the capital available to the user's organization. The centrality of the economic dimension is not clear from the existing research, but the present study did identify it across all user archetypes. It might be, that the concern exists primarily on an organizational level and that it is on the organizational level where differences can be found, rather than on a user-level.

Technical Dimension

The technical dimension as defined in this research had a great deal of corroborating evidence in existing literature. The main conceptualization was technical accessibility and the ease of access to OGD (Abella et al., 2017; Cranefield et al., 2014; T. Jetzek, 2016; T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a; Lakomaa & Kallberg, 2013). Technical accessibility was furthermore discussed within effort expectancy from the unified theory of acceptance and

use (Saxena & Janssen, 2017; Anneke Zuiderwijk et al., 2015), and as processing capabilities of OGD systems (Alexopoulos et al., 2016).

Technical access is considered a requirement for the ability to execute the business plan (Lakomaa & Kallberg, 2013) and maintaining a library of content (T. Jetzek, Avital, & Bjorn-Andersen, 2014). Several papers included the aspect of discoverability of data (Cranefield et al., 2014; A. Zuiderwijk et al., 2015). In this study, discoverability did appear in a few interviews; however, not enough evidence was found to incorporate it into the findings. The lack of evidence could be because the data utilized by the respondents was mainly nationally scoped and available through the basic data programs' data distributor. Some observations showed that when the case organization tried to develop and brainstorm about new and innovative services, data discovery could present a challenge.

The unified theory of acceptance and use of technologies includes the construct "effort expectancy" (Venkatesh, Morris, Davis, & Davis, 2003). Effort expectancy contained partial evidence for data accessibility and is found to be negatively linked to intention to use OGD. The negative relation suggests that when data is difficult to access the use-value is lower (Saxena & Janssen, 2017; Anneke Zuiderwijk et al., 2015). Similarly, processing capabilities of the infrastructure are found to be essential to achieve the objectives of users (Alexopoulos et al., 2016). To this end, services can be considered to provide a level of processing capability within the technical infrastructure.

In this research, the link between effort and use was not straightforward, since the user archetypes exhibited a considerable variation in relation to what type of access would be considered challenging, and whether they preferred access to services or raw data. Effort expectancy would thus largely depend on the user archetype's skillset. For example, data scientists would consider a high degree of effort expectancy if data was published as services. However, data visualizers would consider the effort expectancy of working with raw datasets much higher. Furthermore, business developers would consider effort expectancy as a positive trait, since it leads to a competitive edge.

Existing research suggests that utilizing OGD through a system has a positive influence on the future behavior of the users. However, weak evidence is found to support it (Alexopoulos et al., 2016). The weak support is not surprising when considering the findings of this study. Differences were found in how technical openness influences user archetypes, hence, it would be difficult to have highly significant findings when investigating a broad set of users.

5.4.2 Data-quality

Within existing research data-quality is widely used. There is a variety of subdimensions quoted which did not always corroborate with the perspectives of the user archetypes found in the present study. Often, existing research contains only limited findings of relevance to provide substance to the subdimensions. The condition of data-quality was found under the notion of data governance (Eckartz et al., 2016; T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a), data-quality (Cranefield et al., 2014; Eckartz et al., 2016; Lakomaa & Kallberg, 2013; Purwanto et al., 2017; Safarov et al., 2017) and technical usability (Attard et al., 2016b; Eckartz et al., 2016; T. Jetzek, 2016; F. Welle Donker & van Loenen, 2017).

Existing research identified data-use to leads to data-quality (Cranefield et al., 2014), this was similarly found to be the case in this research, providing some evidence that data-quality might to a certain degree depend on the use of data.

One study found that data-quality is the most discussed aspect in research, and hence, the most important condition to use OGD (Safarov et al., 2017). Data-quality was not the most discussed dimension in the present research. Instead, data-openness was given more importance.

A study found high data-quality leads to less processing necessary for subsequent users of the data (Lakomaa & Kallberg, 2013), and that users require a level of quality to use data (Cranefield et al., 2014). These aspects were not accurate when considering the findings in this study since user archetypes considered data-quality to be more fluid. Instead, the processing and quality were perceived as bound to the context where data was used, and not necessarily constant.

Sustainability Dimension

The sustainability dimension includes the perception of data trustworthiness and if data delivery is reliable. Existing research including the sustainability dimension found that uncertainty of data supply introduces risk for the private-sector users and that it is essential that the ability to continuously collect data will be guaranteed (T. Jetzek, Avital, & Bjorn-Andersen, 2014). The continuity of data supply was evident from the present study as well, in the risk concerns voiced by the business developers. In existing research, the trustworthiness of data relates to data coming from a respectable source (T. Jetzek, Avital, & Bjorn-Andersen, 2014), the present study found a similar aspect as user archetypes viewed the reputable or authoritative source to be of some importance.

Existing research included the aspect of delivery time (F. Welle Donker & van Loenen, 2017). Delivery time was related to how many days it would take to receive the data. These long waiting times between requesting and receiving data was not discussed much during the interviews. Meeting observations did, however, hint at this concern. Furthermore, in the present research, the discussion was not related to days, but rather a discussion of the response time of services or how quickly one could download a full dataset. Not enough evidence was collected to theorize the aspect of data delivery explicitly.

Comprehensibility Dimension

The comprehensibility dimension includes the need for metadata. Existing research considers metadata a central part in facilitating data use and value generation (e.g., Jetzek et al., 2013b; Purwanto et al., 2017). Some existing research includes metadata as an aspect of other constructs, such as resource governance (T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjørn-Andersen, 2014a), or “usable” (F. Welle Donker & van Loenen, 2017). Several other aspects of data-quality were included under these overarching dimensions as well. In the present research, the inclusion of metadata was considered most relevant for data scientists as metadata was relevant for judging the accuracy of data. It was not clear from the present research, to what extent metadata itself was the key for business developers and data visualizers’ comprehension, and how much of their knowledge was a result of the data scientists sharing the insights with the organization.

Existing research found that it was important for users to “know precisely and in a scientific manner the methodology of how the data were produced” (A. Zuiderwijk et al., 2015, p. 205), and that users need documentation on the API to understand the content of the data (Purwanto et al., 2017). In this research, these aspects were lacking on a meta-data level, albeit the user archetypes considered them to be necessary.

Accuracy Dimension

This research considers accuracy to include objectivity and accuracy of data related to the specifications, and closeness to reality. The notion of accuracy has been found in existing research as technical standards (Abella et al., 2017) and a subsection of interoperability which focuses on standards (Eckartz et al., 2016; T. Jetzek, 2016). In this research, technical standards mattered primarily to the software developers as it influenced their ability to maintain local datasets.

In some research, interoperability is considered a conceptual affordance of liquid data which includes the idea of “semantic and syntactic interoperability” (T. Jetzek, 2016, p. 98). This partially overlaps with the accuracy dimension of this research, in relation to the standardization. Yet, the conceptual interoperability also included e.g. the notion of standardized metadata, which this research instead considered falling under

the comprehensibility dimension. The reason for the discrepancy between concepts might be explained by the present focus on the user-perspective instead of the public-sector perspective which the existing research takes. The user archetypes were considered with standards to make their work easier, but also cared that the standards were representing reality. Whereas, standardization of data and the explanation of standards through metadata might be more readily grouped when planning the OGD initiative as these two aspects are interlinked.

Functionality Dimension

It was challenging to find corroborating evidence directly related to the functionality dimension aside from the aspect of interoperability (Eckartz et al., 2016; T. Jetzek, 2016). One article from the literature review includes interoperability as an affordance of liquid open data, where the emphasis was given to having clear semantics and syntax of data (T. Jetzek, 2016).

One reason for the difficulty of finding corroborating evidence might be, that functionality in the present research was considered the capabilities of datasets to interoperate, but also that they had functionality in terms of data scope, and amount of information. Another reason might be, that data scientists to whom this aspect mattered the most, are not typically consulted in existing literature which investigates how to facilitate the use of OGD.

The above sections presented the findings related to use-conditions influence on the user archetype's perceived use-value. Furthermore, it investigated existing literature for corroborating evidence. The next section will provide a discussion of the central lessons to be learned from the present study.

Condition	Literature Support	Dimensions	Literature Support
Data-openness	(T. Jetzek, 2016; T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a; Safarov et al., 2017; F. Welle Donker & van Loenen, 2017)	Legal	(Abella et al., 2017; Attard et al., 2016b; Cranefield et al., 2014; T. Jetzek, 2016; T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a; Safarov et al., 2017; A. Zuiderwijk et al., 2015)
		Economic	(T. Jetzek, 2016; Lakomaa & Kallberg, 2013; F. Welle Donker & van Loenen, 2017)
		Technical	(Abella et al., 2017; Alexopoulos et al., 2016; Cranefield et al., 2014; T. Jetzek, 2016; T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a; Lakomaa & Kallberg, 2013; Saxena & Janssen, 2017; Anneke Zuiderwijk et al., 2015)
Data-quality	(Attard et al., 2016b; Cranefield et al., 2014; Eckartz et al., 2016; T. Jetzek, 2016; T. Jetzek, Avital, & Bjørn-Andersen, 2014a; T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; Lakomaa & Kallberg, 2013; Purwanto et al., 2017; Safarov et al., 2017; F. Welle Donker & van Loenen, 2017)	Sustainability	(T. Jetzek, Avital, & Bjorn-Andersen, 2014; F. Welle Donker & van Loenen, 2017)
		Comprehensibility	(T. Jetzek et al., 2013; T. Jetzek, Avital, & Bjorn-Andersen, 2014; T. Jetzek, Avital, & Bjørn-Andersen, 2014a; Purwanto et al., 2017; A. Zuiderwijk et al., 2015)
		Accuracy	(Abella et al., 2017; T. Jetzek, 2016)
		Functionality	(Eckartz et al., 2016; T. Jetzek, 2016)

Table 4: Corroborating Literature

6 Discussion

The discussion section will start by summarizing the findings. Section 6.2 describes the contributions to research, followed by section 6.3 which provides recommendations for practice. Lastly, section 6.4 discusses the studies limitations.

6.1 Summary of Findings

Investigating how organizational users frame the influence of use-conditions on the perceived use-value led to several valuable insights. Firstly, four distinct user archetypes were identified within the case organization, based on how they used data in their everyday work. The groups included business developers, data visualizers, software developers, and data scientists. Their OGD use ranged from conceptual to direct data manipulation (see section 5.1).

Secondly, the use-conditions *data-openness* and *data-quality* were identified as relevant to all four user archetypes. *Data-openness* includes a legal, economic and technical dimension. *Data-quality* includes the dimensions functionality, comprehensibility, sustainability, and accuracy (see section 5.2).

Thirdly, each user archetype had distinct ways of framing how use-conditions influence the perceived use-value of OGD and perceived the relative importance of the dimensions differently (see Table 3). A few shared frames were found across all groups, which suggest some perceived influences depends on organizational level factors. For the most part, the archetypes demonstrated differences in their framing. An example of contradicting perspectives within data-openness is that business developers considered complexity in technical access beneficial in providing a competitive advantage. In contrast, the data analysts considered complexity in technical access to complicate their data collection efforts, and hinder experimentation. In relation to data-quality, data scientists were more concerned with the functionality of data, the software developers cared more about accuracy, and the data visualizers and business developers cared primarily about the comprehensibility. Hence the user archetypes require different things for use-conditions to positively influence their use-value. An example of a shared frame is economic openness, where free access to data was considered an essential enabler for use by all user archetypes.

Lastly, the premise on which the user archetypes judge how use-conditions influence the use-value of OGD can be summed up as follows. Business developers consider data in terms of its ability to facilitate business opportunities, and lower administrative overhead. (i.e. “can data solve the problem?” and “are there restrictions to data-use which affect the commercial application?”). Data visualizers consider data in terms

of how it can convey the necessary information to the end-user (i.e., “what requirements do the customer and end-user have to data”? and “can data display what is necessary to convey?”). Software developers consider data in terms of how solutions can integrate the data consistently (i.e., “does data have a consistent format?” and “can data be integrated without workarounds?”). Data scientist consider data in terms of how it can be combined and analyzed (i.e., “are there keys across datasets which facilitate integration?” and “how does the data compare to its specification?”).

6.2 Contribution to Research

6.2.1 User Archetypes

The user archetypes are an essential addition to research as they enhance our understanding of what types of OGD users exist. The presented classification of user archetypes has the following advantages:

First, the user archetypes were classified by how they work with data. This stands in contrast to the more holistic classifications in literature, where the focus is on overall objectives or motivations to use OGD (Magalhaes, Roseira, & Manley, 2014; Smith & Sandberg, 2018). The respondents in the present research exhibited characteristics of multiple groups in the existing classifications. As such, the presented user archetypes avoid the volatility arising from focusing on the outcome of use, by instead centering on the way a user utilizes data.

Secondly, the classification does not depend on demographic indicators to identify differences among users. The unified theory of acceptance and use of technologies (Venkatesh et al., 2003) is applied in OGD research related to user motivation (Saxena & Janssen, 2017; Anneke Zuiderwijk et al., 2015). This research-stream identify differences between users based on demographic variables such as gender and age. This risks that research reflects biases stemming from either gender or age distributions in the fields where OGD is used the most, which includes the data/IT sector (Magalhaes & Roseira, 2017). The suggested user archetypes are one way of introducing the contextual and specific conditions lacking when applying the unified theory of acceptance and use of technologies framework (Anneke Zuiderwijk et al., 2015). The archetypes can furthermore enable investigation of OGD users over time, without being impacted by fluctuations in the age and gender distribution of the workforce.

Lastly, the proposed classification does not distinguish between the direct users in terms of 1) the legal nature of organizations, i.e., whether it is a non-governmental organization, or commercial business (Safarov et al., 2017), or 2) include two potential analytical levels, i.e. organizational and individual (citizens, developers, researchers). The archetypes in this research have the potential to span these pre-

existing classifications for investigating how the use-value can be facilitated. For example, the converging frames identified within the economic dimension indicated that use-conditions might have some degree of organizational level impact. By understanding the user-perspectives, it became possible to identify organization-level shared frames, which indicates that some dimensions might depend more on the organizational reality than they do the individuals. Furthermore, the user archetypes could also be used to explore differences regarding the legal nature of organizations. For example, by comparing the user archetypes in commercial vs. non-governmental organizations.

6.2.2 Granularity in the User-perspectives

The findings demonstrate different user archetypes perceive similar conditions to influence their perceived use-value in different ways. The difference in perception is significant because it suggests the intended impacts of facilitating use vary depending on which user archetype one aims to enable. It implies putting user archetypes at the center and viewing the way they use data as determining the facilitation of data-use. This proposition does not contradict existing research's use of skills as a condition for using OGD (T. Jetzek, Avital, & Bjorn-Andersen, 2014; Safarov et al., 2017). Instead, it embraces the plurality of skills necessary for creating innovative solutions from OGD and put this at the center to facilitate the insights of different user archetypes. The change of perception moves forward our understanding of the expected and actual impacts of OGD (T. Jetzek, Avital, & Bjorn-Andersen, 2014), by proposing that use facilitation depends on how users apply data in relation to their role, and not necessarily on a variety of user-goals (F. Welle Donker & van Loenen, 2017). Hence, research investigating why expected values are not realized should instead adopt a more granular user-perspective.

6.2.3 Levels of Analysis

This research contributes to overcoming the “chicken and the egg” paradox (T. H. Jetzek, 2015). By focusing on use-value for organizational users rather than abstract benefits and impacts. This research highlights that valuable insights can be found when understanding the perceived use-value. The findings showed that user archetypes articulated which benefits they experienced (the use-value) and could link the experiences to the use-conditions which contributed to realizing the use-value. Existing research proposes a set of societal level measures for assessing the impacts (F. Welle Donker & van Loenen, 2017). While these can be beneficial in assessing the overall impact of an initiative, they seem far distanced from the reality of the organizational users. The lack of understanding the organizational user archetypes might be a contributing factor for the difficulty to facilitate the motivation of public-sector employees working with OGD publication (SEISMONAUT, 2017). The lack of motivation might be related to the values at the

societal level being abstract and difficult to comprehend. Instead, public-sector employees could benefit from seeing the actual impacts that their work with OGD has on the individuals. Thus, it would be beneficial for research to move beyond the present focus on the societal level of analysis and adopt a multilevel perspective including societal, organizational and individual levels of analysis (Lepak et al., 2007). Moving to a multilevel perspective can help move beyond the use of abstract impacts, to discover the actual impacts on users, organizations, and society at large. By understanding the OGD user archetypes, it is possible to expand and enrich the user-perspectives to have a better understanding of who are users, the effects they produce, and whom these effects benefit.

6.2.4 The Dimensions of Use-conditions

The findings showed that different dimensions of the use-conditions mattered to the user archetypes, and some dimensions seemed to matter on an organizational level. When comparing this study's findings with the corroborating evidence, it was clear that there were differences in the importance attributed to the dimensions. For example, existing research focused more on legal openness than economic openness. While the viability of the case organization's solutions depended on having appropriate licenses that did not constrict commercial use (Cranefield et al., 2014). The findings from the economic dimension demonstrated that data could be legally free, without it being economically free. These payments were a barrier to utilizing data freely within any solution. The strong representation of the legal dimension relative to the economic dimension in existing literature suggest a bias in the research efforts which might be related to two aspects. 1) previous research has focused mainly on the public-sector perspective, to whom the question of policies and licenses is a central concern of the publication process (Sayogo & Pardo, 2013). 2) Existing research including user-perspectives is likely to talk with representatives who could fall under the category of business developers.

One existing case study within a multi-national organization demonstrates that the economic dimensions are not highly important (T. Jetzek, Avital, & Bjorn-Andersen, 2014), however research on IT entrepreneurs, demonstrates that charges mainly hinder the access to data (Lakomaa & Kallberg, 2013). This contrast suggests that indicators such as the size of the organization and objectives might play a significant role in the economic dimension. On the other hand, technical openness matters to a larger extent on the individual level and depends on the way the users worked with data. Hence, this research suggests that the identified conditions and their dimensions provide a framework which, combined with the level of analysis, can provide interesting insights for which use-conditions are essential to different users.

6.3 Contributions to Practice

The findings in this research led to identifying five important impacts for public managers who aim at publishing OGD to enable the ability for the private-sector to innovate with it, and create new solutions:

6.3.1 Get to Know Your Users

This research shows that use-conditions influenced user archetypes in different ways. Public managers responsible for developing OGD initiatives should make sure to understand the different users. So-called democratic dialogues (Davies, 2010) between open data providers and users (Anneke Zuiderwijk et al., 2015) are taking place in Denmark. However, the user-side of democratic dialogues risks being represented mainly by business developers, which could result in a skewed basis for prioritization of OGD initiatives. For example, data scientists considered the use-value of data lower if the complexity of collecting it was high, the business developers considered the use-value of data higher as it makes it easier to achieve some competitive advantage. To balance the perspectives of users, efforts need to be taken to include the different user-perspectives in the democratic dialogues. Including multiple user-perspectives can ensure sensitivity to the different use situations and avoid the “false-positive” that by following a single user-perspective, the best way to facilitating data is achieved. Knowing the different users, can provide a more diverse perspective on the requirements for stimulating OGD use, and hence facilitate public managers prioritization. Previous research focuses on broad groups of users such as citizens (Safarov et al., 2017). By looking at the present findings, however, the broad categories might complicate the understanding of how to facilitate utilization. Instead, the proposed user archetypes provide a useful distinction which might surpass the contextual application of OGD.

6.3.2 Identify Overarching Principles Instead of End-goals

It has been found that different use-contexts require different usability needs (Ruijter et al., 2017). However, by focusing on the objective and end-goal of use, the publication process needs to adapt to an array of needs. Instead, if the focus is what makes data useful to a specific user, i.e., when data has use-value for different user archetypes, then the end-goal of use might change, but facilitation of a set of core principles of data-openness and quality can persist. The presented conditions with dimensions could, for example, be a starting point for distinguishing between principles to facilitate the unrestricted access to data, and the principles for how legislative and economic schemes can support the use.

Existing research suggests that if the private-sector influences the release of datasets, then only the private-sector will reap benefits from the publication (Ruijter et al., 2017). Compared to the findings in this research,

the underlying premise of the argument seems valid; if only one user archetype is consulted, it might hinder the general usability of the data. However, according to this research, the implication was found on a user-level, instead of a sector-level.

Overall, it is important to recognize that public managers cannot expect to facilitate innovation by focusing on supporting the end-goals. Instead, it is in focusing on the use situation, and combining the perspectives of the users that general principles can arise to facilitate a more objective use-value from data.

6.3.3 Focus on the Use-value of Data

Public-sector advocates who want to promote further publication should start focusing on the use-value of data, instead of abstract notions of, e.g., transparency (Bertot, Jaeger, & Grimes, 2012), accountability (Lourenço, Piotrowski, & Ingrams, 2017) and civic engagement (Conradie & Choenni, 2014). The findings demonstrated that significant benefits were not just an outcome of the OGD use but experienced in the act of utilizing data. However, these direct benefits from opening data which are seemingly not made explicit or broadly acknowledged in existing literature. It is the impacts of the use-conditions on the use-value which opens the door to understand the direct impacts of the public-sector efforts on the users. Hence, to ensure the motivation for publication of data will not be lower (SEISMONAUT, 2017), the public-sector could focus on the use-value impacts as a way to increase motivation for publication of data. Focusing on the user archetypes identified in this research would be an excellent place to start. This research found that by focusing on the use-value, it was possible to get insights into the specific ways in which data is beneficial to the user archetypes as well as the examples where user archetypes applied data to provide benefits for a customer. Existing research acknowledge that metadata is necessary for facilitating the understanding of data (Purwanto et al., 2017). The data scientists used metadata to assess the accuracy of data, and it was not clear if metadata itself was enough knowledge for the remaining archetypes, or if they too needed to know about the accuracy. Yet, it implies that there might be aspects of metadata such as accuracy which is not appropriately defined in current metadata. Nevertheless, by focusing on the use-value enabled discovering the differences in how the use-conditions could benefit different user archetypes.

Thus, the findings lead to the proposition that focusing on the use-value experienced by the user archetypes could be a first step in ensuring continued motivations for OGD publication as it facilitates an understanding of the reality of the users.

6.3.4 Design an Inclusive Infrastructure

Without data being valuable in the broadest possible realm of use-scenarios, it will hinder the innovation potential. For data to be usable it needs to be user-friendly, or manageable (F. Welle Donker & van Loenen, 2017). However, since different user archetypes vary in their skills and capabilities, what makes data publication user-friendly varies a great deal. The findings suggested that different user archetypes cared differently about the way access was provided. For example, data visualizers preferred access to services, while data scientists preferred access to raw data files. It has been suggested that processing capabilities of the infrastructure are important for users (Alexopoulos et al., 2016). Services can be considered a factor which enhances the processing capabilities of the infrastructure. However, this was not considered a useful capability by the data-scientists, while the business developers considered it an opportunity. While the case organization had skills to overcome the complexity in processing, it was evident that a mix of highly capable software developers and data scientists were necessary to ensure that data was available for innovation. The situation demonstrated the need to have access to raw data as this would make it easier to create services which were not existing. The raw data-files were considered the optimal way to ensure most possible users can access datasets.

The implication is that public managers who work with OGD publication and have any ambition to enable innovation for society at large should consider if it is possible to provide access to raw data as a standard. Once raw-data is available, different services which enhance the applicability of OGD data can be published to ensure a broader use-value for other users. However, by ensuring raw-data access and appropriate level of comprehensibility of the data, the public managers enable the entire range of user archetypes to understand and explore the datasets potential outside the original purpose set out by the public-sector. A critique in existing research is that individuals might not have the appropriate skills and technological knowledge to apply data (Susha, Grönlund, & Janssen, 2015b). The lack of skills has seemingly become an underlying reason why the government should focus on publishing easy to use services. However, the findings of this research highlighted the dangers of only publishing data as a service as it complicated the general use-value to create innovations.

Hence, public managers should reflect on which type of users they are considering in the design of OGD initiatives or infrastructures, and which users they do not include. By considering the different user archetypes, public managers would have a better basis for designing an OGD publication process which can facilitate the broadest possible use.

6.3.5 Focus on Data-transparency Before Data-quality

Practitioners should ensure that the specifications of the existing data include information about how data is produced. This point supports findings from existing research which implies that it is necessary to provide information on how data is produced, calculated or estimated (F. Welle Donker, van Loenen, & Bregt, 2016; A. Zuiderwijk et al., 2015). Data production knowledge is important especially for the data scientists since it provides insights into how accurate any measure is. For the business developers and data visualizers, this insight was used subsequently for advising customers.

The results showed that data-quality is bound to the context of its application, why the user archetypes did not know whether data was useful before they could understand, view, and work with it. The significance of this is that it challenges the dominating assumption in literature, where data-quality is considered a barrier to use (T. Jetzek, Avital, & Bjørn-Andersen, 2014a; Purwanto et al., 2017). Instead, the findings imply that data-quality is not always a hindrance to the use as long as users know how the overall data-quality is. By knowing how data is calculated, users can begin to understand the accuracy of data, and whether the way data is created suits the contextual application they need.

Achieving perfect data is an impossible goal. While existing research considers good data-quality to correlate with less, or easier processing (Lakomaa & Kallberg, 2013), this study finds that it is not always the case. If the data format delivered does not match the one that the users need, there still is a need for a great deal of processing before data is useful. Public managers cannot expect to foresee all possible use situations, and part of the benefits of data for the organizational users are, that they can have access to free and trustworthy datasets which is a sound basis for new solutions. However, if the user archetypes do not know how data is calculated or produced, they might be reluctant to apply it in innovative data solutions. Public managers should provide transparency in the data production, and allow for feedback on data-quality as this might provide novel insights into how data can become more objectively applicable.

6.4 Limitations

Like any other research, this has a set of limitations. The first relates to the subjective bias of the author. The philosophical foundation for this study was pragmatism, which is value-driven (Saunders et al., 2016). Some necessary measures were taken to lower the influences of the researcher's values: 1) A structured case-study methodology was applied, which included data-theory-model alignment (Pan & Tan, 2011). 2) A reflective diary was kept throughout the research process to keep a more objective perspective (Saunders et al., 2016). 3) Respondents were consulted about the meaning of the central analytical quotes. 4) A well-established qualitative analytical approach was adopted for the research process (Corbin & Strauss, 2015).

The intention was to create new theoretical propositions to test in future research. Yet, the categorization and analysis of the material is a result of the authors own interpretation, why it might reflect a degree of the author's subjective bias.

The second limitation is the geographical context of the study. The contextual reality of Denmark can be different from other places. Denmark has a highly developed technical infrastructure, which is the case for the publication of data as well. The technical maturity might be the reason why certain conditions in the literature were not considered highly significant in this research's findings. For example, both the technical infrastructure and skills of the users were considered as explicit use-condition in some research (T. Jetzek et al., 2013). However, the present research considered skills as embedded in the archetypes, and the archetypes discussed the technical infrastructure in terms of technical data-openness. Different dimensions of the use-conditions might be more critical in other geographical contexts where OGD maturity is lower. An example related to the legal dimension might be; if a country does not have the necessary open data policies in places, users might perceive the legal dimension as the biggest challenge, and not care about how technical data-openness is provided. Furthermore, it is possible that as the maturity of the policy aspects increases, users instead emphasize other dimensions of data-openness and data-quality.

The third limitation is about the generalizability of the findings. This research adopted a single-case study approach of a smaller commercial organization specialized in geographical data. The case study approach introduced the following limitations; 1) the abductive approach did introduce some complexities in the comparison of subjective interview accounts with quantitative insights from literature. However, as a step towards demonstrating the link between use-conditions and use-value, it was beneficial as it showed the relevance of existing measures related to the subjective user-perspectives. 2) The small size of the case organization meant that the representation of user archetypes was similarly small. 3) The case organization was a geographical data specialist. Geodata is a more developed OGD area (F. Welle Donker & van Loenen, 2017). The perceptions of users from the geodata industry might not apply to other contexts, both in terms of industry, and related to the type of data applied. Furthermore, it is possible that other sectors, or contexts could find more user archetypes than the ones suggested in this research.

The above sections presented a summary of the findings and discussed the implications for research and practice as well as the limitations of the study. The following section concludes this research by first summarising the paper, and then provide suggestions for future research.

7 Conclusion

The expected impacts of OGD are still not realized, and the public-sector is struggling with how to facilitate OGD use to reap the expected benefits of innovation with OGD. This research aimed at providing insights into the missing user-perspective, by taking an abductive approach to investigate the case of a Danish organization whose solutions involve the use of OGD. The study identified four organizational user archetypes within the organization, who differed based on how they worked with OGD. The research investigated how the user archetypes framed the influence of the use-conditions on the perceived use-value of OGD. The data was corroborated with existing research to identify points of contrast and comparison. The main contributions of the study include the distinction of the four user archetypes, and how they frame the influences of the use-conditions on their perceived use-value of OGD differently. Secondly, the research provides recommendations for research and practice about how the findings of this study can be leveraged moving forward. The present research might be a good starting point for understanding why the proposed benefits of OGD are not realized yet. However, it raises more questions than it answers. As such the following opportunities are identified for future research:

Future research needs to solidify the user-perspectives. No agreed upon user classification is adopted in OGD research. Yet, this study found that differences existed in how use-conditions facilitate the work with OGD of the user archetypes. It is necessary for OGD research to strategically move the needle forward on the user-perspective, by exploring if the presented user archetypes apply in different settings, and if there are more archetypes on an individual level. There might also be user types of importance on the organizational level. Insights into the broader organizational influences might be approached by adopting a user-perspective, through focusing on the shared frames. Future research should, furthermore, investigate the user characteristics of the archetypes to approach ways of quantifying the user archetypes in ways where they are incorporable in assessment and adoption frameworks.

Future research should compare the actions taken to facilitate use by the public-sector, with the perceived use-value of data-users. Existing research does not explain why open data cannot live up to its expectation (F. Welle Donker & van Loenen, 2017). The explanation might be approached by comparing the actual use-conditions put in place, to the impact experienced by users. The comparison can provide an enhanced understanding of which actions are effective in facilitating use, and what actions do not benefit a broader group of users. The findings in this research shed light on some general principles, but by making a comparison of the intended consequences of actions with the experienced influence, it is possible to identify further actionable insights.

Furthermore, this study took an abductive approach which relied on the knowledge and insights of the author. Thus, it might be possible to collect new and beneficial insights to extend the findings by conducting a broader review of the OGD literature through the lens of the proposed archetypes. Such a review might shed more light on how current findings around facilitating use corroborate with the archetypes.

Lastly, Future research should investigate whether the user archetypes apply to other contexts. This entails investigating other geographical contexts of similar or different maturity levels. Furthermore, it includes comparing to similar cases and cases with differences in organizational contexts, the legal nature of organizations, the industry, the organizational size, objective of utilizing OGD and finally the type of data used.

The challenges of facilitating use and realizing the potential impacts of OGD are rising with the implementation of OGD initiatives across the globe. As such, it will continue to be a concern for public managers responsible for the OGD initiatives. By identifying a set of user archetypes, this research aimed at providing a more structured understanding of the existing user perspectives, for investigating how to facilitate the use of OGD. The findings provided a step towards understanding why the promises of OGD are not yet realized.

8 References

- Abella, A., Ortiz-De-Urbina-Criado, M., & De-Pablos-Heredero, C. (2017). A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems. *Cities*, *64*, 47–53. <https://doi.org/10.1016/j.cities.2017.01.011>
- Alexopoulos, C., Loukis, E., & Charalabidis, Y. (2016). A methodology for determining the value generation mechanism and the improvement priorities of open government data systems. *Computer Science and Information Systems*, *13*(1), 237–258. <https://doi.org/10.2298/CSIS151210001A>
- Attard, J., Orlandi, F., & Auer, S. (2016a). Data driven governments: Creating value through open government data. In A. et. a. Haumeurlain (Ed.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9860 LNCS, pp. 84–110). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-53416-8_6
- Attard, J., Orlandi, F., & Auer, S. (2016b). Value Creation on Open Government Data. In *49th Hawaii International Conference on System Science* (pp. 2605–2514).
- Attard, J., Orlandi, F., & Auer, S. (2017). Exploiting the Value of Data through Data Value Networks. In *Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance - ICEGOV '17* (pp. 475–484). New York, New York, USA: ACM Press. <https://doi.org/10.1145/3047273.3047299>
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, *32*(4), 399–418. <https://doi.org/10.1016/j.giq.2015.07.006>
- Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2012). Promoting transparency and accountability through ICTs, social media, and collaborative e-government. *Transforming Government: People, Process and Policy*, *6*(1), 78–91. <https://doi.org/10.1108/17506161211214831>
- Borah, P. (2011). Conceptual Issues in Framing Theory: A Systematic Examination of a Decade's Literature. *Journal of Communication*, *61*(2), 246–263. <https://doi.org/10.1111/j.1460-2466.2011.01539.x>
- Bowman, C., & Ambrosini, V. (2000). Value Creation versus Value Capture: Towards a Coherent Deinition of Value in Strategy. *British Journal of Management*, *11*, 1–15.

- Chu, P.-Y., & Chiang, S.-Y. (2014). A Study of Establishing Evaluation Indices for Open Government Data. In *Proceedings of the 14th European Conference on eGovernment* (pp. 393–401).
- Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly*, 31(1), S10–S17. <https://doi.org/10.1016/j.giq.2014.01.003>
- Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (4.). California: SAGE Publications Ltd.
- Cornelissen, J. P., & Werner, M. D. (2014). Putting Framing in Perspective: A Review of Framing and Frame Analysis across the Management and Organizational Literature. *The Academy of Management Annals*, 8(1), 181–235. <https://doi.org/10.1080/19416520.2014.875669>
- Cranefield, J., Robertson, O., & Oliver, G. (2014). VALUE IN THE MASH: EXPLORING THE BENEFITS, BARRIERS AND ENABLERS OF OPEN DATA APPS. In *Proceedings of the European Conference on Information Systems (ECIS) 2014*. Retrieved from <http://aisel.aisnet.org/ecis2014/proceedings/track23/6>
- Davidson, E. (2006). A Technological Frames Perspective on Information Technology and Organizational Change. *The Journal of Applied Behavioral Science*, 42(1), 23–39. <https://doi.org/10.1177/0021886305285126>
- Davidson, E., & Pai, D. (2004). Making Sense of Technological Frames: Promise, Progress, and Potential. In *Information Systems Research* (pp. 473–491). Boston: Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-8095-6_26
- Davies, T. (2010). Open data, democracy and public sector reform. Retrieved from <http://www.practicalparticipation.co.uk/odi/report/>
- Eckartz, S., van den Broek, T., & Ooms, M. (2016). Open Data Innovation Capabilities: Towards a Framework of How to Innovate with Open Data. In *International Conference on Electronic Government and the Information Systems Perspective (EGOVIS)* (pp. 47–60). https://doi.org/10.1007/978-3-319-44421-5_4
- European Commission. (2015). Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources. <https://doi.org/10.2759/328101>
- Gonzalez-Zapata, F., & Heeks, R. (2015). The multiple meanings of open government data: Understanding

- different stakeholders and their perspectives. *Government Information Quarterly*, 32(4), 441–452.
<https://doi.org/10.1016/j.giq.2015.09.001>
- Herala, A., Vanhala, E., & Krri, T. (2016). Experiences About Opening Data In Private Sector: A Systematic Literature Review. In *SAI Computing Conference 2016* (pp. 715–724).
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446–456.
<https://doi.org/10.1016/j.giq.2011.01.004>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258–268.
<https://doi.org/10.1080/10580530.2012.716740>
- Jetzek, T. (2016). Managing complexity across multiple dimensions of liquid open data: The case of the Danish Basic Data Program. *Government Information Quarterly*, 33(1), 89–104.
<https://doi.org/10.1016/j.giq.2015.11.003>
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014). Data-Driven Innovation through Open Government Data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 15–16.
<https://doi.org/10.4067/S0718-18762014000200008>
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013). The Generative Mechanisms of Open Government Data. In *ECIS 2013 Proceedings* (pp. 1–12). Retrieved from <http://aisel.aisnet.org/ecis2013>
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014a). Generating Sustainable Value from Open Data in a Sharing Society. In *ICIS 2013 Proceedings - International Conference on Information Systems* (pp. 62–82). https://doi.org/10.1007/978-3-662-43459-8_5
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014b). Generating Sustainable Value from Open Data in a Sharing Society. In B. Bergvall-Kåreborn & P. A. Nielsen (Eds.), *Creating Value for All Through IT: IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2014, Aalborg, Denmark, June 2-4, 2014. Proceedings* (pp. 62–82). Berlin, Heidelberg: Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-662-43459-8_5
- Jetzek, T. H. (2015). *The sustainable value of open government data*. Retrieved from [http://openarchive.cbs.dk/bitstream/handle/10398/9183/Thorhildur Jetzek.pdf?sequence=1](http://openarchive.cbs.dk/bitstream/handle/10398/9183/Thorhildur%20Jetzek.pdf?sequence=1)

- Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly*, 30(4), 508–513. <https://doi.org/10.1016/j.giq.2013.05.012>
- Klein, H. K., & Myers, M. D. (1999). A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. *MIS Quarterly*, 23(1), 67. <https://doi.org/10.2307/249410>
- Kvale, S., & Brinkman, S. (2009). *InterViews: Learning the craft of qualitative research interviewing*.
- Lakomaa, E., & Kallberg, J. (2013). Open Data as a Foundation for Innovation: The Enabling Effect of Free Public Sector Information for Entrepreneurs. *IEEE Access*, 1, 558–563. <https://doi.org/10.1109/ACCESS.2013.2279164>
- Lepak, D. P., Smith, K. G., & Taylor, M. S. (2007). Value Creation and Value Capture: A Multilevel Perspective. *Academy of Management Review*, 32(1), 180–194. <https://doi.org/10.5465/amr.2007.23464011>
- Lourenço, R. P., Piotrowski, S., & Ingrams, A. (2017). Open data driven public accountability. *Transforming Government: People, Process and Policy*, 11(1), 42–57. <https://doi.org/10.1108/TG-12-2015-0050>
- Magalhaes, G., & Roseira, C. (2017). Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*. <https://doi.org/10.1016/j.giq.2017.08.004>
- Magalhaes, G., Roseira, C., & Manley, L. (2014). Business models for open government data. In *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance - ICEGOV 2014* (pp. 365–370). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2691195.2691273>
- Mazmanian, M. (2013). Avoiding the Trap of Constant Connectivity: When Congruent Frames Allow for Heterogeneous Practices. *Academy of Management Journal*, 56(5), 1225–1250.
- Müller, S. D., & Jensen, P. (2017). Big data in the Danish industry: application and value creation. *Business Process Management Journal*, 23(3), 645–670. <https://doi.org/10.1108/BPMJ-01-2016-0017>
- Open Knowledge Foundation. (2018). The Open Definition. Retrieved October 30, 2018, from <https://opendefinition.org/>

- Orlikowski, W. J., & Gash, D. C. (1994). Technological Frames: Making Sense of Information Technology in Organizations. *ACM Transactions on Information Systems (TOIS)*, *12*(2), 174–207.
- Pan, S. L., & Tan, B. (2011). Demystifying case research: A structured–pragmatic–situational (SPS) approach to conducting case studies. *Information and Organization*, *21*(3), 161–176. <https://doi.org/10.1016/j.infoandorg.2011.07.001>
- Pereira, G. V., Macadar, M. A., Luciano, E. M., & Testa, M. G. (2017). Delivering public value through open government data initiatives in a Smart City context. *Information Systems Frontiers*, *19*(2), 213–229. <https://doi.org/10.1007/s10796-016-9673-7>
- Purwanto, A., Janssen, M. F. W. H. A., & Zuiderwijk-van Eijk, A. (2017). Towards an Open Government Data Success Model: A case study from Indonesia. In *Proceedings of the European Conference on e-Government, ECEG (2017) Part F129463* (pp. 154–162). Academic Conferences. <https://doi.org/978-1-911218-38-8>
- Repo, A. J. (1986). The dual approach to the value of information: An appraisal of use and exchange values. *Information Processing & Management*, *22*(5), 373–383. [https://doi.org/10.1016/0306-4573\(86\)90072-5](https://doi.org/10.1016/0306-4573(86)90072-5)
- Ruijter, E., Grimmelikhuijsen, S., Hogan, M., Enzerink, S., Ojo, A., & Meijer, A. (2017). Connecting societal issues, users and data. Scenario-based design of open data platforms. *Government Information Quarterly*, *34*(3), 470–480. <https://doi.org/10.1016/j.giq.2017.06.003>
- Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, *22*(1), 1–24. <https://doi.org/10.3233/IP-160012>
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods for Business Students. Research methods for business students* (7th ed.). Pearson.
- Saxena, S., & Janssen, M. (2017). Examining open government data (OGD) usage in India through UTAUT framework. *Foresight*, *19*(4), 421–436. <https://doi.org/10.1108/FS-02-2017-0003>
- Sayogo, D. S., & Pardo, T. A. (2013). Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data. *Government Information Quarterly*, *30*(SUPPL. 1), S19–S31. <https://doi.org/10.1016/j.giq.2012.06.011>

- SEISMONAUT. (2017). *Når myndigheder udstiller offentlige data - målgruppe- og behovsanalyse i forbindelse med udstilling af data i myndigheder.*
- Smith, G., & Sandberg, J. (2018). Barriers to innovating with open government data: Exploring experiences across service phases and user types. *Information Polity*, 23(3), 249–265. <https://doi.org/10.3233/IP-170045>
- Sun, T. Q., & Medaglia, R. (2018). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, (June), 1–16. <https://doi.org/10.1016/j.giq.2018.09.008>
- Susha, I., Grönlund, Å., & Janssen, M. (2015a). Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries. *Information Polity*, 20(1), 19–34. <https://doi.org/10.3233/IP-150353>
- Susha, I., Grönlund, Å., & Janssen, M. (2015b). Organizational measures to stimulate user engagement with open data. *Transforming Government: People, Process and Policy*, 9(2), 181–206. <https://doi.org/10.1108/TG-05-2014-0016>
- Tavory, I., & Timmermans, S. (2014). *Abductive Analysis: Theorizing Qualitative Research.* <https://doi.org/10.7208/chicago/9780226180311.001.0001>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425–478.
- Viscusi, G., Spahiu, B., Maurino, A., & Batini, C. (2014). Compliance with open government data policies: An empirical assessment of Italian local public administrations. *Information Polity*, 19(3,4), 263–275. <https://doi.org/10.3233/IP-140338>
- Welle Donker, F. M. (2016). *From Access to Re-use: A user's perspective on public sector availability. Architecture and the built environment.*
- Welle Donker, F., & van Loenen, B. (2017). How to assess the success of the open data ecosystem? *International Journal of Digital Earth*, 10(3), 284–306. <https://doi.org/10.1080/17538947.2016.1224938>
- Welle Donker, F., van Loenen, B., & Bregt, A. (2016). Open Data and Beyond. *ISPRS International Journal of Geo-Information*, 5(4), 48. <https://doi.org/10.3390/ijgi5040048>

- Worthy, B. (2015). THE IMPACT OF OPEN DATA IN THE UK: COMPLEX, UNPREDICTABLE, AND POLITICAL. *Public Administration*, 93(3), 788–805. <https://doi.org/10.1111/padm.12166>
- Yin, R. K. (2018). *Case Study Research and Applications : design and methods* (6th ed.). Los Angeles: SAGE.
- Zeti, C. florin, Bartalis, A. M., Dobre, A. C., & Pitu, C. I. (2017). The role of data opening in creating added value. *Annals of the „Constantin BrâNcuși” University of TâRgu Jiu, Economy Series*, (6), 177–182.
- Zuiderwijk, A., Janssen, M., & Dwivedi, Y. K. (2015). Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Government Information Quarterly*, 32(4), 429–440. <https://doi.org/10.1016/j.giq.2015.09.005>
- Zuiderwijk, A., Janssen, M., & Susha, I. (2016). Improving the speed and ease of open data use through metadata, interaction mechanisms, and quality indicators. *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), 116–146. <https://doi.org/10.1080/10919392.2015.1125180>
- Zuiderwijk, A., Susha, I., Charalabidis, Y., Parycek, P., & Janssen, M. (2015). Open data disclosure and use: critical factors from a case study. In P. Parycek & N. Edelman (Eds.), *CeDEM 2015: Proceedings of the International Conference for E-Democracy and Open Government 2015* (pp. 197–208). Krems: Edition Donau-Universität Krems N.B.

9 Appendices

9.1 Appendix I: Interview Guide

Pre-interview:

- Inform the respondent of the format of the format of the interview, as well as what it will be used for.
- They are informed that I have an NDA, and the information published will be approved by the CEO.
- The respondent will receive all translated quotes for comments before utilized in the report.

Consent: Does the interviewee agree to be recorded?

Part 1: Background

Question 1: What is your background and role in the organization?

Question 2: What open government data is part of your daily work?

Part 2: Expected use-value

Question 3: Why are open government data interesting in relation to your work?

Question 4: Why are open public data useful in your work?

Part 3: Realized use-value

Subjective Value-In-Use; Role of information (data) on a knowledge task.

Question 5: How does open data play a role in your work-tasks or projects?

Question 6: How do you work with open government data?

Question 7: How are open data part of the projects or tasks you work with?

Question 8: How do you decide what data are useful or relevant in relation to the projects or tasks you have?

Question 9: How would you describe the usability of OGD in relation to the projects you are working with?

9.2 Appendix II: Data categorization

Conditions	Dimensions	Exemplary Quote
Data-openness	Legal	<p>“It [open data] is about someone making it available with a relatively free license. [...] For me, it is not open data if they say you can use it, but not for commercial purposes” [R5]</p> <p>“but there are elements of the infrastructure which are not bound or ensured by law” [R10]</p> <p>“License aspects are about what you can use data for. It is a demand for the organizations use that data can be used commercially and that you can earn money on solutions using the data.” [R1 - note]</p>
	Economic	<p>“Data is not always as accessible or open in reality. They can be a bit more complicated to get a hold of. [...], e.g., BBR-data which has also been bought free [prior to the basic data program]. But in reality, it costs a lot of money to get them” [R4]</p> <p>” To go out and say that you can buy something, that is a barrier to start with. Then rather say here is something for free, then there are a lot of people who can use it” [R2]</p>

		<p>“It became affordable to make the mistakes which are traditionally made in a business. [...] we need perhaps 20-50 bad ideas, and then we get a good one. Let us say each idea would cost just 100.000, right. That is expensive in data purchase to arrive at the last idea where you perhaps can earn a million” [R5]</p> <p>“then it barely matters that data is free, if the calculation software costs a bloody fortune” [R6]</p>
	Technical	<p>“We want data. And this is a big weakness for the data distributor. [...] it is not data they are distributing [...] they distribute services on top of data. [...] Which requires that they’ve considered all kinds of services that one could wish to create on top of the data. If they haven’t, well, then it is difficult to build the service, because it is difficult to access the data behind [the existing services]” [R6]</p> <p>“When it [data publication] becomes service oriented. Then it is up to the creator of service to define how I will get the addresses out of the system” [R11]</p> <p>“We collect all data in-house, from CVR and from local plan data, and indexes them ourselves [...] our search product can aside from searching in the existing public search registers, search in the local plans and CVR” [R3]</p>
Data-quality	Functionality	<p>“it is necessary that data is nationally covering to create interesting solutions unless the customer e.g. is a municipality or the who needs help to improve their processes.” [R1 - note]</p>

		<p>“They [the datasets] are connected on a logical level, but not on a physical level. [...] Physical connection is that I cannot delete a building [in the dataset] if someone lives in it. [...] where if datasets are [logically connected], then I can delete the building even if someone lives there” [R7]</p> <p>“OSM [open street map] contains a lot of different information compared to Geo Denmark. e.g. information about parks, attractions etc. these data can be very valuable” [R8]</p>
	Comprehensibility	<p>“then I, of course, think about whether there is something in these data we can use in some way [...] to create value for others” [R1].</p> <p>“As long as you know the quality level of data, then they are very useful. [...] data should be published with [...] some detailed specification which states what you can expect to be correct in this dataset, and what can you expect is not accurate” [R7]</p> <p>“it can be very difficult to get an overview over [...] all these metadata which exist for it [the data set]. It is quite important to know how they are registered when you need to use them” [R8]</p> <p>“Metadata is such a boring word, but it is really important in understanding the data-quality. So, I think [...] it is very relevant also when you're advising a customer, that you know what you are providing advice on” [R9]</p>

	Sustainability	<p>“It would be clear-cut [to display the road net] because then there would be an authoritative register that contained all the traffic in Denmark” [R7]</p> <p>“[data] should have a certain quality for them [the public-sector] to use it. Where you as a private [organization] might be able to compromise” [R9]</p> <p>“They actually come from an authoritative source [...] there is not a crown stamped on, but something similar. There is a somewhat higher trust in the [data] since these are being used to govern society.”. [R12]</p>
	Accuracy	<p>“If there is doubt about exactly which road and house number and zip code it is [the data field], then it can be difficult to e.g. geocode it [data], so you can display it on a map” [R4]</p> <p>“When there are mistakes in some of these data we pull in [...] for example] suddenly there are two building 1 on the same property. That is a mistake in data which should be corrected. But it implies that they systems using these data are not doing too good” [R11]</p> <p>“If we get an erroneous data extract that one time a week we run it [the script] and find out that it [data] is erroneous. Then we would rather keep the old data, which are 2 weeks old, instead of collecting data which is one week old and broken” [R12]</p>