TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

John-Philipp Vogt 201676IVGM

ANALYSIS OF OPEN GOVERNMENT DATA USAGE IN HAMBURG

Master's Thesis

Supervisor
Eric Blake Jackson
MSc

Tallinn 2022

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia Teaduskond

John-Philipp Vogt 201676IVGM

AVATUD VALITSEMISE ANDMETE ANALÜÜS HAMBURGI NÄITEL

Magistritöö

Juhendaja
Eric Blake Jackson
MSc

Tallinn 2022

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: John-Philipp Vogt .....
(signature)

Date: December 16, 2022

Abstract

Throughout the past decade, national and local governments across the globe underwent increased efforts in order to generate and release the data they use and work with – coined as OGD – in the pursuit of more citizen participation and trust, enhanced transparency and accountability, bolstered economic and scientific added value and more. While a diverse range of possible effects through the release are hypothesised, most actual effects remain to be proven. A possible reason for this lack of proof seems to be the lack of knowledge about why particular users engage with OGD, what they actually do with it in the pursuit of which effects.

This thesis' main objective is to explore this particular user attitude and behaviour in the case and context of the Free and Hanseatic City of Hamburg, in order to give an initial insight into the situation based on first-hand empirical data collected.

In this study, the qualitative analysis method of reflexive thematic analysis is used to analyse 9 semi-structured interviews in order to give insight into motivation, interaction, and desired effects in connection with OGD use, which resulted in 4 themes.

The thematic analysis revealed that OGD re-use, in the context of the FHH, seems to involve a highly diffuse, and heterogeneous process which involves various data from various sources in various different use cases. While this is the case, motivations for use seem to involve multiple facets, however transparency as a fundament for trust, participation, inclusion were identified as most prevalent, along with practicability and inherent value perception of OGD. Desired effects mostly aligned with motivations, however the data showed that users find that OGD use will or should contribute to a better informed citizenry and thus empower the public.

Keywords: open government, open government data, data usage, data utilisation, users, thematic analysis, Hamburg

The thesis is written in English and contains 64 pages of text (88 pages in total), 6 chapters, 5 figures and 6 tables.

List of abbreviations and terms

OG	Open Government
OD	Open Data
OGP	Open Government Partnership
OGD	Open Government Data
FHH	Free and Hanseatic City of Hamburg
IS	Information System
TAM	Technology Acceptance Model
UTAUT	Unified Theory of Acceptance and Use of Technology
TOE	Technology-Organisation-Environment
FAIR	Findability, accesibility, interoperability, and reusability
CKAN	Comprehensive Knowledge Archive Network
DCAT	Data Catalogue Vocabulary
GDP	Gross Domestic Product
PU	Perceived Usefulness
PEU	Perceived Ease of Use
SU	Subjective Norm
PE	Performance Expectancy
EE	Effort Expectancy
SI	Social Influence
FC	Facilitating Conditions
RSS	Resource Description Framework
API	Application Programming Interface
MRQ	Main Research Question
SRQ	Sub Research Question
OSINT	Open Source Intelligence
CAQDAS	Computer Assisted Qualitative Analysis Software
OSINT	Open Source Intelligence
ODI	Open Data Institue
HmbTG	Hamburg Transparency Law

Table of Contents

List of Figures

List of Tables

1 Introduction

With the prospect of more citizen participation and trust, enhanced transparency and accountability [1], [2], bolstered economic and scientific added value [3]–[5] and more, governments all over the globe are committing themselves to "[i]ncrease the availability of information about governmental activities" [6]. This phenomenon is regularly labelled with the terms *Open Government* and *Open Government Data*.

Since 2011, over 70 national and 100 local governments have endorsed the Open Government Declaration of the Open Government Partnership [6]–[8]. By doing so, they not just commit themselves to fundamental principles of human rights and good governance [6]. They also "acknowledge that people all around the world are demanding more openness in government. They are calling for greater civic participation in public affairs, and seeking ways to make their governments more transparent, responsive, accountable, and effective" [6].

While this quote does not sum the declaration in it's entirety, it illustrates the core values bound to joining the partnership precisely. The membership of all anglo-saxon nations – Australia and New Zealand, the U.S.A., Canada, the U.K. and Ireland – as well as all economic and political heavy-weights from Europe and Latin America – Germany, France, Italy, Spain, the Netherlands, Brazil, Mexico, Argentina, Colombia, Chile – indicates a practically global commitment to values such as civic participation, democratic accountability and governmental transparency.

As the example of the OGP makes the case for a global phenomenon, it is not the only wide-spanning, and at least regional, declaration of political intent. The European Union found in 2003 already that the harmonisation of dissemination and re-use of public sector information plays an important part in shaping the European internal market. For almost 20 years it has promoted and revised it's Open Data directive. It found, and still does today, that Europe evolves "towards an information and knowledge society" and that the data generated and held by the public sector is a "fundamental instrument for extending the right to knowledge, which is a basic principle of democracy" [9], [10]. This early insight acknowledged, and included already, the importance and impact of re-using and sharing research data.

And the list goes on: In 2009, former U.S. President Barack Obama, signed his Open Government Directive. And in 2013 the G8 proposed and committed to its Open Data Charter. These documents emphasise values similar to the OGP, and to some extent to those of the E.U.: Transparency, participation, collaboration, inclusion, accountability, sustainability and value for democratic and economic development [11], [12].

While some significant nations remain excluded from the previous examples, such as China, India, Japan, and Russia, it is fair to say that a global political interest in the concepts of Open Government and Open Government Data is on the rise.

Through an academic lens and at the highest level of abstraction, both of these terms encompass notions around transparency, participation, and collaboration in relation with the public's interaction with their governments and administrations mediated through the use of ICTs. They are, in fact, tightly interrelated and parts of the same phenomenon. However, due to issues regarding inherent complexity, in parts vague definitions and convolution of meaning, both terms require some deeper discussion[1] [13].

Along with the political investments into said phenomenon, unsurprisingly, a wide scientific interest took hold. The years around 2010 are consequently demarcated as a starting point for intensified research and political engagement. Points of departure are former U.S. President Barack Obama's Open Government Directive and the founding of the Open Government Partnership [1], [14]–[18]. Since then, the scientific literature on the topic has increased dramatically [13], [14], but is seen at large as still quite young and emerging [1], [19], and as quite disparate and immature [13], [20], [21]. The numerous systematic reviews in a relatively short amount of time attest to the emergence and high activity of the field. But it also shows the disparity through the demonstration of a high diversity of viewpoints, subtopics, origins of interested scholars and their disciplines, and attribution of importance [1], [13], [14], [22]–[25].

Be that as it may: The expected, or theorised, effects that this global phenomenon is purported to cause are many things, except contested. There is a remarkable range of positive outcomes, generally held in the literature, which *should* materialise – in one form or another. The aforementioned increased "availability of information about governmental activities" is supposed to lead to more accountability, trust, and satisfaction, a reduction in corruption, increases in social cohesion, better deliberation, feedback and participation mechanisms, including participation in policy formation, spurred economic development through co-created public services and applications, improvements or optimisations of public sector processes and an increase in complex problem solving capabilities [1]–[4],

[1] See 3.1

[13], [14], [24], [26]–[28].

As intuitive as many of those proposals may seem however, most of these are accompanied by several negative implications and tangible problems both in theory and in practice. To the point: Many of the theorised effects have yet to be empirically validated [13], [19], [24], [26], [29], [30]. The overarching issue here, for both OG and OGD research and practice, is a certain *imbalance between hypothesised or desired, and measured and materialised outcomes of efforts* [13], [24], [29], [30].

In other words: Expectations are not met. Although considerable academic and political resources are devoted to OGD, neither camp could clearly verify if and why the desired effects actually occur. In practical terms this means that, as long as neither camp is able to verify the causal relationship between efforts and effects, a reasonable argument why those efforts should be sustained cannot easily be formulated.

This overarching issue has various facets which could be explored. While this paper acknowledges that key issue as such, it focuses on one facet that is seemingly contributing to this inability [24]: By default – that is because of the anonymous nature of OGD use for the sake of privacy protection, at least in the case of the FHH [31], it is illegal to passively gather usage data on OGD users merely through their interaction with a respective OGD Web site. As a result, decision-makers of OGD policies at large know very little, if anything, about concrete users and their patterns of OGD usage [19], [24], [32], [33], which is problematic for obvious reasons. Unless one side reaches out to the other and a two-way communication channel of some kind can be created and sustained, this issue cannot be resolved.

Further, this facet has another interesting characteristic which exacerbates the problem: The usage of OGD by potential users seems problematically low. In the literature, this circumstance is referred to as the under-utilisation of published governmental data [24], [34]–[36]. This paper is concerned with the problem of under-utilisation of OGD on the one hand and the problem of the lack of knowledge around usage patterns of OGD users on the other. It posits that these issues are not just problems in and of themselves, but are also negatively affecting the overarching academic issue.

Therefore, the central research problem addressed in this paper is:

How and why do users interact with OGD in the FHH?

1.1 Objective and Overview

Considering the problematic situation outlined above against the background that the re-use of published OGD is anticipated to lead to various positive outcomes it is crucial to understand how this re-use is actually playing out in reality. As this paper will show, there are various investigations into this problem to shine light on the issue. But however much is known about a specific context defined as, for example, two countries, one agency, or one particular OGDI, it remains questionable how much value those findings have for any one political entity that administers it's own OGD which does not share the context. After all, it is the political entity that administers, which must make the decisions regarding the configuration of strategic and operational aspects of releasing it's OGD. Therefore, the most valuable knowledge this political entity can have to make decisions regarding its OGD policies comes from within its own context, from its immediate users of its own OGD.

The immediate objective of this study is to provide an analysis of empirical data, gathered first-hand, from within its own context, in order to give an evidence based, credible and insightful outline of how and why users of OGD interact with it. The results are intended to present an outline of actual re-use, not a positivistic statistical or probabilistic generalisation of user behaviour and expectations. First and foremost the results aim to inform users and decision-makers of OGD in the FHH about the reality of re-use. While this might simply be interesting to both parties, the decision-makers ought to at least use this information as data points to look for more, and in the best case, to support the development of OGD policies.

The extended objective is to use those findings in order to evaluate the validity of current theory used to explain interaction with OGD. That is the findings strive to contribute to analytical generalisations to theory explaining interaction with OGD. These results primarily address researchers of OGD.

In order to achieve this objective this paper continues as follows:
This introductory chapter continues with the authors motivational background for this research project. Then it proves the need for this research and articulates the precise research questions to be addressed in the study.

Chapter 2 justifies and explains the overall research design and strategy. Further it details the processes of data generation and data analysis in order to comply with the given research strategy and to establish credibility of the research. It shows and documents in detail why the view is held and the approach is chosen, and offers arguments to possible

alternative ways of dealing with this topic.

Chapter 3 reviews the literature on the topics of OG, OGD and the utilisation of OGD. As this study's background, it discusses and defines those terms and lays the groundwork to understand the key concept of this study – OGD utilization. It then draws from previous findings of related works on the issue to present the conceptual frame in which the data generated are viewed.

Chapter 4 describes the case of OGD and it's use and users in the FHH.

Chapter 5 presents the thematic analysis of the research project, and discusses all themes individually.

Chapter 6 concludes the paper by summarising and presenting all relevant conclusions and finally providing comments on limitations and future research prospects.

1.2 Motivation

It is the author's conviction, that better informed societies are enabled to make better informed decisions. Open Government Data offers to provide that information as a basis to make better informed decisions on. Given the appropriate access and political instruments to participate, it seems a matter of effort, not a matter of possibility to hold the relevant form of governance accountable. While time and money are notoriously short in supply, there are some questionable democratic developments to be observed, both in Europe and elsewhere. Thus it seems ever more important to try and use those resource one has access to, to try and support a free, equitable and empowered society. Especially those resource which seem both promising and under-exploited – the worst waste is that of potential.

Admittedly, the situation is a different one for different nations, but progress needs to be actively stimulated everywhere when one demands accountability. Thus, the author is taking it up for him to address this issue in a context where he believes he can make some sort of impact.

The goal of this paper is, as stated, to present some sort of empirically validated answer to the MRQ, which reflect both an academic concern as it does from the practitioners' standpoint. Second to that, the motivation is two-fold. First, to practically contribute to the current deficiencies which are obvious and strikingly hindering in the context of the FHH and elsewhere, as pointed to in the introductory remarks. The results of this study hopefully contribute to draw a picture of those users who actively engaged in the use of

OGD in the FHH. Those findings ought to be of value, firstly, to somewhat perceive the user-persona and places to look for more, in the pursuit of tailoring strategic objectives for the OGD policy area. Secondly, the results ought to be able to place a marker in what way the sentiment is pointing in terms of the value-proposition which is pursued within the topic of OGD. Placing the figurative marker should indicate a direction in which to look when one tries to find the kind and place of effects that should be expected.

As tangible deliverables this project aims to provide propositions in reference to the posed research questions in order to begin or continue the exploration in the same or other places, and to be in a position to ask more precise questions with regards to what actually happens with OGD and why.

1.3 Research Gap

In their literature review on utilisation of OGD, the authors of [24] point out that the key issues relating to OGD are bound to the demand-side, i.e. the user and his or her use of data, rather than the supply-side, i.e. the publishing of data. Meaning that the literature can not give proper indication but mostly makes assumptions about *who* really uses OGD and *how* this is supposed to happen, and as a corollary, out of which *motivation* and in the pursuit of which *effects*. Put bluntly, [19] find that decision-makers may simply not know who uses their OGD. To quote a finding from [36] regarding engagement with OGD users from the perspective of a U.S federal department head: "I have no idea who is using our open data sets". [30] find that OGDIs do not to address their users appropriately in terms of skills, profiles, support and interaction possibilities. These findings show that OGD decision-makers are not aware of their users' capabilities and demands, if they are aware of them at all.

In a recent, comprehensive review on citizen engagement with open government data [25] and with regards to motivation (i.e. driving and inhibiting factors of citizen engagement) the authors find that the literature mostly does not focus explicitly on neither driving nor inhibiting factors for OGD engagement, which seem to be highly context-dependent and thus yields sometimes inconsistent results. Therefore, the authors have "excerpted empirical data that indicates the *demand, needs,* and *interests* of citizens, *purpose* of using OGD, *motivations* for participating in OGD engagement" and "data that indicate the challenges, difficulties, problems, impediments, and barriers felt and experienced by citizens before engaging with OGD or during the OGD engagement" [25, p. 12, 14] (original emphasis). Both [24] and [25] reveal that the individual users' drivers and inhibitors have not been explicitly studied along with the types of usage that users pursue. [24] conclude that several linkages deserve more attention. Specifically the link between

who uses OGD and 1) how this particular user will do so, and 2) what type of effect is aimed for. Thus, more effort in addressing individual OGD usage is warranted.

While [25] point out that most of their reviewed studies investigate OGD in European countries, including Germany, to the author's best knowledge, there has not been any academic work done in the context of the FHH or any other political entity that addresses this particular issue in this particular scope (neither in English nor in German).

Finally, [29] lament that the lack of frameworks and quality evidence, mismeasurement, and time-lags between efforts and effects are contributing to the difficulty to accurately establish connections between OGD utilisation and outcomes. These findings indicate that – next to the unknowns about the users in particular – the effects that are supposed to materialise, are not just difficult to measure, but even difficult to discover and pin-point. And this does not even address the issue of causal relationships. With more defined evidence of where effects are desired, and what processes are employed in order to do so, this difficulty might be alleviated from.

While literature reviews as meta-analyses exist (e.g.: [24], [25]) those only provide assumptions or generalisations serving as hints to characteristics of a complex phenomenon playing out in differing legal, cultural, social, technological and thematic contexts. Those findings can not hold as reliable evidence for the actual use in a defined context as there are too many possibly important factors that are uncontrolled. Thus, this paper posits the research gap as the inability to appropriately shed light accurately on the individual users' profiles given a particular context. Their demands and motives to utilise OGD, their actual interaction with OGD, and the anticipated outcomes of their interaction. It sets out to explore those variables in defined a context (FHH) while focusing on in OGD interested individuals who voluntarily self-engage in value-creation through OGD. This lack of knowledge about OGD utilisation then, is contributing also to the literature's difficulty to reliably find and measure actual outcomes of OGD efforts.

1.4 Research Questions

Based on the research problem and gap, the necessary MRQ to be addressed is the following:

MRQ: *How and why do users interact with OGD in the FHH?*

This main research question encapsulates the central research problem outlined above.

The wording aims for users who are, or have been, actively engaged in OGD and are not only part of an assumed user group. The MRQ is decomposed into four SRQs that reflect the identified lack of knowledge both from the academic standpoint and the practitioner's vantagepoint. Precisely, the SRQs define this paper's interpretation of *interaction* in the MRQ, which will be answered collectively by the following SRQs.

The first

SRQ1: *What are the users' motives behind their OGD utilisation?*

addresses motives explicitly. Thus *why* actual OGD users are driven to even turn to OGD in the first place. It seeks what pulls individuals toward the use of OGD. Next to the why question, this SRQ also has bearing on *who* uses OGD in the FHH. Meaning, current research [24], [25] suggests to stratify users according to occupation (sometimes referred to as capabilities) such as journalist, researcher, citizen or developer; Or roles such as developer, deliberator, or designer. Based on the assumptions with regards to who uses OGD, a valid question would seem if it is even appropriate to use those labels to cluster user groups.

The second and third SRQs

SRQ2: *How can the actual usage be described?*
SRQ3: *What skills and tools are employed?*

addresses the point of *how* the actual work with the OGD looks like. It aims to uncover what kind of knowledge-creation is pursued by a given user, and what skills and instruments are employed in the process. In other words, what enables utilisation in the first place, which could also be seen as constraints. These answers have bearing on the demands users have with regards to working with OGD, on possibly lacking functionalities of the Web site they retrieve the data from and possibly on quality characteristics of the data they work with.

Finally, the fourth

SRQ4: What are the users' desired effects of their OGD utilisation?

This question addresses if and if yes, which effects are desired by the user. The purpose of this SRQ is to address the issue of where one ought to expect to find effects of OGD use,

and to see if motives align with desired effects. Further, it aims to explore what specific types of effects are pursued by which kind of user and what this particular user employs to pursue them.

Taken together, the answers to those questions should be sufficient to draw a sketch of a given user's profile, including motivation for engagement, skills necessary for knowledge-creation and effects desired with regards to their OGD use and OGD in a bigger picture.

2 Research Design

At the outset of this research project the most appropriate strategy to find answers to the posed SRQs was perceived to be a qualitative survey. The initial proposal presumed to be able to obtain empirical data through interviews to find generally applicable answers to the question of how the user's interaction with OGD can be described. The aim was to find converging patterns through thematic analyses of the obtained data. However, after the collection of the first few interview opportunities, several discussions with possible samples and as the literature review and general research progressed it became clear that this strategy could not be pursued any further. A survey strategy's core advantages are to generate much data with little effort within a rather short amount of time. Further they lend themselves to statistical generalisation and exploratory endeavours. However, what they typically lack in capability is to provide in-depth analysis and any causal inferences. Finally, once set up, they are not flexible for adaption [37].

It slowly became clear that the sample size necessary for making any sort of statistical generalisation was not going to be obtained, which in hindsight, seems obvious. After all, one of the fundamental prompts for this study is the fact that very little if anything is known about actual users of OGD. This lead to a dilemma in which either one of the two defining characteristics according to [37] of a survey had to be departed from. Either attempting to be satisfied with a smaller sample size and thereby sacrificing statistical generalisability or departing from a standardised and systematic way of generating and analysing samples by adapting the procedure and criteria in which samples are obtained halfway through the project in an effort to increase the sample size.

After careful consideration a case study research strategy has been determined and adopted.

2.1 Research Strategy

According to [38] a case study research strategy is particularly desirable when attempting to expand in-depth on complex social phenomena playing out in contemporary contexts, and where the behaviour cannot be sufficiently controlled. Further, especially when it comes to "how and why" questions a case study research ought to be considered. Lastly, a case study research does not focus on statistical generalisation of a population or universe

but on analytical generalisation with respect to theoretical approaches or explanations of phenomena. As [38] points out that there is no set-in-stone approach to do case study research, there are some core principles, discussed below, which ought to be respected for a study of this scope.

This research project is fundamentally of exploratory nature and therefore adopts the aims and characteristics of an exploratory or descriptive case study according to [38]. Regarding the prerequisites for selecting the case study research approach in principle, the emergence of those principles were observed during the research project. OGD use is indeed a rather complex, multi-faceted social phenomenon which is on the bleeding edge of current research, and not really dating back further than ten years. It demands to be studied in depth in a real-life context, and as it turns out, generalisations around OGD will usually lead to misleading or inadequate results. While this behaviour can, in principle occur anywhere, the setting aimed for in this study usually happens, or has happened, in the anonymity and privacy of the home, or within a civil association or professional context. In this respect the behaviour is beyond any control.

[38] claims that five components are of special importance when designing a case study. The first component, the case study's questions, have been discussed in 1.4. The second component, the propositions it sets out, do not play a major role in exploratory research designs. The third component, the case, is defined and bounded in chapter 4. The fourth and fifths components, the logic linking the data to the purpose of the study and the criteria for interpreting the findings will be discussed next.

2.2 Methodology

Regarding methodological considerations, case study research design offers high flexibility in terms of available and specific data generation and analysis methods. One particular aspect that is important for case studies is the involvement of multiple sources of data [38]. The following describes and explains both generation and analysis procedures for all kinds of data used.

2.2.1 Data Generation

As 1 made clear, the focal point of this study is revolving around the users interaction with OGD. Parts of this interaction are stipulated as motives as drivers (SRQ1). Those are underlying personal or otherwise motives fundamental to *why* individuals engage with OGD in the first place. Second are characteristics of interaction and enabling or

facilitating factors in forms of soft and hard skills, and tools (SRQ2), thus pointing at *how* the interaction with OGD looks like and what components it contains. Lastly, it is the question of what superordinate aims and effects are expected or pursued due to the interaction with OGD. As all of those components aim at very personal, at times emotional topics regarding ontology and ideology. And as humans just simply do interact with OGD the most plausible instrument to generate those data are interviews. Therefore the primary mean of generating data are semi-structured interviews. This type of interviewing lends itself to explore feelings and sentiments, and topic in depth and detail. Further, semi-structured interviews allow the interviewer to pursue interesting tangents which may appear leading to a richer perspective on the topic [37]. However, the weakness in that approach lies in the potential bias due to poorly articulated questions, resulting in response and reflexivity bias. In order to mitigate those biases, to the degree which it is possible, the interview questions are closely aligned to the verified and analysed research questions. The general approach and the actual question examples can be revisited in the appendix. The general layout of the interviews was structured into three thematic sections as follow:

1. Motivations and drivers for the engagement (SRQ1),
2. Daily Business - how the actual interaction looks (SRQ2/3) and
3. Pursued or expected effects through the interaction or use of OGD (SRQ4).

To each interviewee the structure was introduced as such at one point. Either during the set-up of the interview or immediately before its beginning. The selection of the candidates did not exclude any user or user group by default. The selection criteria originally followed non-probabilistic snowball-sampling [37][1] which had it's zero-point at the main authority for the relevant OGD portal in the FHH. From there on the recommendation for further interview candidates followed and were pursued accordingly. Effectively this means the last request for each interviewee was to suggest further people who would be able and willing to talk about the topic. As discussed in the beginning of 2 this did not lead to enough samples (i.e. interview candidates) so the mode of discovery was enhanced through OSINT. Namely through generic Web searches via ordinary search engines, searching through related contacts and posting activities on social media (mainly Twitter and LinkedIn). Finally, through posts on Reddit and use of channels on Discord the search realm was wide enough to obtain enough sources of data.

To ensure maximum comfort for the interviewees during the interviews, it was agreed with each interviewee to keep the interviews anonymous as in to obfuscate any person-

[1] [38] does not recommend to use this language in case studies due to the misconception that the case study is aiming to generate universal, statistically relevant generalisations. This is not the case here. This example is a remnant from the previously employed research strategy discussed in the beginning of 2.

ally identifiable information. However, all interviews were recorded. The subsequent transcription of the interviews was realised with OpenAI's "Whisper" [39] which is an open-source, general purpose speech recognition and translation model. Whisper transcribed all interviews into text in the original language, German. Then the author manually revised and cleaned for word errors, punctuation, grammar, mannerism, interjections and so on, to arrive at workable, semantically correct German transcripts. All interviews have been analysed in their original language, German. Then, LibreTranslate [40] was used to translate all transcripts into the English language in order to have English transcripts at disposal. LibreTranslate is a free and open source machine translation API offering text translation among 56 languages. Then the transcripts were revised again to ensure proper semantics. Personally identifying data have been removed. All operations with both tools were conducted locally and offline. The full list of interviewees can be found in the appendix, in table 6.

Due to the nature of the research problem, that is the general ignorance of who actually uses OGD in specific contexts, there does not seem to be any other accessible way to generate a more primary data in order to triangulate findings. Plausible ones could be participant or event observation insofar there is access to a specific event, or the opportunity to observe users during their interaction, which were not presented in this study's context. For this reason, secondary data are used to supplement and relate findings, which stem on the one hand from the academic literature in form of a thorough literature review in chapter 3. From it, a wide range of preceding findings on the matter of OGD utilisation and engagement, with special respect to the research questions, is presented and discussed. This review culminates in a rich framework to illustrate the authors approach to the topic, and to provide empirical and theoretical foundations to base the analysis on. On the other hand they stem from archival and legal data in the form of legislation, protocol documentation from committee meetings and plenary sessions of the Hamburg parliament. Notably also, one comprehensive evaluation report has been analysed and is used to supplement the primary data. Those are found in its own chapter, chapter 4, as it directly and exclusively discusses the context of the FHH.

2.2.2 Data Analysis

The data generation method resulted in a large corpus of textual data derived directly from actual users of OGD, with direct bearing on the issues delineated in the introductory chapter 1. As the literature review, and the from it derived research questions and conceptual frameworks, preceded the conceptualisation of the interview structure, it is clear that a firm conceptual picture of OGD utilisation and engagement was present already before the data generation and analysis. Considering the exploratory ambition of this study, the decision

regarding the choice of the data analysis instrument must take into account the qualitative type of data to be analysed, the preconceived notions of what ought to be observed in those data, as well as the necessary freedom to explore and uncover new, perhaps alternative, or rival explanations of the phenomena under study. On the basis of those considerations, the method of reflexive thematic analysis based on [41], [42] is chosen.

Properly conducted, thematic analyses can give a coherent, rich and detailed account of complex phenomena through themes derived from qualitative data through the researcher's identification, analysis and reporting. While sound guidelines exist (i.e. [42]), there is a degree of freedom in the design of this instrument. However, some critical aspects have to be considered in order to come to the promised outcomes. Most importantly: The role of the researcher him or herself; the role of theory; and the role of the dataset [41].

From an interpretivist standpoint common in qualitative analyses, it is key to realise that *I* as a researcher, play an active role in the production of data, findings and conclusions. This circumstance precludes the ambition to find universally valid laws or even patterns in order to discover natural constants of the phenomenon. The focus thus, is on the exploration and the understanding of a phenomenon under study in a given context without being able to know everything about it. That is, the observer is bound to influence the measurement, as it were. This consideration includes the possibility of conflicting or contradictory results, and entails a certain lack of objectivity.

Regarding the use of theory, the author acknowledges, as highlighted above, there is a strong preconceived picture of the phenomenon of OG, OGD, and OGD utilisation, albeit it being based on authoritative and/or peer-reviewed findings from contemporary academic literature and relevant global organisations. This inevitably forms the notions through which the issue is viewed and this circumstance is acknowledged, even explicitly formulated in section 3.4. From that fact, derived is the deductive approach to the analysis. Which is rather theory-driven as frameworks for explanatory purposes are employed to drive the analytic process. For the thematic analysis this means that a certain set of initial conditions are defined, which serve as the units of analysis – the initial codes. Again, these are derived from the literature and are presented in chapter 5.

Finally, the role of the data set is to serve as a basis for the thematic analysis. In the frame of this study, it serves to provide a detailed account of mainly three particular aspects: Why users engage in OGD, how they do it, and what, if any, end goals are in mind while doing so. While these aspects might be explicitly embedded in the data, this analysis does not aim to restrict itself to semantic description of findings. Therefore, implicit, interpretative findings are consciously aimed for, resulting in both semantic and latent themes to be

constructed.

With these points in mind, the analysis ought to be coined as a reflexive deductive semantic and latent thematic analysis. It is conducted as follows, with strong reference to Braun and Clarke's six-phase model [42]:

2.2.2.1 Familiarisation

In the first phase, each interview transcript is edited and re-read while revisiting each recording of the interviews to compensate for Whisper's shortcomings and to produce analysable data items. LibreTranslate's output is revisited and treated similarly. Each interview is treated in the three thematic sections as described in 2.2.1 (Motivation, Interaction, Effects). Each section is analytically viewed through the frameworks presented in 3.4.1, 3.4.2, and 3.4.3. The purpose of this phase is to become immersed in the data, produce clean data items, and view the whole data corpus in the light of the derived frameworks.

2.2.2.2 Coding

In the second phase, the whole data set is worked through again, with the aural and visual aid of the recordings. At this point, all data items are introduced to the CAQDAS Delve [43] to support the analysis. The purpose of this phase is to produce first code labels, synthesised from the connections between the initial codes and the data items, both at the semantic and latent level.

2.2.2.3 Generating Initial Themes

In the third phase each code and their data items are reviewed to draft initial themes, collating the grouped codes into their first thematic iteration. The author actively crafts candidate themes by grouping coded data items. The purpose of this phase is to establish shared patterns as meanings, both explicit and implicit across the data set – provisional themes ought to be internally coherent and address the research questions.

2.2.2.4 Developing and Reviewing Themes

In the fourth phase the current themes are revised and critically evaluated based on their central organising concept with respect to the codes in them and their relevance for the posed research questions and frameworks. Further criteria is the sensibility of their relationship to each other, and their meaningfulness regarding the previously reviewed literature and findings from chapter 4.

2.2.2.5 Refining, Defining, and Naming Themes

In the fifth phase the current themes are evaluated and refined based on their strength regarding their internal coherence and sensibility. Each theme is captured by a synopsis and a fitting title in order to be prepared to be presented as final. A final theme effectively captures the essence of all its codes, relates meaningfully to the others, and provides useful insight into the research questions. Phases two to five may be iterated over again until there is insufficient marginal gain of insight.

2.2.2.6 Writing up

In the sixth and final phase a report on the final thematic analysis is presented. The analysis summarising the thematic findings, weaving them into a coherent unit of analysis and presented in the final chapters 5 and 6.

3 Literature Review

As stated in 1, this chapters purpose is to lay the groundwork for the analysis. While the case of OGD in the FHH has its dedicated chapter, this chapter serves to define the case of utilisation of OGD. That is it delineates the conceptual boundaries of utilisation for this paper's analysis. This necessarily includes the wider topic in which the case is settled. According to [38] a case is classically a single person or an organisation. However, a case may also be a more intangible event, or entity, such as "citizen participation", "organisational learning" or "social movements" (p. 29). The case under study is then embedded in the wider context. The MRQ allows for derivation of the main case – the interaction of users with OGD. As this chapter discovers, this is often referred to either utilisation of or engagement with OGD by a user-entity, which is precisely what the case is defined as. According to the research questions derived from the research gap the bounding of the case is set as follows. The engagement with OGD by a given user is caused by a want or need – the motive of his engagement. The actual engagement is characterised by the type of knowledge creation pursued and the skills and tools employed. What follows from this driven engagement is a purpose to be pursued, or a specific effect to be desired, or a specific benefit realised. These three components together bound the case. Further bounds are set by constraining the interaction within the FHH – that is within its geographical borders. This is to ensure a somewhat stable context. Finally the context in which the case happens to be situated needs to be defined and discussed which is what this chapter turns to next.

3.1 Open Government and Open Government Data

As a known and discussed concept OG is not a child of the 21st century. It's usage can be traced back a half-century. [44] show that the term *Open Government* has its roots in the 1950s of the U.S.A., where newspaper editors and reporters pressed for the public's right to know about sensitive governmental records and proceedings. Therefore, in its origin, the concept was tightly related to public scrutiny and the resulting governmental accountability to the public, and in corollary to media.

Open Government *Data*, on the other hand, seems to stem from the merging of the term OG and *Open Data*. Open Data finds its roots in the same country, albeit twenty years

later. Then, the scientific and academic community was concerned with the re-usability and compatibility of machine-readable and other scientific research data [44]. Hence, the *Open Data* component in OGD stems mainly from concerns regarding licensing and re-use, compatibility, and collaboration, and in corollary stood in close relation to academia and science.

[44] also claim that during the merging of the two concepts – OG and OD – from the late 1990s on, fuelled by the proliferation of the public internet, came in parts with an "increase in accessibility [of governmental data that] was dramatic" (p. 191). But, eventually, the merging came also with a loss - the succinct distinction between the former concept's representation of accountability on the one hand, and the latter's representation of interoperability and re-usability of data on the other. The authors argue that, today, the use of the terms in practice could be too fuzzy and vague to have any real meaning at all. The conflation of both terms in OGD could leave room to emphasise and interpret either one or the other in practice.

Indeed, a reoccurring theme in OGD related research is the difficulty to clearly define and outline the concepts of OG and OGD. The reasons for that have been demonstrated to be the inclusion of a wide range of stakeholders and initiatives in practice [14], [15], various interrelated concepts and some confusion about their use [14], [45], resulting in adoption of fitting definitions of the term OG and OGD and their modification for their use in the context, see [46], [47]. For example, while [44] stress the importance of the disclosure of sensitive government information within the term OG, [46] portray OG as having transparency, collaboration and participation as key factors, without taking political topicality into account, while [48] refer to OG as a certain problem solving capability of the public sector.

The authors of [13] attended to this very issue of definition and conception of OG and OGD in their systematic review. The authors conclude that there is indeed a "lack of conceptual refinement of open government" (p. 7), and that many scholars are not even attempting to define the terms Open Government or Open Government Data. Although this is the case, where definitions are proposed, those mostly adhere to the view that *Open Government* as a social-political concept is understood in two major dimensions. First, the willingness of public institutions to publicise various data (or information) – i.e.: *transparency*. And secondly, the willingness to facilitate necessary instruments to receive feedback from the public and then also to act upon it – i.e.: *participation* or *collaboration*. [49] illustratively refer to this duality of open government as "vision and voice" (p. 1). In their analysis, Open Government is conceptualised as "the extent to which citizens can monitor and influence government processes through access to government information and access to

decision-making arenas" (p. 13).

It must be noticed that those viewpoints are rather concise which seems fitting for such a broad concept. Yet they leave incentive for discussion about certain topics regarding the concrete implementation of the "extent" such as skills needed, the actual desire to do so, the digital divide, competency and education of the citizenry to make decisions, or matters of ideology. Further, they make hardly any mention about ICTs, although the transparency dimension certainly implies the use of them. At this point it is helpful to supplement this discussion with the concept of OGD.

This paper agrees with [46] insofar that OGD should be viewed as instrumental in the operationalisation of OG. That is the word *data* in OGD should be emphasised, so as not to confuse OGD with software instances such as portals, applications or web pages for parliamentary documentation that *mediate* governmental data. Or with specific governmental (or otherwise) initiatives that *make use* of governmental data. And especially not with policies that *address* the core values embedded in the concept of OG. Thus, this paper defines OGD in the, more or less, technical sense of *Open Data*: "Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)" [50]. Open data then may be considered Open *Government* Data when it is in principle generated on the basis of tax-funds, regardless if directly or indirectly [5], [51], [52]. Therefore, OGD and OG refer to noticeably different concepts. OGD is a much narrower, and more defined concept which plays a concrete role within the transparency dimension of Open Government. That is Open Government Data represent those information released in the effort to be more transparent. Assuming a truism then, which is that most decisions made are based on information, OGD plays at least partial role in Open Government which is its active role in the transparency dimension.

Here it must be noted that, due to legislative differences[1] across the globe, the data falling under this OGD definition must vary from nation state to nation state. Meaning what governmental data exactly constitutes OGD relies in detail on state-specific legislation and policy concerning various aspects such as disclosure practices, data (set) types preferences, data quality, and obfuscation practices due to preservation of privacy and state secrecy.

To alleviate from the quality related drawbacks, various and rather technical (and thus more objective) standards are often applied to publish, and to classify and rate OGD quality and constituency. The most used are:

- Tim Berners-Lee's Five-Star data deployment scheme [14], [53], [54],

[1] For context in the FHH see chapter4

- the NIST's FAIR-Data Principles [5], [35],
- the Open Knowledge Foundation's Open Definition [5], [14], [35]
- CKAN'S knowledge networks for government [14], [17], [18], [35], [51], [54], [55],
- W3C's Data Catalogue Vocabulary [14], [35], [51], [55], and
- the Open Government Working Group's Eight Principles of Open Government Data [14].

These standards can be used to define and classify, and rate quality of OGD (e.g. 5-star data, FAIR, Eight Principles of OGD) – and to implement publishing and storing or to improve usability (e.g. discoverability, interoperability, contextuality) of already published data (CKAN, DCAT).

Concluding then a reasonable OGD definition comes in two parts. An abstract prescriptive one, and a technical normative one:

> OGD can only be those data which are generated or used through the employment of public funds by some national, or sub-national authority. And which are subsequently disclosed, free of charge, by the relevant public authorities. Disclosed data must be open in the sense of the Open Definition.
> To be considered usable OGD, those data should be machine-readable, human-understandable, freely accessible to anyone, with the use of appropriate licensing and non-proprietary data formats to facilitate interoperability. Furthermore, they ought to be easily findable, up-to-date, complete, well documented and structured, attached to relevant meta-data, and in the best case, to be uniquely identified and linked with other relevant data and context.

3.2 Effects and Benefits of OGD

[16] provide a reasonable account of the evolution of OGD research. They conclude it as going through four phases of research focus from 2009 on, as depicted in Table 1.

Table 1. *Phases of OGD research according to [16].*

Nr	**Phase**	**Time frame**
1	OGD Launch	2009 - 2011
2	Evaluation and Learning	2011 - 2014
3	OGD Adoption and Use	2014 - 2017
4	Implementation and Comparison among Countries	2018 - today

Findings with relevance for this paper in regards to effects and benefits, stem mainly from the first two phases. These phases are concerned with contextual factors (i.e.: legal, cultural-institutional, operational, technical, political, social, economic) enabling or constraining proposed effects and benefits, barriers, and adoption of OGD practices and policies. Thus, those findings are concerned with OGD at large. Generally speaking, the core potential benefits and effects, along with barriers of different types are well explored and understood [24], [26], [27], [34], [56].

Typologically, the effects could well be clustered in dimensions such as *social and political*, *good governance*, *economic* and [24], [26], [28], and *operational and technical* [26], [28]. Additionally, they could be ascribed to impact the micro (individual), meso (organisational) and/or macro (societal) level [13]. However, to view those effects as isolated from each other is misleading. In practice, these ought to be expected to overlap and interact [26], and thus have some synergistic characteristics. E.g.: More employment in the data economy resulting in higher tax revenue (*economic*) could have positive *social* effects through more funding for healthcare and education; Being more capable to use and re-use data to streamline public administration's processes (*operational and technical*) could have spillover effects on cost-effectiveness of public agencies (*good governance*) and their capability to better deliver services (*social and good governance*).

Moreover, the previously cited works ([13], [24], [26], [28]) show that the encapsulation of effects within those dimensions is somewhat arbitrary. Their attribution depends on the interpretation or definition of terms such as *social value* or *good governance*. Considering this circumstance, with the fact of overlapping and possible reciprocity of impacts, it would perhaps prove more valuable to ascribe certain dimensions to actual effects. Nevertheless, this begs the question of what degree of abstraction is appropriate to cluster those effects.

What has clearly been established since, is that the expected effects of OGD efforts are mostly concerned with social value and good governance principles. Social values are, for example, improvements in outcomes of public policy areas [24], or private service innovations regarding transportation, education or culture [57]. Effectively, those impacts could be found on all three levels but impacting both the economic dimension and the good governance dimension.

Viewing *good governance* in the way the United Nations present it, as depicted in Figure 1, leads to a good case for consolidating most dimensions into those rubrics. Predominant effects of OGD utilisation are purported to be more transparency, democratic accountability, participation in decision-making, collaboration, trust in government, or reduction in corruption [14], [24], [26].

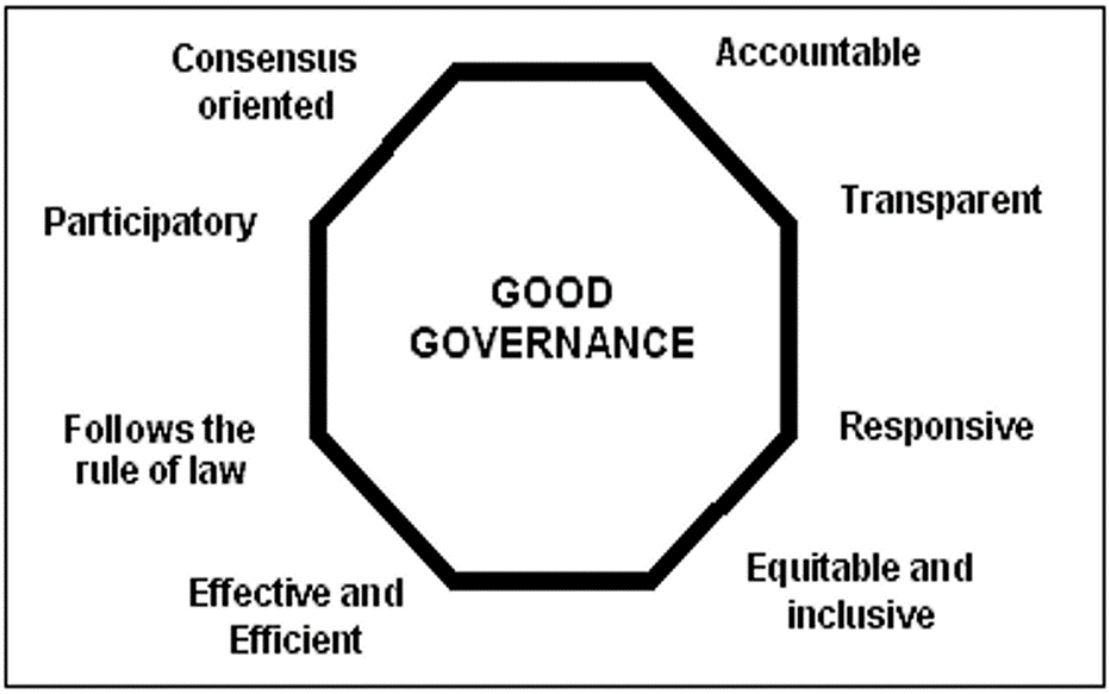

Figure 1. *The eight dimensions of Good Governance according to [58].*

In [58], the point "Effective and Efficient" [58, p. 3] refers to the responsible use of resources to "produce results that meet the needs of society" from the perspective of the government. Although this claim intentionally and unmistakably refers to the use of natural resources, by extension, this claim also ought to apply to financial (i.e. tax revenue) and human resources used by the government. Similarly, the point "Responsive" [58, p. 3] refers to institutions being able to "serve all stakeholders within a reasonable timeframe". Therefore, by claiming OGD utilisation may have positive effects on good governance principles, this paper argues to include efficiency and responsivity gains for the public sector as an organisational institution as potential benefits from OGD utilisation. Thus, from a high-level perspective, social, good governance, and operational and technical effects of OGD will be collectively viewed as *socio-political* effects, due to their interconnectedness and convergence in the concept of good governance. Effectively, this leaves effects stratified in two broad categories: 1) Socio-political effects and 2) economic effects.

Here it must be noted that there are a wide variety of theorised and empirically found barriers that constrain those effects [26], [28], [59]. However, it is not the focal point of this study to explicitly uncover and explore barriers to utilisation by user groups. Moreover, those effects are in parts highly context-dependent. For the sake of brevity, elaborations are withheld until necessary in chapters 5 and 6.

Albeit it being clear what should constitute expected effects and benefits, the evidence for those materialising is rather scarce - and where they are, the results range from mixed to contrary [1], [24], [60] for the socio-political type of impact. For example, while [2] find that evidence suggests "significant and positive, although weak" [2, p. 1224] effects of OGD on trust in public institutions, [61] corroborate that finding but argue that experimental and case studies would show negative effects on trust. Reasons for that are estimated to be the youth of the research, the broad definition and application of terms surrounding the discussion of what constitutes a democracy and the measures applied accordingly [1], methodological bias [61], the high complexity of technology - stakeholder - interests - institutions relationships of the OGD topic [60], [62], and the general difficulty to pin-point actual effects and their evidence [29].

With regard to the economic type of effects, the outlook is quite similar. Although it is widely assumed that OGD will precede economic stimulation in the form of more innovation, employment, GDP growth, and new enterprises, processes, products and services [24], [26], the literature, and accordingly the evidence, of the actuality of those outcomes is also scarce [4], [62]. However, some evidence does point to the importance and value of OGD for the private sector, with respect to geo-spatial and cadastral data [63], [64], for example. But there is also evidence for the contrary, pointing to a relative higher importance of OGD towards the socio-political-types of effects [65].

In sum, hypothesised types of effects of OGD utilisation are well understood, and could be viewed in appearing in one or more dimensions: Economic, social, political, good governance, operational and technical. To add granularity those ought to impact on one or more level: Micro, meso, and macro. For abstraction those dimension may well be consolidated into socio-political and economic. There are considerable difficulties in finding, measuring those with the addition of determining causality. For the purposes of this paper, these types of effects are consolidated into 1) socio-political and 2) economic due to their convergence in good governance principles. However, this does not preclude the sensibility to ascribe them to one or more dimensions or levels.

3.3 Users

As one of the core concerns of this case study is the user of OGD as such, it is indispensable to narrow down what a definition of the user is. Although this study examines the user on the individual scale, it is not the only way to define a user of Open Government Data. For example, there is an interesting, scoped-out take on users of OGD in the sense of stakeholder entities which can, among others, have user-roles. This ecosystem-type view (see e.g., [62], [66], [67]) illustrates the complexity and contextuality characteristics

that come with Open Government Data utilisation. It attempts to draw a multitude of relationships between those entities and derive an OGD life-cycle model from those. It emphasises that OGD utilisation takes place in a highly heterogeneous setting, under the participation of dispersed, independent as well as interdependent stakeholders, which all rely on various *things*, ranging from online tutorials and guides on what to do with data, over data quality rating systems, to governmental policies and strategies.

In the latest refinement of this strand of research, the OGD ecosystem is filled with "actors" (p. 4) which may fill one or more "roles" (p. 4) in the process of generating some value from OGD [67]. For example the government itself (i.e. political figures or parties, and public agencies) may well be considered OGD user(s) as well as providers, among other roles. They may draw from it for their own research and decision-making – leveraging the work of other ministries or agencies. Similarly, NGOs may function as "infomediaries" (p. 2) which process data and broker it between other stakeholders or the public at large. How then go about an organisation such as Transparency International? User, provider, processor, intermediary? All of them? In what function is the organisation what? Can there be a clear distinction? These few examples show how quickly a simple perception of the OGD user can be defined down.

In the context of this study *the user* is defined as any individual that engages with OGD. Engagement in this sense means any intentional effort of any individual to produce some output for some purpose. These outputs may be tangible artefacts such as applications, visualisations or articles [25], or supplemented and enhanced data [68]. A particular motive for the engagement, e.g. economic, political, social, for the fun of use, etc. does not function as an inclusion criterion. Similarly, what specific kind of output is produced does not also. At the core of this study is the usage as such – the utilisation – which reflect the generation of value in the eyes of the user. In fact, these precise motives that are connected to the engagement are subject of this study.

With regard to typical users of OGD, a wider range of types have been articulated. Generally speaking, the literature defines users as citizens, businesses, researchers, developers, NGOs, and journalists [24], government employees [69], entrepreneurs, employees and hobbyist [56] among many more. On the other hand, some case studies expose more differentiated views. OGD users are seen as "civic technology community" and "transparency advocates", and "OGD Beneficiaries" [62, p. 25] as downstream consumers of the resulting products and services. Or they might be "hackers", "students", "start-ups", or "friends" without particular defining characteristics [70, p. 283]. Or they might be "individual members of civil society" (i.e.: citizens) [71, p. 61].

The authors of [25] suggest viewing users as having capabilities and filling roles in OGD engagement. Whereas the former maps on to what one could call occupation of the individual (student, teacher, researcher, employee, consultant, developer, analyst, journalist, etc.) which might not directly be visible during engagement. The latter relates to the individual's function within the engagement. It is observable and might be more than one defined role per engagement such as designer, promoter or domain analyst. Those roles often map onto specific capabilities.

The point being, the definition of *user* is also highly context dependent, along with their relative importance to the matter. It is subject to the setting and lense of any particular study – conversely it means: In a geographically small area of high political authority, such as the federal state of the FHH, general statements about users and their motives might lead to some useful assumptions, but the literature suggests that, due to context-dependency, the patterns ought to be different in detail.

3.3.1 Motives

When examining the intention of adoption of OGD by users, it is practically impossible to ignore the research done on the acceptance and use of OGD technology. Numerous works have been published, that aim to determine factors that are supposed to explain and predict the adoption of OGD technology, mainly utilising the IS success model, TAM and UTAUT [19], [20], [33], [46], [72]–[74], based on the works of [75], [76].

The results show how the constructs[2] influence the adoption of ISs by users. SI, PE [33], and EE, PE, SI [74] are posited as main influences, but also FC in Eastern countries like Bangladesh [20] for example[3]. What seems to be a drawback of those models is that they do not aim to predict the continuous usage intention of ISs [20]. Moreover, for what those results are worth, they do not aim to make any claims about the users themselves, with respect to their defining characteristics relevant for understanding the utilisation itself. That is: They do not aim to explore what driving forces are behind the engagement with OGD by the users at the most fundamental level, they do not aim to give any insight into what the users actually do, or want to do, or what ends are anticipated as a result of their engagement. They only aim to assess the momentary intention of adoption of a given technology.

This conclusion is reflected in the meta-analysis of [25] mentioned in 1.3. There is little literature exclusively focusing on particular motivations of individuals for considering en-

[2] Please refer to tables 4 and 5 in appendix 1 for a brief overview of the constructs used.
[3] All of the acronyms meanings can be taken from table 1

gagement in OGD. However, the authors' synthesis of driving factors of citizen engagement yields useful results – a comprehensive list of drivers in six categories:

- Personal factors (intrinsic motives such as fun of it, learning, intellectual challenges),
- performance related factors (extrinsic motives such as usefulness of outcome, recognition or career related),
- economic factors (e.g.: economic motives such as financial rewards),
- social factors (e.g.: networking, solving social problems),
- technical factors (e.g.: system or data quality, functionalities) and
- political factors (e.g.: political participation, public good, trust in government) [25].

This synthesis culminates in a proposal of a conceptual model in which those factors influence the OGD engagement. This relationship is mediated by a seventh category of items – Citizen's profiles – which consists of gender, age, education, resources, capabilities, awareness, competency, experience and voluntariness. The authors of [25] find that the "three most frequently mentioned factors that drive citizens to engage with OGD are related to performance-related motivation," [i.e. extrinsic motivation], "intrinsic motivation, and political interest" (p. 14). The weakness of these results however, lie in their various origins. As those results come from a vast array of papers published from all over the world, utilising various methods, from various contexts they may serve as a list of references but can not serve as a checklist for practitioners for their local OGD initiatives, nor can they serve as a predictive model for a context-specific study like this one. To illustrate: While [25] find that extrinsic motivation is the most mentioned factor for engagement in OGD, [72] find that, in the German context, extrinsic motivation "turns out to be completely irrelevant" (p. 315). Thus, vast generalisation can lead to contradicting results.

Drilling into the literature to find those publications which are more closely in line with this study, the list becomes short. Six publications are identified ([46], [56], [57], [68], [72], [77]) to fall into consideration, based of the following criteria: Published not earlier than the year 2010. Conducted within a European legal and cultural context. Not centred around a specific (set of) government-led Open Government Initiatives, which may skew the participants' profile, but rather focusing on voluntary engagement with OGD, with specific focus on users and their drivers or motivations facilitating their engagement.

The two selected studies of Wirtz, Weyerer and Rösch ([46], [72]) identify the same precise issue as this study does. Infancy, heterogeneity, and under exploitation of OGD utilisation paired with the issue of the lack of citizen-focused and usage related studies. Both studies take a similar approach to investigate "citizen expectancy" (p. 566) [46] and "antecedents" (p. 308) [72] of OGD. They utilise multivariate data analysis and structural

equation modelling under the use of the technology acceptance model, supplemented with additional variable constructs. Both rely on the same survey data which includes 210 samples for the whole of the FRG which they acknowledge as a possible weakness. Their results demonstrate the validity of the TAM however they seem insufficient in explaining underlying drivers and motivations. As pointed out above, like manifold similar studies, those results do not go far beyond proving that perceived usefulness and perceived ease of use of a given OGD portal, and further unidentified intrinsic motivation precedes the general German's intention to use OGD. They do not shine light on underlying drivers, actual interactions, desired effects of any user. Further, they do not consider any particular user profiles or the possibility of different evolutionary trajectories of OGD efforts and policies among the federal states.

Although the authors of [57] utilised the UTAUT, their study yields some insightful results. The authors surveyed businesses in the Netherlands and Sweden to assess drivers of 25 individuals to innovate services. Although they precisely aim for businesses, they find that motivations vary both between countries and nature of enterprise. A range of factors play into that conclusion such as the specifics of data supply, cultural and social trends, innovation tendencies, skills available or personal investments in issues. The top three motives to innovate with OD are an amalgamation of, in that order, intrinsic (interest and curiosity), extrinsic (economic added-value) and decent support and data quality (infrastructural or technical arguments). The authors conclude that "generalising in the open data domain creates inadequate expectations. The drivers and dynamics behind the development of social innovation projects are different from those aiming for commercial success" (p. 32).

In a rather broad study, which is highly similar to the intention of this one, Davies [68] investigated users and uses of *data.gov.uk*, using interviews, surveying, and participant observations. He found motivations in a set of six categories:

1. Government focused (understanding government, promoting efficiency and accountability),
2. Technology innovation focused (driven by creating new platform/tools, linked data technology),
3. Reward focused (driven by recognition and profit),
4. Digitising government (driven by technology-driven improvements for government),
5. Problem solving (driven by specific domain challenges), and
6. Social/public sector enterprise (improving/providing public services).

Although not intentionally excluding voluntary sector workers Davies found that very few

subjects in his study claimed that particular background.

[56] interviewed 19 OGD users who procured data through a specific OGD marketplace in Sweden and found that, first, employees were rather driven extrinsically, i.e. to complete work tasks and were not interested by knowledge sharing. Second, entrepreneurs were rather driven by a mix of intrinsic and extrinsic factors, i.e. to attain project goals and the challenge along the way. And third, that hobbyist were driven by the fun of exploration and manipulation of data. Those users were much more interested in sharing knowledge and insights, and were much more affected by administrative obstacles such as data availability and authorisation procedures.

[77] interviewed OGD users ranging from software developers and academics to local authorities and international experts. The subjects were mainly from the Republic of Ireland as well as other Anglo-saxon and European countries. Although nearly all interviewees were concerned with transparency as an objective, the authors found that motivations to engage in OGD range form political, to community oriented, or purely personal factors such as the fun of exploration.

What all those works [24], [25], [46], [56], [57], [68], [72], [77] have in common is that they point to the lack of understanding and exploration of user focused OGD utilisation in general, and specifically with regards to who uses OGD and why. This seems to be somewhat false as there are a number of publications, many of which included in the discussions of this paper, that attempt to shine light upon that issue. However, this seems to be somewhat true as well, as findings are corroborating, contradicting, and differing altogether, depending on scope, context, and theoretical and methodological background. To be descriptive about the findings in one word: They are inconclusive. Particularly [25], [29] point out that there is a lack of consistent use of frameworks, which is certainly true. Moreover, the current research seems to lack a specific depth regarding the interaction of users with OGD and their underlying drivers with special respect to any longitudinal studies. To illustrate: What can consistently be found and used are a differentiation between internal and external, or intrinsic and extrinsic motivation to engage in OGD at the more scoped out level.

1. Extrinsic motives (e.g.: future career concerns, usefulness for job-performance purposes, direct instruction [56], [72], [78], [79]), and
2. intrinsic motives (e.g.: fun and enjoyment, data exploration, altruism, intellectual challenges and learning [56], [70], [74])

But even in specific terms, the extrinsic – intrinsic differentiation holds. The authors of

[48] have examined possible motivations for engagement in a specific setting – namely an OGDI involving querying citizens for ideas and opinions on three policy areas: Family, education and innovation over a course of eight weeks. Even though this setting was quite specific in focusing on social issues, a wide range of levels of engagement and different motivations for engagement have been found. The participants were motivated by both intrinsic (discussion, ideation) and extrinsic (dissatisfaction, political improvement desired) factors. In a similar vein, [70] analysed a three-month contest in order to find out what motivates the public to engage in Open Data Innovation. Again, the scope was quite specific: An innovation contest was held by a public transport organisation. Compared to the previous example, however, the aim was to develop a digital service to make public transportation more fun, efficient and accessible. Again, a range of motivational factors have been found, both of the extrinsic and intrinsic type. In this case the outlook of financial gain was present, although not guaranteed, and most participants were start-ups or third party developers. While intrinsic factors outweighed extrinsic ones somewhat, both were found and analysed.

The point is that the scope of analysis does not seem to significantly impact the classification of the results. Meaning, it seems insufficient to answer the question of *why* certain people engage in OGD with extrinsic–intrinsic arguments. Moreover, a mixture of a wide range of motivations are to be expected in any case, however their balance and exact constituents (as in underlying motives) depends on what is looked at and seemingly influenced by the particular lens. Finally, any attempts to generalise in a wider sense seems to be inadequate to yield conclusive results.

These findings highlight the context-dependency regarding the outcomes of an analysis of motivation for OGD engagement. They show the wide range of motives which could possibly affect engagement on the one hand, and on the other, the fact that findings are uneven to say the least. This leads to an argument for studies closer to the user in a defined context with respect to legal, political, cultural and geographical factors. An attempt should be made to constrain the scope as much as needed in order to determine a profile of engagement factors within one political entity. This suggestion is based on the defining characteristic of OGD. OGD originate from political entities such as administrations, ministries, and agencies concerned with a specific set of people. Those within their legal, political and geographical authority.

In sum, the research does have evidence about drivers and motives of interaction with OGD. Yet, the results are mixed to inconclusive. Generalisation are misleading, therefore the literature can make only *assumptions* about the users in specific contexts [24]. Thus, especially when it comes down to a geographically and politically defined area, those

assumptions turn out to be weak hinge-points to base substantial and goal-oriented technological and strategic considerations on. What factors influence adoption of a given OGD portal is well understood, however not in terms of underlying motivation and drivers to utilise OGD as such. However, the literature can tell what those ought to be in the general picture, but shows also at the same time that these often do not hold in specific terms. The same logic applies to user-groups. Those are defined in general and also found in specific case-studies, along with various other type-definitions.

3.3.2 Utilization Types and Skills

This study views engagement with OGD as any intentional effort of any individual to produce some output for some purpose. From this perspective, engagement refers rather to the origin – the *why* – of the effort (i.e. motifs and drivers) to produce for some purpose. OGD utilization, however similar in meaning, refers rather to *what* this engagement can be described as, what it requires and what purpose it entails. In their review of OGD utilization the authors of [24] classify the found types of utilisation as either analytic or synthetic knowledge-creation. Both types include knowledge-creation through OGD expressed "typically as a field of study or practice" (p. 15), such as data analytics, innovation, decision-making, smart-city or new services (see figure 3). The authors find that, by a large margin, innovation is the key concept discussed in the literature. However, they concede as well that innovation might both be seen as a type and an effect of OGD utilization and might be broken down into multiple dimensions as well. With respect to concrete interaction, few studies have been conducted to explore, and much less to explain, *how* actual usage can look like [24]. From those which did, the results are diverse and abstract [80], and scarce which, again, point to the contextual nature of OGD research [24].

Some far-fetched but concrete inferences can by made through [18], [19] from web traffic analyses, which define some sort of usage on OGD portals in New York and Philadelphia. Users seem to create maps, browse, search, and filter data sets, view pages, link to RSS feeds or social media, and download and print data. While this approach might be feasible in the U.S. American context, this kind of data collection is defacto an illegal practice, at least in the context of the FHH [31].

Those works which aim to analyse the interaction of users with OGD explicitly, subsume their findings in various concepts, with varying granularity and related meanings. For example, [80, p. 213] find that users go through the phases of "start, identify, acquire, enrich and deploy". Those phases contain various activities (e.g.: explore and assess, extract, cast, map, concoct, etc. [80, p. 224]). Their findings indicate that the user process may vary dramatically in kind and complexity. An initial condition such as the driving

factor for engagement, which could be either data, demand, or idea driven, would impact the configuration of the whole process as such.

On the other hand, a study similar-in-kind to this work from 2010 [68] proposes five high-level processes of OGD utilization. Those range from low to high complexity. Accordingly, OGD utilization processes range from simply searching and browsing, and creating visual and textual conclusions ("extract" and "report"), to more complex technical operation like creating software to support exploration, the provisioning of APIs, and the creation of new products and services ("interface, data" and "service") [68, p. 26]. These results show the broad range of possible interactions with OGD, their possible difference-in-kind, as well as their range in in complexity. This hints at contextuality again, and to the fact that a wide set of skills are in question when utilising OGD.

In 2012 already the authors of [26] point out that technical skill-sets of users needed to meaningfully generate value with OGD might be underestimated. Technical skills relating to statistics and programming in tandem with the time needed to deploy those ought to be rare among those who wish to benefit from them – thus more OGD might exacerbate the digital divide. Specific technical skills of users needed, as much as general cognitive abilities (e.g.: knowledge, experience, professional competencies) too, are correctly viewed as conditions or moderating factors influencing OGD utilisation [24], [25]. Those needed to innovate with OGD are seen as short in supply [81] and more important than financial assets in order to generate value from OGD [57]. However, not only specific skills are needed – [82] express, for on open data based learning for students, the need for critical thinking, teamwork skills, and general research skills, along to a competency of global citizenship.

While specific users bring specific sets of skills and demands when offered OGD (e.g.: [83]), [69, p. 235] find, in a concise review of needed skills of OGD users, that "the literature is scarce in relating types of users and needed skills". What can be inferred logically, is that the skills needed or employed must be dependent on the type of data, the individual's desired use of the data, that is the task complexity, and the data's quality and findability. With respect to technical and managerial skills needed and employed the authors of [69] find that those that are presented in the ODI Data Skills Framework [84] (see figure 2) largely map to those found in the literature. However, the authors find also gaps in the ODI framework in skills pertaining to "the ability to ask meaningful questions about specific public policy problems, identify ways to answer the questions using data, and report findings from the analysis" even though the ODI Data Skills Framework is a comprehensive work, covering a broad set of skills and application.

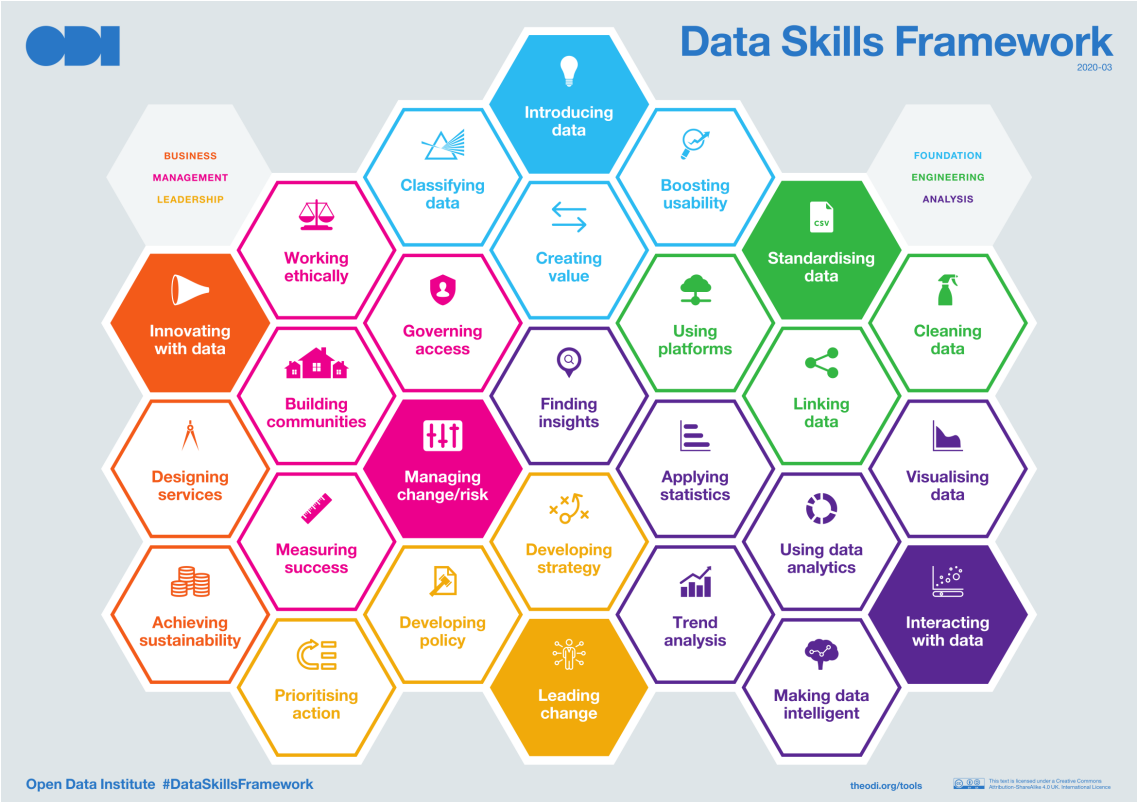

Figure 2. *Open Data Institue Data Skills Framework [84]*

Additionally, some authors discuss and promote *data literacy*. This concept does not seem to be universally agreed on through a single definition – that is it is layed out differently. However, it finds application in research (e.g.: [85]). This specific perspective on the concept follows from a synthesis of various literacy concepts (digital literacy, informational literacy, multiple takes on data literacy) into one [86] and derives core competencies of data literacy:

- Understanding data (e.g.: types, purpose, origin),
- finding and/or obtaining data (e.g.: sources, quality assessments, methods for obtaining),
- reading, interpreting and evaluating data (e.g.: comprehension, analysis, summarizing),
- Managing data (e.g.: reference management, databases), and
- Using data (e.g.: tool usage (SPSS, Excel, R), representation, ethics).

A differing view on data literacy is expressed by [87] defined as “Critical Data Literacy is the set of abilities which allows one to use and produce data in a critical way” (p. 117). This view proposes data literacy as a reflective process between subjects and context attainable through stages (i.e.: Investigation, Thematisation, Problematisation, Systematisation) as it is based on education pedagogy. However, the cited “abilities” needed are similar to those

of [86]:

- Data reading (e.g.: origins, interpretation, contextuality),
- data processing (e.g.: technical processing through tools),
- data communication (e.g.: visualisation, representation, ethics), and
- data production (e.g.: publication practices and tools, data formats).

Lastly, a very current study points out that once OGD initiatives are up and running, the top two reasons why they fail are the platforms' "functionality and support" and their "inclusiveness" [30, p. 288]. The former meaning that they fail to address the "lack of knowledge among users, the lack of advanced search support, the lack of data analysis support tools, the lack of data visualization and interpretation tools, the lack of interaction media on the platform and the lack of a help desk" and the latter meaning that they fail to address the "the lack of skills of some groups, the subsequent exclusion of these groups in OGDIs and the digital divide, meaning that not all groups can use the platform and participate equally" [30, p. 288]. This is a clear indication that OGD decision-makers and administering stakeholders are *not* aware of their users – neither in terms of skills and interaction, nor expectations, nor motives.

In sum, types of utilisation are hypothesised and empirically verified. The literature is rather clear on necessary technical skills needed in order to work with respective data, and acknowledges other skills needed for a broader context, not necessarily connected with technical data. Moreover, it is acknowledged that further-reaching skills are necessary such as understanding policy formulation and programs to effectively act upon OGD. The literature also shows that the user is rather not a focus of examination, and therefore demands more attention. Frameworks exist to conceptualise user interaction with OGD, but at the end of the day, it is evident, that OGD decision makers are not aware of the skill and usage patterns of their respective users, although comprehensive frameworks for such purposes exist.

3.4 Conceptual Framework

OGD utilisation as a framework (see figure 3) is a set of related components describing the generation of value through the re-use of OGD by any type of user-entity. Based on the framework, this paper adopts the view that individual *OGD users* make use of one or more *type of utilisation* to cause (or that causes, only) one or more *effects of utilisation*. Types of utilisation are analytically distinguishable, at least, between analytic and synthetic use[4].

[4] See Section 3.4.2.

Types of utilisation effects can be more closely distinguished between social, economic and good governance, but are somewhat consolidated and viewed as binary in terms of either socio-political or economic[5] Users are either direct users, who actively engage with OGD (relevant for this paper), and indirect users as the aforementioned beneficiaries of OGD utilisation effects, which do not (irrelevant for this paper). Furthermore, both types and effects of utilisation are, in some way, constrained by contextual *types of OGD utilisation conditions*, which are of either technical of social nature. Due to its originality, and close relevance to the topic, this framework is of highest appropriateness for this work.

In particular, a user engages in OGD if there is an *individual's conscious cognitive engagement* with a certain piece or set of OGD – regardless of whether or not the individual is part of a public sector organisation, and regardless of the specifics surrounding the type of utilisation or the kind of data involved. Therefore, this definition excludes circumstances such as the mere publishing of OGD, or the use of OGD in an Web or mobile application, where the user does not necessarily needs to be aware of his or hers engagement with OGD.

As a baseline framework to view utilisation of OGD, this papers uses Safarov's [24] utilisation framework, see figure 3.

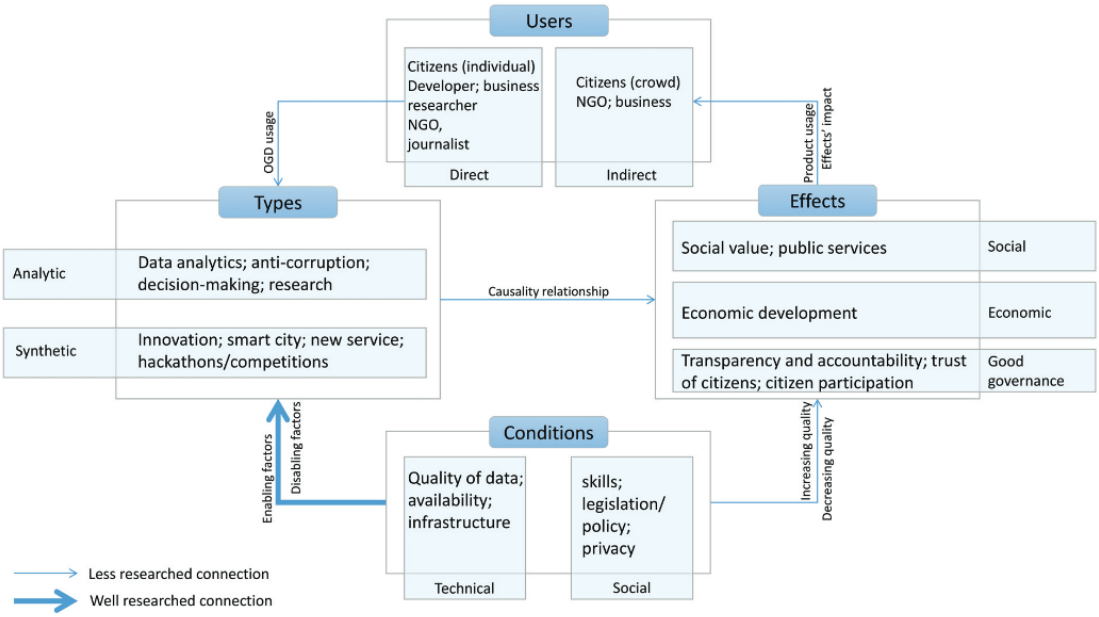

Figure 3. *Utilisation Framework adopted from [24]*

[5]See section 3.4.3.

3.4.1 Why users engage

As eluded to in 3.3.1 there are a number of proposed models to describe motives for engagement in OGD. Two frameworks seem useful to structure one's thinking about underlying drivers or motives. First, as a highly similar-in-kind study by [68] suggests, drivers ought to be found accrued into "six overlapping clusters" (p. 22). These have been extracted through interviews and questionnaires without restriction in terms of users, and thus are deemed appropriate for this study.

Table 2. *Overview of Motivational Clusters according to [68]*

Name	**Description**
Government Focus	"wanting to better understand government and to promote efficiency and accountability"
Technology innovation focus	"interested in creating new platforms and tools, and in semantic-web/linked-data technology"
Reward Focus	"seeking recognition and/or profit"
Digitizing Government	"[s]eeking technologically driven improvements in efficiency and functioning of government"
Problem Solving	"using OGD to meet specific challenges"
Social/public sector enterprise	"using OGD to provide services in/to the public sector"

Second, drawing from the literature of open innovation, and with respect to the former classification of extrinsic and intrinsic motivations, [88] studied participation incentives in the realm of open innovation and open source software communities. These are characterised as intermediaries who "maintain an online open innovation community for members who are ideating online collectively or individually" (p. 441). This basic idea is close to the spirit of collaboratory community work attributed to the open data field, and consequently should pose an appropriate view on OGD utilisation as well. The authors identified, similar to the findings from [25], that motives for participation in online communities are threefold. The following framework is extended with the findings from [25]

1. intrinsic [25], [88],
2. extrinsic [25], [88],
3. social motives [25], [88],
4. economic factors [25],
5. technical factors [25], and
6. political factors [25].

Intrinsic motifs refer to "inherently interesting or pleasant" (p. 443) activities where the task itself is the motivation to participate. Extrinsic motives, on the other hand, are those where the consequence of the activity is the motivating factor. Social motives refer then to those where the "good of the group enters one's utility equation" (p. 445). These categories of motives relate to "altruism, [...] and [...] commitment to the community" (p. 445) [88]. Which are those motives that seem to be likely as core for those members of the community which aim to pursue the good governance goals in the socio-political type of effects. Economic factors are somewhat similar to extrinsic however with clear financial prospects in view [25]. Technical factors reflect aspects such as good data quality or quantity, user friendliness, specific capabilities or functionalities [25]. And lastly, political factors represent the interest in government and the expectation of interacting with it through participation in order to improve trust or decrease corruption, for example [25]. Thus, taking into account the proposed collaborative character of OGD, this framework seems highly useful to categorise incentives, or drivers and motives behind OGD utilisation.

3.4.2 How users engage

Refering to the utilisation framework in figure 3, it follows that certain users employ certain types of utilisation to pursue a certain type of effect. Adopting the view held in [24], this paper proposes to see general utilisation pattern in two categories - analytical and synthetic utilisation. These concepts stem also from open innovation research, and refer to knowledge-creation processes. Analytical knowledge-creation refers to problem solving processes and capabilities. They typically result in "know-how" (p. 1047) [89]. It is rather interpretivistic in nature. Synthetic knowledge-creation on the other hand, refers to the "understanding and explaining of the natural world" (p. 1047) [89]. From which follows the know-why kind of knowledge [89]. This process is rather positivistic in nature. This framework is applied in [24] which has direct bearing on utilisation types, and therefore is deemed as appropriate for this study.

Moreover, another highly-appropriate and valid viewpoint is exposed, again, in [68]. The authors argue that there are five distinctive "processes of OGD use" (p. 2), which have briefly been mentioned in section 3.3.2. These processes have been observed in a very similar study, and thus are expected to be found in this study as well.

Thus, utilisation patterns are expected to be found in either of two dimensions. (1) Analytical and (2) synthetic. Within these dimensions, the utilisation-types should be appropriately classified as one or more of those processes from Figure 4.

Additionally, what comes as a logical conclusion, the knowledge-creation processes of

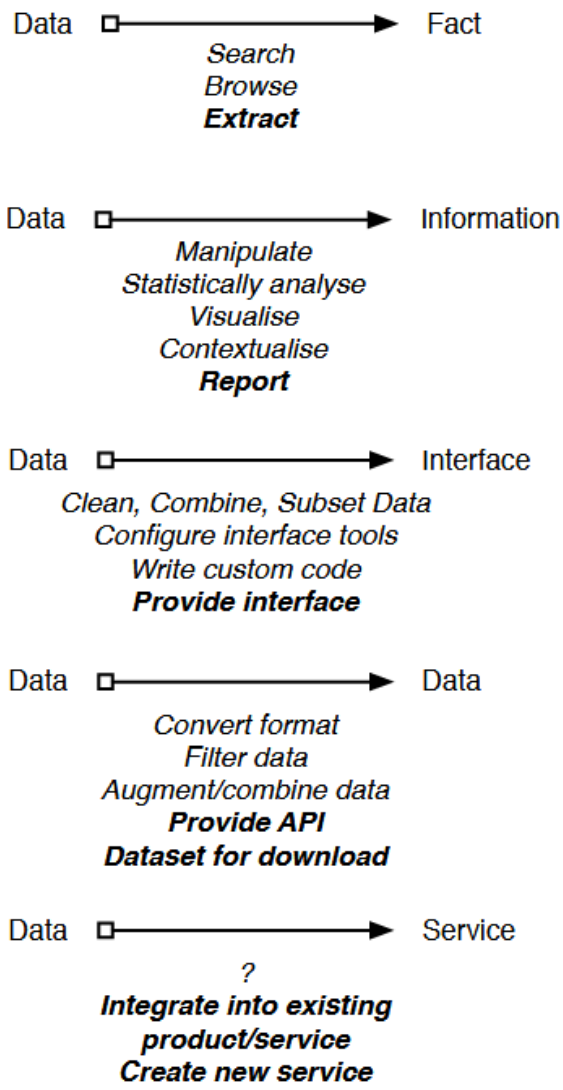

Figure 4. *High-Level OGD utilisation processes according to [68]*

OGD will need to be inevitably facilitated by some hard skill and/or tool. In the light of [30], OGD initiatives mainly fail because of the lack of "functionality and support" (p. 388). Thus it is naturally necessary to uncover what facilitating skills and tools are used for the employed knowledge-creation process. Regarding structure around skills to be employed by users in their OGD engagement the ODI framework from figure 2 is applied.

3.4.3 What effect users anticipate

Finally, to come full circle so to speak, it remains to know what the actual end of the OGD utilisation is aimed to achieve in order to draw a full sketch of a characterisation of OGD utilisation. As this paper pointed to in section 3.2, it is sensical to view the proposed effects of OGD utilisation as in three categories, of which *good governance* type of effects are those which the literature is considering the most anticipated. While there needs to be an

overlap in those effects, due to their inherent nature and the interconnectedness of context, it seems appropriate to coarsely consolidate the effects in two dimensions – *socio-political* and *economic*. While that may be the case, it is similarly feasible to consider not only a type of effects, but also an abstract *location of impact*. Therefore, his paper proposes to view those effects in 2 × 2 dimensions:

(1) Either affecting primarily

(a) **external** towards the public at large, or primarily

(b) **internal** towards public administrations, agencies and governments. And

(2) either of rather

(a) **economic** (e.g.: Cost-effectiveness internally; Increase in tax-revenue or GDP externally.) or of rather

(b) **socio-political** (e.g.: Higher trust in the public sector or government externally or internally; Increase in political participation and decision-making externally.).

This model-view is visually illustrated in figure 5, where x-axis corresponds to (2) and the y-axis to (1) from above.

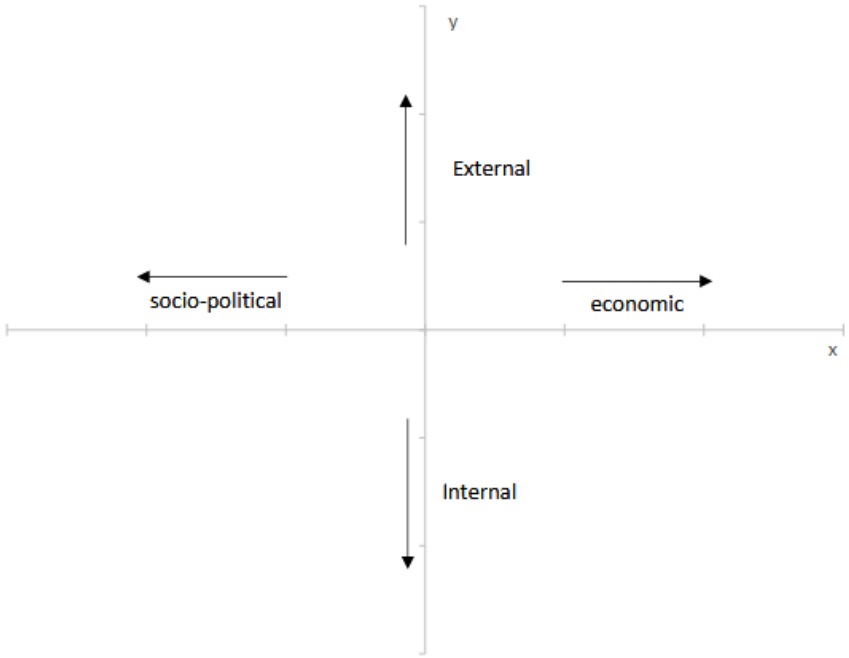

Figure 5. *OGD Effect Model*

4 OGD in the FHH

As this study is an exploration of users' interaction with a particular piece of technology and data in a particular place, it is reasonable to offer some justification for the choice as well as to shed some light on the context which emerges as a result.

4.1 Scope

As discussed above, vast generalisations regarding the topic of OGD utilisation lead to inadequate results. Therefore, findings from one instance should not be assumed to hold true in the context of another. It is clear that any effort to publish data for re-use by any agency is directed at some user – or generally speaking the public. It is precisely not the argument of this paper that the data published in the FHH are intentionally directed at a particular user – for good reason, although this argument could be made. First, this conclusion would be drawn based on findings from the extant literature which would mean generalising findings, which was proven to be inadequate. Secondly, in doing so it would, by default, restrict the scope of this paper to a specific user group and thereby forfeiting any expectation to give some sort of balanced analysis of user interaction with OGD in the FHH. Moreover, this would even further decrease the potential candidates for this study, which is, given the general ignorance about the OGD user at large, not desirable.

To come to a coherent set of findings, it should be the aim to reduce the scope of analysis to such a degree that all potential data sources are dealing with the same external contextual factors (i.e. data quality and availability, legal, political, institutional). While it is obviously impossible to control for all factors being equal, a restriction to a single, city-sized Federal State should constrain those factors reasonable tightly. In this respect, the FHH, in which exists a comparably mature [90] Open Government Data portal over which it has full authority, is a reasonably good choice. A logical next step to further constrain the context would be to restrict the scope of analysis to a particular kind of user, data or data set. However, this again, would restrict the potential data sources and possibly yielding insufficient data. For that reason the scope is drawn around users of the entirety of OGD available in the FHH.

4.2 Origin and Legal Basis

The central Open Government Data portal, the *Transparenzportal* (literally: Transparency Portal) is found here. The basis for which is a Freedom of Information Act-type law called the *Hamburgisches Transparenzgesetz (HmbTG)* (literally: Hamburg Transparency Law). The spiritual origin of the law is a true grass-roots movement in which people from Hamburg used social and political instruments available to formally demand a change in public policy. The process started formally in October 2011 [91] and led, according to Transparency International [92] and the Hamburg parliament [93] to one of the most progressive transparency statutes in Germany if not in Western Europe.

The initiative named *Transparenz schafft Vertrauen* (literally: Transparency creates trust), organised through a cooperation of various NGOs and civil associations[1], collected over 15,000 signatures within six weeks, and thus motioned a draft bill for deliberation to the Hamburg parliament. Given recent political circumstances, the initiative effectively compelled the parliament to act upon the motion.

At its core, this draft bill proposed to turn the former *Informationsfreiheitsgesetz* (literally: Information Freedom Law) from a passive law, where the release of governmental documents needed to be explicitly applied for, for a fee in many cases, into an active law by introducing a general duty to publish information only with selected exceptions. The proposals at the centre of the initiative were to increase the ease of access to information which ought to make corruption more difficult, guard against tax waste, increase trust in politics and administration, improve administrative processes and enable civic participation [91]. The law's first edition was unanimously adopted on June 13th 2012 [94] after previous adaptation through intensive deliberation with the initiative, subject matter experts and a committee. Albeit the draft bill underwent some changes, the core of it has not been removed. The resulting bill adopted was hailed as a renunciation of official secrecy for the benefit Open Government Data principles – a "quantum leap" (p. 2) [93].

Today, the *HmbTG* is still held as one of the most extensive Freedom Of Information-type law in Germany. According to a ranking held by the two civil associations, Open Knowledge Foundation Germany e. V. and Mehr Demokratie e. V., the legislation ranks at the very top [90]. Compared are the legislation from all 16 federal states in Germany under the aspects of the proactive publishing of information, range of default inclusion of

[1] Mainly:
- Chaos Computer Club e. V.
- Mehr Demokratie e. V.
- Transparency International Deutschland e. V.

For more information mostly in German, see this page.

obligation to publish, management of exceptions, fees, and competencies of the freedom of information commissioners [95].

4.3 Evaluation

After five years, in 2017, the *HmbTG* and the OGD portal underwent a legal and technical evaluation [96]. The legal analysis found that 98% of the data objects published originated from ministries and courts. Given the amount of data published, the report found that the bodies subject to publication were following their lawful obligation to large extent. No substantive fiscal impact or hindrance in operative capability were found. Thus, the implementation of the law resulted in the initiative's desired effect without substantially strangling the ministries and courts financially or operatively by and large.

The technical analysis of the portal included a user analysis, their motives to use, their usage intensity; the implementation from the perspective of bodies subject to publication; the user-friendliness of the portal; and finally an assessment whether the information published on the transparency portal contributes to the satisfaction of the information interest of its users. The following summarises the report's results regarding user groups and motivations for use.

The user (hereafter: private users) analysis was conducted via a questionnaire which resulted in 385 complete and 27 partial responses. According to that survey, the users were 43,5 years old (SD = 13,1; n = 352), predominantly males (57,5% against 33,7% female, rest n. a.), highly educated (68% claimed a university degree), and more than 70% seemed to claim good or very good competence in political matters. Most claimed to live in the FHH (73.3%) and to hold German citizenship (90%). Moreover, 896 employees of the FHH (hereafter: public users) were surveyed (49 partial responses) which, regarding age and gender distribution, were somewhat different comparing to the public (47,1 years (SD = 10,9; n = 623), 45% female against 39,8% male, 15,2% n. a.).

For the user analysis the private users were asked to tick one or more boxes that fit the function in which they access the portal. The results show that the private users largely claimed to be "private individuals with personal interest" (p. 140) (61.7%). Remaining frequent options were some entity (i.e. company, individual or group) with economic interests (17,2%), scientists (12,9%) and approximately 7,8%[2] have been found to relate to public employees, students, NGO or civic initiatives[3].

[2]Derived from the replies to the "others" category. See footnote below.

[3]The remaining options were:

- Parliamentarian – 1%

Regarding the private users' reasons why they would use the portal for personal reasons, the results show that more than half of the users (52,2%) chose "interest to get information/overview about a topic" (p. 141), 41,5% chose "interested in a specific procedure", and "general political interests" was ticked by 28,4%. Personal interest in specific administrative or court process was chosen only by 10,4% and 3,2% respectively. Regarding the private users' reasons why they would use the portal for professional reasons, the results show similar behaviour for administrative and court processes (2,7 and 10,9% respectively). Interested to procure information/overview about a topic was chosen by 27,2% of users, for scientific or economic purposes was chosen by 18,4 and 15,5% respectively, and for political purposes was ticked by 10,2% of private users.

Regarding the motivation for use, those public users who specified to have used the portal more than once in the preceding twelve months were asked to specify whether they used it for private or professional, or for both reasons. 79,6% specified to have used it for professional reasons only, and 6,6% for private reasons only. 13% of public users specified both and less than 1% specified none.

The concluding remarks of the report state that the largest share of users are by far private individuals, while other user groups still find significant representation. Regarding the usage frequency, a wide range of usage intensity has been found. The private users have a much higher frequency than the aforementioned public users. The public users were by far using the portal for professional purposes.

While this evaluation report gives first important insights into usage patterns of OGD in Hamburg, the evaluation did not intend to uncover underlying reasons for engagement or tried to uncover individuals use-behaviour regarding specific interaction, or knowledge creation types. Furthermore, no kinds of long-term anticipated effects were attempted to synthesise.

- Journalist / Media – 5,1%
- Lawyer – 1,9%
- Interest group – 3,9%
- Union – 0,7%
- Political Party – 0,7%
- Other associations – 2,2%
- Operators of internet forums, who over consultation for specific topic – 1,9%
- Others – 20,4%

5 Thematic Analysis

Before starting with some declarative findings first, regarding the research questions and the situation in the FHH, a disclaimer ought to be held in mind. Some of following data stemming from interviews must be seen in the light of the past five or even ten years. Meaning the experiences and opinions reflect project work that, in a few instances, started or happened further back in the past. Those particular views may or may not accurately reflect the current technical state of a discussed information system.

As mentioned in section 2.2.1 a first, preliminary, open-ended, interview was conducted with the relevant authority at the Ministry for Culture and Media of the FHH. The purpose of this interview was to explore the topic of OGD in the FHH from the provider perspective to gain more insight about the situation and viewpoints from that perspective. It quickly became apparent that those issues which were identified in the literature, see 1, were dominant in the specific context of the FHH as well. Namely first, practically no insight or knowledge of value may be extracted through any interaction with the transparency portal regarding user groups or characteristics, or behavioural patterns. As the IP addresses are obfuscated automatically and are not saved it is not possible to infer any useful assumptions. Effectively, only data requests are logged, thus a certain popularity of a specific data object may be inferred. However, there is no way to tell who or what is behind this request – it may well be an automated bot, or a script of some form, from practically anywhere on the planet that is the source of those requests. However, the requests may be classified into coming from inside the FHH-net or not. This allows the conclusion that approximately a third of all requests on the transparency portal originate from within networks connected to the FHH-net. This leads to the assumption that the data objects availability on the transparency portal are somehow of value to the administration. However, all data on the transparency portal are originally from within the FHH network – meaning there should not be need to resort to the transparency portal as there are internal systems for accessing those data. What is unclear as of now, who this exactly encompasses. For example, other meta-portals such as GovData, as the federal open data portal for Germany, harvest the data on the transparency portal in Hamburg. It is not clear to the author to what degree this might impact the amount of requests logged.

Secondly, the unknowns about the user, that is the lack of a feedback-loop between data

provider, transparency portal and data user is impeding the further development of the transparency portal as such because it is not known what general demands are held, or what specific improvements are needed to facilitate a specific use case. The use case is simply not known. Further then, this leads to a practical operational problem in the mid- to long-term term. Not being able to show how and why the portal is used, and what practical outcome this entails, it becomes increasingly difficult to justify funding to develop it further.

Turning toward the actual analysis, table 3 displays the final themes, their relation to the topics under study, and a brief synopsis of their central organising concept.

Table 3. *Theme Summary Table*

Relation	**Theme**	**Central Organising Concept**
Motivation	Transparency Precedes Trust and Inclusion	The perception that the use of OGD, or inherently in having OGD at one's disposal, enables and contributes to trust in public institutions as it is a fundamental ingredient for any form of inclusion.
Motivation	Operational Utility	The notion that use of OGD is either useful or necessary to accomplish specific goal, due to OGD inherent characteristic of not being possible to be procured from anywhere else.
Interaction	Multiple Sources Of OGD	The distinct view that OGD will be procured from multiple sources and via different media, depending on the type of information needed.
Effects	Public Empowerment	The expectation that the proliferation of OGD will have a positive impact on the capacity of the public to be better informed.

5.1 Motivation

To recap, the particular research question aimed to address in this part of the analysis is SRQ1.

> *What are the users' motivations behind their OGD utilisation?*

The initial code labels for this part of the thematic analysis were composed of Davies'

motivational clusters (the 'Name' column of table 2) and the subsequently listed motives and factors synthesised from [25], [88] (extrinsic, intrinsic, social, political, economical, technical and political motives or factors). The expectation was that those keywords were sensible in application to the exercise, and would lead to clusters of motivations as the first step to analyse motives. This expectation was largely fulfilled.

Transparency Precedes Trust and Inclusion

This theme's conceptual core is the matter of transparency of the public sector at large – mostly meaning public agencies, the administration as an institution, and the ministries as the government. Transparency here should be understood as a willingness to provide simply information, or quantitative data, that would allow the user to be able gain insight into a given topic, process or decision. This form of transparency, which could be abstractly referred to willingness, seemed to be seen as inherently positive and therefore desirable because it allows the user to contribute to a certain capacity for empowerment of the citizenry or the voter. Although the specifics of this perception, that is the way this view was being articulated, varied, this notion formed a clear pattern throughout almost the whole data set.

The first observation that could be made was that a clear predominance of social and political motives prevailed. From within those code-clusters, any type of direct economic motives was absent. Moreover, no interviewee expressed the intention to provide a specific technological service as a public service to the public sector as a result from using OGD. Thus, it did not seem to be valid to speak of co-produced public services in order to obtain a financial gain. Although the intention of actually creating some service did surface, however the motives for which were not of the economical type. Those underlying motives were almost exclusively clustered under social and political motives and factors, or 'government focus'.

To reify the notions of social and political motivations: Social motives were coded as anything pertaining to the abstract greater good of society. That is those motives or drivers which were closely in line with the findings from [25], [88], where the benefit of others (as the society at large) was within consideration. For political motives a specific concern with the functioning of the public sector regarding efficiency and effectivity was coded, along with matters relating to trust and general political interest. What became increasingly clear is that those two motivational aspects could not really be detached from another in this specific setting of exploring open *government* data. Where one is interested in promoting transparency and collaboration opportunities with the administration, the motives are, almost automatically, to be seen as both political and social. Conversely, where one is

interested in promoting a better informed citizenry, when concerned with public budgets for example, then this is also automatically both socially and politically motivated interest. Therefore, it did not seem to be sensible to clearly distinguish, or try to find and craft distinctions, between either politically or socially motivated interests.

Loosely spoken, across the whole data set, the latent interpretation of the data always led to some form of value proposition arising from the availability of OGD in connection with transparency. Either influencing the effectiveness (i.e. the perceived correctness of an outcome) of a decision-making capacity of public bodies (or the executive part of the notion of 'government') with the inclusion of the public, or the intention to have a better informed public. The value 'transparency', if it can be seen as such, was prevalent in connection with being worth something in and of itself and as a mean to further utilisation. Mostly, this value was in connection with being able to have more information or data to work with – particularly in regards to being able to inform others about certain aspect of the government or the administration. While the specifics where expressed differently, the end goal was almost always very similar.

For example, while considering what the fundamental impetuses to engage in OGD are, one interviewee sees how transparency of the polity is seen as fundamental to enable participation in decision-making:

> "Well, in principle, it's about enabling the public to participate in the political process at various levels[1]. And for that, the state simply has to be transparent, so that people can firstly understand what is actually happening. And secondly, they can exert their influence at specific points. If I don't know what is actually happening, what decisions are being made, then I can't participate."

From the civil advocacy standpoint, elaborating on underlying motives to come to a more inclusive democracy:

> "In other words, more transparency, more participation, more debate, more dialogue, so that we can arrive at viable, sustainable solutions and results that can be used to strengthen trust in politics and in the administration."

From a volunteer-standpoint, who engaged with budgeting data in order to make the budget more accessible and understandable for anyone interested gave a very concise, but highly elucidating comment on why the volunteer did what he did:

[1] The interviewee refers to federalist structures here.

> "Precisely. To empower the citizens to understand it, if they want."

From the parliamentarian perspective, a clear socio-political motive arose. Although OGD were used from a utilitarian standpoint (see below), the motives for engagement stem from value-perception for the society in relation to transparency and communication between the public and the public sector:

> "As a politician, I would say that my main job is not so much the question of why to use data in this form, but why to make it usable. So that is the political job or the mandate. To make data publicly accessible and that is, so to speak, that is my motivation. To achieve that, that's my job. Making it usable because it ensures a better, faster flow of information, because it ensures comparability of data. Transparency is a value in itself, in my eyes, because it promotes trust and makes it easier to understand decisions, but of course it also enables an exchange between governmental and non-governmental organisations."

Operational Utility

This theme should be seen as a sibling theme to the former. A distinctive pattern of practical utility of OGD could be seen. As obvious as it may seem that such a pattern would occur, there is an interesting aspect to it. At the core of this theme is the inherent value of OGD in coming from public bodies. This means that OGD seems to have an inherent value of being a data source that cannot be replaced with other sources. This characteristic seems to be another underlying driver, next to the more moral importance of having transparency as such.

This pattern was visible through interspersed code labels gathered under problem solving, extrinsic, and in few cases technical factors. Extrinsic factors relate to motivations that directly relate to the outcome of an activity, not the activity as such, while technical factors relate to data quality and constituency. Problem solving was associated with being able to use OGD to overcome specific challenges. Still, this property is somehow inherently linked to the transparency concept due to the data's origin.

For example, the perspective from a journalist aiming to fulfil his obligations towards his employer and the public at the same time:

> "When we want to, and have to, and should report on the development of the corona-pandemic, then we need data for that. And in specific areas these can

only come from the public sector, so that we can satisfy the public's need for information."

From the viewpoint of a politician the sheer availability of OGD through the Hamburg Transparency law translates into practical usefulness as much as into a morally valuable:

"The verifiability of the statements of the administration, for example. Does that, what we have been presented with, correspond at all to what is then also found in the transparency portal, is it congruent? That has been a benefit and the fact that I don't have to do it at all, but the sheer fact that it is verifiable, has of course led to the reliability of the administration's statements being much, yes, much higher. So it was no longer the case that just anything was said, because in case of doubt I can check something like that. I think that this verifiability is of high value."

From a city-planner perspective, the quality aspect of the data was primarily important. However, it must be noted that those cadastral data are centrally kept and available only through governmental entities when it comes to have a whole metropolis equally covered. Here, the motive for using OGD was a matter of a professional challenge combined with high quality data:

"A large construction company approached us and asked if there was a possibility for us to look for potential building sites in the city of Hamburg, within the city limits, so to speak, for possibilities for increasing the density of residential properties [...] and there are of course opportunities for a project developer, if he finds a good plot of land, to put a residential building on it. To put it simply: [We asked ourselves] can the whole thing be done systematically somehow at the city level, so that you don't just say you're looking at individual areas. [...] At the beginning, it wasn't quite clear that you could actually do this via geodata, via this open data, which Hamburg also makes extremely available, that you can actually do this very well because the city of Hamburg and the transparency portal simply make an incredible amount of information available that you can then put together. So in Hamburg there is [the data availability and quality very high] so that is really very good again compared to other federal states."

To put both of these themes into perspective, it is necessary to point out that various other other reasons have been either directly stated or could be easily interpreted into the data within a latent analysis. However, those aspects have not been prevalent and strong enough across the whole data for any additional themes. This does indeed echo findings from the literature. There is a broad set of motivational factors at play, and any attempt at generalising about them does seem inadequate. What seems to be the case is that certain democratic values seem to be a foundational, underlying factor in the engagement with OGD. Extrinsic and economic factors do play a role but a subordinate one. What has to be taken into account in this respect, is that the data are actually meaningful for the sake of fostering inclusion in anything – be it deliberation, decision-making, or simply being kept in an information loop.

What must also be accounted for is: Who is actually vocal about their engagement with OGD? While it being clear that finding OGD users as such, even ones affiliated with any organisation, is already not an easy task, those who would want to advance the topic in general are ought to be expected to be more vocal and in praise of OGD. For the context in the FHH this does not mean that democratic values and the intent to promote OGD are the key motive for people to engage with OGD. What it does mean is that there are organisations and people within civil society ready and expecting to work and collaborate with the public sector on the basis of OGD in which they see high value.

Regarding the conceptual framework for this part of the analysis, it is evident that a true explanatory framework for motivation is non-existent. Davies' motivational clusters quickly appeared to be a quite narrow frame to fit motives in. Which doesn't really invalidate the framework, but rather exposes a different perspective on utilisation in the UK. For an exploratory study to structure thinking about motives, it seemed almost distracting more than stimulating.

However, clustering the motivation-related codes into extrinsic, intrinsic, social, political, economical, technical and political motives or factors prove to be much more nuanced, and provided the necessary analytical navigation-space to collate the code labels and understand the data items. What could also be made out was that every coded data item was somehow fit for one or more categories, which made it possible to to view and analyse different motives in different light. But it also lead to a large initial overlap of motivations, which needed to be atomized to be analytically clearer.

Concluding for the motivational part of the analysis it can be said that the findings from the literature are largely reflected. Motivational underpinnings seem to vary: a wider range have been articulated in some form. However, in this particular data set the value of

transparency for the benefit of trust and inclusion was most prominent. Therefore, socio-political motives as underlying for the engagement with OGD are empirically verified in the context of the FHH. These include motives and drivers expressed as wanting to strengthen trust in governance, wanting to better inform citizens (or have a better informed citizenry) or to enable inclusion or participation in an abstract form.

From a theoretical perspective the framework used was valuable to initially describe motivations but it does not seem to have strong explanatory power of why users want to use OGD. However, for exploring the topic it seemed supportive.

A final observation must also be pointed out: Many interviewees had some kind of origin story or anecdote to offer when inquired about their own motivations and drivers for engaging with OGD. This resulted in a temporary theme called *Problem Orientation* in relation to interaction with OGD. The theme was retired but formed the basis for the theme for interaction. The reason was that by saying that some problem was at the basis of engaging with OGD would have been too easily formulated and also not true. Those stories were not always clearly a fundamental motivating factor but rather a medium of explanation and illustration.

5.2 Interaction

This part of the analysis is concerned with the following research questions:

SRQ2: *How can the actual usage be described?*
SRQ3: *What skills and tools are employed?*

For this angle of the analysis, Davies' Utilisation Processes (see figure 4), and the dichotomy of synthetic versus analytical knowledge-creation were used as initial categories to cluster code labels. The expectation was that those would help in structuring the data and lead to apparent patterns in use. This expectation was only met partially. The Utilisation Processes were helpful to cluster actual use instances into degrees of complexity of use. This helped to structure the data immensely but did not lead to any sensible patterns, thus not leading the analysis anywhere. Further, the classification into either synthetic or analytic knowledge-creation did not seem appropriate as the analysis progressed. The reason for that was the impossibility to trace the final knowledge-creation to the actual use of OGD within the use context. Here, the analysis has to be supplemented with some reflexive introspection.

It was the authors assumption that, in the context of the FHH, OGD would be those data procured from the transparency portal. Afterall, the Hamburg Transparency Law is the fundament on which the publication of OGD is based on. Based on those data, some kind of work would have been done in the pursuit of some goal. This was true in most cases, and therefore a valid assumption but it is not close to the real picture at all. At the core of the assumption lay that OGD can only come from a, or one, Open Government Data portal. Although the definition exposed in 3.1 does not necessarily imply this assumption, this perception was prevalent during most of the time of this research. However, what the data items clearly suggest is that, from the perspective of the users, there is no such clear distinction. OGD can come from many places. As a matter of fact, the transparency portal seems to be just one outlet, or source that is used when a user is being inquired about their actual use of OGD. What is not immediately apparent is why this view is being held. A quite personal suspicion is that it could be related to certain semantics. While the international literature mostly uses the term *Open Government Data*, in German the term used to refer to the data from the transparency portal is *Offene Verwaltungsdaten*. This translates literally into *Open Administrative Data*[2]. On the other hand there is the argument of pragmatism. To reach a given goal, data are procured from wherever possible to draw conclusions. Those might be Web pages of MPs, specific agencies, the Senate of Hamburg, or any other source that seem suitable in that case. This led to the realisation that one concrete pattern would revolve around the multiple sources used when elaborating the interaction with OGD.

Multiple Source Of Open Government Data

At the core of this theme is the issue of where OGD is procured from, which also relates to the perception of what constitutes OGD in the eyes of the user. Most interviewees exposed the fact that the use of the transparency portal, which was prevalent, only constitutes one instance of all of the possible sources of information used. This means that the data indicated that the engagement in OGD, in the cases observed, entailed an integration of data from multiple source in order to pursue a given objective.

To begin with it is sensible to contrast the transparency portal to another, frequently mentioned source – the parliamentary documentation. To highlight the difference in those two sources:

> "No, they are actually two different worlds of information, you could almost say. So I visit the parliamentary database [i.e. parliamentary documentation] when I know that the parliament has discussed a topic or I assume: the par-

[2]To verify: Common online translation engines will most likely reproduce this translation.

> liament must have dealt with the topic, it can't be otherwise, then I go to the parliamentary database and know what the parliamentary processes are. I can ask for the relevant keywords, whether it must have been a committee meeting or whether it must have been a parliamentary debate or whether I am looking for a verbatim record perhaps of a debate, which is sometimes quite important, then I know, okay, I can find that in the parliamentary documentation. But when it comes to finding the Senate documents, the documents of the administration, I can't find them in the parliamentary database. So there are two different levels."

What this excerpt highlights is the fact that not all information necessary is procured from the transparency portal when considering OGD in a wider sense. Meaning that the transparency portal alone does not seem to suffice in obtaining information related to the *government*[3]. Which raises the question: What is included in a particular users' perception of OGD?

While elaborating on the daily activities involved into maintaining and operating the site Abgeordnetenwatch the interviewee expounded:

> "So everything I have just mentioned actually comes from the website of the Hamburg Parliament. Either from the protocol data or from the MP profiles published there. Sometimes you have to google the MPs to find their contact email addresses or something similar. These are very specific cases. And every time there is a parliamentary election, the state election administration is added as a data source, from which we obtain the data of the candidates and make it available. [...] The parliamentary documentation is included. And apart from that, there are also profiles of MPs and an overview of parliamentary groups. I assume that the parliament itself is the operator of the site. But I don't know exactly what the responsibilities are."

During these discussions about the daily use, that is those processes in which OGD were employed from the perspective of the user, it became obvious that various Web Sites related to governmental institutions, such as the parliament or the general Hamburg City Web Site were manually scraped in order to supplement data otherwise obtained from the transparency portal. This could imply the circumstance that the transparency portal is

[3]It is necessary to emphasise here that there might be some difference in interpretation of what the term 'government' includes. To illustrate: How to delineate the terms 'state', 'government', 'executive branch', 'administration' from one another?

simply just for a specific kind of OGD, relating mostly to the Senate, and the administration as such. Or that OGD from the user perspective, is simply all data related to the government, regardless of whether or not it stems from an open data portal.

In a completely different use-case called SmartOpenHamburg, in which OGD played a role to effectively simulate traffic and transportation flows in the city via agent-based modelling, the data procurement is fundamentally different. But the theme is the same: Data from the transparency portal constitute only a part of the data used. Other data were procured through any means necessary, e.g.: Open-Street-Maps. Among the sources was another data portal run by the city – the Urban Data Platform, and a dedicated access to specific IoT-data from a public agency. When inquired about where the data used are from the interviewee explained:

> "Well, that is of course very variable. So what is relevant for Hamburg, when you do spacio-temporal simulations like we do, is that the Urban Data Hub is extremely relevant. [...] We work with them and of course they have a huge amount of interesting data that we can use more or less directly in our models, so to speak."

Later when asked about the amount of data used from the Urban Data Hub, the interviewee illustrated the quantitive ratios of the data above used for the whole project:

> "No, it's actually so that I would say, so 50% or below. Because, the default dataset comes from Open Street Maps. So the demographic data, something like how many lanes a street has, where there are traffic lights and so on, so the things that are not necessarily freely available in open street maps. That's something we get from the portal. [...] You have in all the projects that we have in this kind of way, you have a significant data integration, so from many different sources in principle."

In sum, this theme is illustrating the fact that when OGD are used, it is hardly ever utilised by itself. It seems that using OGD means trying to obtain information from various sources in order to integrate those for the purpose of reaching a goal or objective. A common occurence was that during the interviews the term *Open Government Data* was hardly ever used by the interviewees, although the German language does make heavy use of anglicisms. Instead, the data used where simply referred to as *open data*, *ministerial data*, *data of the public hand* or *open data of the city*. This raises the impression that during the

use of publicly available open data for some purpose, data related to the administration or the government or otherwise were also used but only played a subordinate role in the whole picture.

Other than this shared pattern, no further common themes, that were pervasive enough, could be synthesised across the data that would give insight into usage patterns. From this fact can be inferred that the use processes varied dramatically in kind, depending on the particular user and use case. What is meant by this absence of a common theme regarding the complexity of use is that across the whole data set different use cases were coded against all of the processes from figure 4. For example: For interviewees 2 and 4 applied only data extraction and data reporting. Here, data from the transparency portal, and other sources, were used for mainly fact-finding and the occasional processing of numbers perhaps. For interviewees 5 and 8, in addition to the previous processes, also data-to-interface would apply as the data needed to be cleaned, was combined with other data, and involved writing custom code in order to work with it. Further, for interviewees 6 and 9 the processes data-to-fact, data-to-information and data-to-service could apply. In these use cases the data were used to provide a complete new service. On the one hand, this was the multi-agent simulation mentioned above, on the other hand it was an internal service for a city planning office in order to more effectively find potential development sites. All of these use cases are inherently different that it did not seem valid to develop any type of theme around them.

From this realisation also stems the absence of a concrete pattern of skills and tools used for the data processing. Those mentioned were specifically related to the particular use case. While a range of software tools were mentioned, only a couple concrete ones occurred more than once across the data items, thus could not contribute in any significant matter to the analysis. The two tools that were mentioned briefly by three interviewees were the general-purpose scripting language python, and a data visualisation tool called data wrapper.

One interesting pattern that did start to become viable to become a theme around skills needed and employed needs to be pointed out however. Three interviewees mentioned that a certain basic knowledge around handling administration in general, administration specific speech or basic political science knowledge is helpful in order to effectively deal with OGD. However, all three interviewees talked about a different data sources but still referred to the same concept. The first example is referring directly to the transparency portal:

> “Well, for me it’s actually the case that I have an administrative background

> and that helps me a lot, because I know the technical terms just for finding data. So, for example, when I enter in the transparency portal: What does the city spend money on? With luck, you still get to the budget, but if you enter *Doppelhaushalt 2021-2022*, you get that immediately. That's where it starts. Access to data is simply becoming easier – knowing administration specific terms are totally useful. Or, for example, OK, how can I communicate with them? That sounds very, very basic, but in fact it's often a case of saying, hey, you can write them an email. Then you briefly look up the authority and then the organisation chart and see where you can get the right position. That helps quite a bit, that's what always helped me, when you understand the administration a bit."

One interviewee mentioned this immediately when asked for useful skills in order to work with OGD in connection with the parliamentary documentation:

> "You actually have to know how things work in parliament, what decisions MPs make, how it actually works, when there are roll-call votes, how it works with motions and so on. Of course, you need a certain basic knowledge of political science to be able to understand it well, let's put it that way. It's supposed to be generally comprehensible, but you have to work your way into it. For example if we now have new people in the team, you can't just say, "You bring the vote from Hamburg onto the page." You first have to explain to people where they can actually find it, what does it mean, how can you understand the whole thing?"

The third excerpt is from the city planner in reference to a missing oblique aerial photo on the Geoportal of the Metropolitan Area Of Hamburg:

> "But it was somehow offline in between and then I wrote to them briefly, then they took it online again straight away, which I thought was kind of cool that it happened so quickly. I was actually positively surprised again. So, on the whole, it worked well most of the time. So I think you have to know your way around a bit. From my professional background, I am used to communicating with various authorities and I always make sure that I get in touch with the right person. That's important, because if you get the wrong one, sometimes it's a dead end or something. That's why I can also imagine an experience that

> is different, that you simply run into a wall or something. So I wouldn't want to rule that out now."

Another interviewee response, but only after being prompted by being asked about the necessity to be able to comprehend administration and politics-specific speech:

> "This is very difficult for me to assess because I have been moving in these contexts for so long that I probably can't see the wood for the trees any more. But if I were allowed to guess, I would say that I would strongly assume so, that it is the case, yes."

Although this notion was expressed three times unprompted, it was not pervasive enough to be actually considered a theme across most parts of the data set.

Regarding the conceptual frameworks used to view this part of the analysis: The different processes of data use from figure 4 were very helpful and sensible to structure the data albeit only to initially serve as a ground for any further consideration. From there on, after multiple rounds of coding, the codes became highly granular and consequently became almost unworkable as they were scattered along all processes. This seemed to be due to all together quite different use cases of each interviewee. This leads to the inference that OGD utilisation in practice is an extraordinarily heterogeneous activity. In retrospective this framework seems useful to describe concrete singular OGD projects as a starting point to further dissect them and to compare them with one another. To classify general OGD engagement into those categories does not seem plausible. The reason for that seems to be that the interviewees from this study were referring to multiple different projects in different times. If the scope would have been specific data sets, or specific services, then the processes would be of better use. Similar logic would apply to the dichotomy between analytic and synthetic knowledge creation. In the early coding phases, this classification made sense. But the more it became obvious that the engagements would be heterogeneous, the less meaningful this classification did seem. In order for those two frameworks to be sensible for a thematic analysis, the data would need to be structured around a specific type of data, or a specific data set – or alternatively a specific type of service or application based on open data.

Regarding the ODI framework from figure 2 the data obtained from the interviewees did hardly overlap. The main reason for this was that virtually all interviewees seemed to find it hard to articulate specific soft-skill necessary to work with data. The complication here

was that the interviewer did not want to try to give examples of skills in the framework in order to counteract response bias. What followed were in some cases the need for basic political science skills, or administration-domain-specific knowledge. This points again to the heterogeneity of OGD in the wider sense. OGD may include free-flow parliamentary speech records, minutes from a meeting, draft bills, vaccination data, oblique areal photos, maps, engineering blue prints to name a few.

5.3 Effects

The final research question under analysis is the fourth one.

> SRQ4: *What are the users' desired effects of their OGD utilisation?*

For this part of the analysis the initial codes derived from the self-synthesised framework from figure 5 were used. The codes were classified belonging into one of four possible configurations: Either of an economic type of effect internal to the public administration or externally toward the public at large. Or of a socio-political type of effect internally to the public administration and its agencies or the public at large. The expectation was to be able to structure the code labels into those categories for the initial set-up for the analysis. This expectation was largely fulfilled. Similarly to the motivation related analysis, an apparent pattern towards the socio-political types of effects became visible. Hence, most of the positive effects, expected or desired, clustered around a better informed citizenry, increased decision-making capacity generally, more trust in government and other similar and typically democratic values.

Public Empowerment

The core concept of this theme revolves around the notion that the proliferation of OGD would lead to a better informed citizenry. This high-level notion was expressed across almost the whole dataset, although more often than not expressed quite latently. Most interviewees expressed a certain potential arising from the availability and use of OGD that would lead to an empowered society through various means. Among those were the ability to tap into a certain collective intelligence of the public by viewing OGD as an abstract generic knowledge resource that should be available to the public at large, which then would lead so far unidentified, abstract effects:

> "I think the greatest potential is, first of all, in the abstract, to give society the freedom to work with it. Just as we as a society provide ourselves with

> an infrastructure: Railways, roads, water, whatever we need, in the 21st century data is also part of it, especially the data of the public. And thanks to ever more technical possibilities, to use a buzzword: big data, there is simply also potential and possibilities for solutions that we have not thought of before or could not think of before, because it would simply have been far too time-consuming to bring together different data."

In the same vein of thought, but with a more concrete vision of what would be achieved through the proliferation of OGD, an interviewee described a tangible potential effect that could be achieved:

> "That we are able to make a better assessment of our environment, so to speak. I believe that this also helps to communicate what kind of condition our environment is in, for example. So it's quite different when I see a map showing where, for example, the groundwater levels in Germany have declined, then I can look at my neighbourhood and say, okay, this will be fine for another 10 years and then it will be difficult to get water here. This simply means that these geospatial data in particular have the possibility to be visualised, so that it is really impressive. And I think that helps us. So at least for those who are visually influenced, I think that's an important aspect."

Next to those kind of latent notions of an empowered public, one interviewee expressed quite semantically that the expectation is to feel more empowered, and to be able to see eye-to-eye with the governing structures.

> "Personally, if I had access, I would feel more independent, I would simply feel more powerful to exchange ideas with whomever about important things in our society, i.e. with my neighbours, with my friends, in my networks and with the administration and with politics. Quite simply, we have the same basis and everyone has access to the same box with the same rights, I don't want to modify them, I don't want to go that far either, but I want to have access, equal access, and that is then the common basis for all of us, for me that is actually the normal state, the desired normal state. We all have access to the same or up-to-date information and can openly talk to each other on this basis without any complications, without any secrets, without any part being covered up."

The same notion appeared also from a journalistic perspective, however with the implicit intent to use OGD to both inform readers (i.e. citizens) and to be part of a check-and-balances system toward the government:

> "Such an overarching goal is more a social one. As a journalist you would have a practical benefit from [OGD], simply the greatest possible transparency, the simplest and most structured access to data. You realise this when you look at the Corona crisis, for example, and realise that depending on which federal state I'm looking at, I have different levels of access to the situation, how is the pandemic, how many people are dying, how widespread is the vaccination and so on. For every detailed question you can think of, you are actually bound to the data access or to the open data that existed then. It's very difficult to just generate other data and we've just realised that the gold standard is to have data that is as granular as possible, free, timely, in technically well available formats and so on, then I as a journalist can do everything I want, which means both describing the situation as well as possible, thus fulfilling the need for information for the readers, and also, in case of doubt, being able to perform a control function, do we actually look at the right metrics in politics when we now decide which measures to implement."

Lastly, from a parliamentarian perspective, the long-term effect desired is transparency with which the public shall be empowered to participate, in the most abstract sense possible. But also a certain value for public service delivery seems present:

> "So I think the main benefit is that transparency leads to trust. And transparency is a form of... yes, of communication and that in turn simply strengthens our democracy. Trust in the state and in state action. And I think that is the greatest benefit. Apart from aspects that relate to flexibility and speed and responsiveness. So simply registering a car no longer takes a day or you have to apply for it, but you can do it online or something. These are aspects that have a real benefit for the individual, but in a larger and philosophical sense, I would say it's about creating the possibility that the citizens' perception is that I have a part in what's going on. I can participate."

This theme shows that potential users are viewing the availability of OGD in general in a good light. It seems as though a potential for the greater good of the society is prevalent. The various effects that are desired could definitely be attributed to socio-political types

of effects and thus pertaining to the public at large. Interestingly, almost no mention was made explicitly that it could also have possible positive effects on the public administration as such. Either performance wise or financial. Similarly only one interviewee made it clear that, for his profession of city planning, there would definitely be potential performance increases on an organisational level. All other hints to desired effects, either for the public administration internally, or for the economy broadly were either side-notes or latent interpretations. Thus, it should be interpreted as a general awareness of possible positive effects for the public administration and for the economy at large, but immediate perception of socio-political values seemed to prevail.

In relation to the literature, this sentiment is also reflected. While predominant effects are hypothesised to be of the good governance (or socio-political type) this seems to be true in this data set as well. It is expected that due to increased release of OGD more useful information is around to be worked with for various goals. Mainly the sentiment of a better informed public seemed to be dominant. However, it does not seem appropriate to hypothesise where concrete effects would occur. This is in the most part because usage seems highly heterogeneous and not particularly organised across the interviewees.

Regarding the framework used for this part of the analysis: It seemed plausible to think of possible effects to materialise either in form of a clear economic type or of socio-political type of effects. The logic for this distinction was the convergence of hypothesised effects from the literature in the concept of good governance (see figure 1 and section 3.2). However, this resulted in most of the possibly observed effects to be classified as socio-political, which could be viewed as a certain bias. To structure the data however, this seemed still helpful. Regarding the abstract location of impact: A differentiation between seeing effects either internally or externally to the public administration or the government seemed less plausible. Regarding the economic axis: OGD are not supposed to be sold for profit to anyone. However, direct access to up-to-date data in some cases are feasible as per the preliminary interview. Some entities may enter into contractual obligations with certain data providers to have more timely data, and a direct access to them via an API for example. However, those instances were not encountered in this study. In regards to the socio-political type of effect it was not immediately clear how socio-political types of effects could be either internally or externally to the public administration: Possible performance gains (an internal effect) would have potential spillover effects, for example. From this perspective internal and external socio-political effects could not be detached from another.

6 Conclusion

The point of this study was to explore the case of OGD utilisation in a defined context – that of the FHH. Further, the objective was to thematically analyse empirical data, gathered first hand from actual users of OGD. Morever, the extended objective was to evaluate the validity of current theory used to describe or explain OGD utilisation.

The motivation and rationale for this project was to tackle the issue of the elusive user and what he or she would actually do with OGD. This issue has been identified as prevalent both in the literature and in the context of the FHH. The question remains what this study has to say about the main research question:

> *How and why do users interact with OGD in the FHH?*

To begin with it must be reiterated what precisely should be considered OGD from the perspective of the user, which will provide insight about how OGD is being used. The author's preconceived notion was that OGD are retrieved from the respective OGD portal in the city – the so called transparency portal. From there, the data are re-used in a way, to produce some sort of output for some purpose. While this does not seem to be incorrect in and of itself, it is only a part of the observed behaviour. When trying to investigate the issue of what OGD is being used for it seems more appropriate not to ask oneself *What is being done with OGD?* but rather *What role does or can OGD play in the use and re-use of Open Data at large?* The data collected strongly suggests that, in the context of the FHH, the transparency portal is but just one source of OGD. During the research it became evident that multiple portals exist, for apparently different purposes. To itemise those that were encountered during the interviews:

- Transparency Portal
- Parlamentary Documentation
- Geoportal for the Metropolitan Region of Hamburg
- Meta-Ver

Moreover, further data have been retrieved from respective sites from ministries and agencies in order to contextualise or supplement data, such as the Web Site of the parliament

or the official city portal. This fact alone turns the issue of OGD re-use into a highly complex and heterogeneous phenomenon. What seems to follow from it as well, is that interaction with OGD can only be uniformly described at a very high level of abstraction. Conversely, this means also that focusing on the transparency portal alone is seemingly implausible, to give an account of OGD re-use in the FHH. It must be realised that OGD users should not be defined over their use of the transparency portal alone. Therefore, it can be concluded that the concept of OGD, in particular it's re-use, is indeed a diffuse, slippery, hard to grasp concept. While this does not invalidate the technical definition from 3.1, it must be realised that open government data in practice should not be viewed as those data which are available through a searchable portal. At least not from the actual users' perspective. If this turned out to be true in general, this could imply serious organisational concerns for decision-makers in any given context. For the FHH, it seems to raise the question of how to orchestrate OGD provision across public institutions in order to effectively distribute and make data available for anyone wanting to reuse it.

Due to this heterogeneity of use a definite thematic-analytic statement of skills necessary for re-use could not be synthesised. While it seems plausible that skills necessary to work with OGD depend strongly on the kind of data procured, on the actual outcome pursued, as well as on the complexity of use. For instance, quantitative data procured from an IoT-network for the purpose of integration into a multi-agent simulation system require a completely different kind of skill set compared with the operation of a Web site gathering data on roll-called voting behaviour or MPs, or compared with the visualisation of budgeting data, or compared with political activism. What seems to be the case though, as pointed out in section 5.2, that a certain need for understanding of administrative processes and administrative specific terms is necessary in order to effectively use different kinds of OGD.

Regarding the question why OGD is being used the results are clearer. The thematic analysis of the data items showed that those users interviewed were fundamentally motivated by socio-political factors, outweighing extrinsic and economic or financial motives by and large. This means that the underlying impetuses reflected the intention to contribute to the greater good of the community or society on the basis that more transparency is a good thing. Moreover, the data showed that most of the interviewees were convinced that their use of OGD could lead to more or improved inclusion into the affairs of the state. Either passively through insights into budgeting or votes in parliament, or actively through political activism in which the OGD were seen as a authoritative source to come into contact with the administration. What this does not mean, is that those motives were isolated from other types. Intrinsic motives such as the sheer fun of playing with data for the sake of exploration and finding patterns, or the opportunity to find like minded people

interested in similar issues were mentioned occasionally. These findings are unreasonable to compare to the findings from 4.3, as the evaluation report [96] exclusively referred to the transparency portal, and the Hamburg Transparency Law.

With respect to the desired effects of their respective utilisation, a certain overlap with motives could be synthesised. In hindsight this seems obvious as a certain outcome of an engagement with OGD can be seen as the initial impetus for the engagement in the first place. In any case, the thematic analysis showed that effects are desired, much in line with motivations, in socio-political and good governance spheres. Most interviewees viewed the potential of the release of OGD as beneficial for the society at large. OGD would pose as a certain generic knowledge resource which could be used to realise a better informed society about topicalities of public affairs. Moreover, a certain potential to tap into a collective intelligence could be observed. Thus particular effects of OGD should be viewed as holistically applying to the general benefit of the community. However, in line with findings from the literature, there seems to be explicit value in the use of cadastral data often used in the context of city planning.

In regards to theoretical models to describe or explain OGD re-use: It is clear that a precise model for explanatory purposed for such a complex social phenomenon will be hard to come by. Generally speaking, motivational factors derived from open innovation studies, such as extrinsic, intrinsic or social motives do make sense to structure data for the purpose of a thematic analysis. However useful for structuring at a high level of abstraction, such motives may serve to describe and cluster latent motives. Given the heterogeneity of OGD use, as was the case in this study, general levels of complexity may also be seen as useful for structuring purposes. However, without precise, validated and causal evidence of what outcomes have been realised due to the re-use of which kind of data, such models seem to not be able to serve for more than descriptive purposes.

6.1 Limitations and critical reflection

As this study aimed to explore a specific social phenomenon in a given context, some degree of uncertainty was expected. Given the circumstance that relatively little is known about the users of OGD and their behaviour, it was also expected that some effort were required to obtain enough data in order to proceed with the project, although initially promising leads were present. During the course of the project, it became increasingly difficult to obtain fitting interviewee candidates for various reasons. The requirements were loosened accordingly: A candidate must have worked with, or ceased to work with, in his or her eyes, with OGD in the FHH. And obviously willing to talk about it. It is perhaps for this reason that the interviewees had very little in common other that that they have

indeed used the transparency portal (or other sources) at least at one point, were interested in the use and proliferation of OGD, and had a fundamentally positive attitude towards the release of OGD. This circumstance must be pointed out as a kind of implicit response bias – possibly skewing the final list of interviewees towards being in generally pro-OGD.

Further it must be pointed out that a key component for case study research is the triangulation of findings, via either data, method, investigator, or theory triangulation. While it was anticipated to integrate data triangulation via a thorough literature review and with intense use of theoretical concepts, this could barely be achieved. The circumstance that little is known about the actual users and their behaviour made this especially difficult.

The data generation and analysis methods must also be addressed. The author was initially convinced of his interpersonal skills to be sufficient in order to conduct appropriate semi-structured interviews. However, as it turns out, interviewing is an extremely challenging task. The frequent revision of the interviews made it clear. One the one hand is the transfer from German to English language. Due to a certain lack of exercise in the author's native language, German, the interviews were in parts riddled with interjections, while grasping for the right terms in German, for example. On the other hand is the balancing act, between trying to steer the interview into the desired direction according to the topic at hand, without trying to place possible answers into the heads of the interviewees, while at the same time managing the natural temptation to let the interviewees explain their thoughts for the sake of personal interest and possible enriching insights, while simultaneously managing time.

Lastly, the topic of data analysis. To analyse qualitative data with a thematic analysis also proved to be a considerably challenge, seemingly against prevailing opinion. This may have been simply a skill issue or pertaining to time pressure at one point, or due to nature of the data collected. What seems most likely in hindsight, is that in dealing with quite intangible and soft topics such as underlying motivations and desired effects of a quite heterogeneous topic such as OGD use turned out to be, is quite difficult.

6.2 Recommendation for future research

Specifically regarding the context of a federal state like Germany, it seems plausible to restrict the context of a study to one political entity, like this study did. In this way important legislative factors, such as transparency laws and other publishing guidelines can be controlled for. Within such context, it seems necessary to focus further on one particular type or kind of data for a specific use case. For example how are transport or cadastral data used and what specific outcomes have been realised, causing what effect.

For the FHH in particular: Highly interesting and useful seems to be to find out what role the transparency portal plays in connection with user-requests originating from within the public administration. As pointed out, approximately one-third of requests on the portal originate from within the administration for so far unclear reasons. Unfortunately, this research project could not obtain any interviewees from within the administrative sphere. It seems plausible that significant value for the Hamburg Transparency Law, and its transparency portal, may be realised by clearing up the value proposition toward the public administration.

Bibliography

[1] E. H. Ruijer and E. Martinius, "Researching the democratic impact of open government data: A systematic literature review," *Information Polity*, vol. 22, no. 4, pp. 233–250, Nov. 2017. DOI: 10.3233/ip-170413.

[2] N. Gonzálvez-Gallego, L. Nieto-Torrejón, and M. C. Pérez-Cárceles, "Is open data an enabler for trust? exploring the link and the mediating role of citizen satisfaction," *International Journal of Public Administration*, vol. 43, no. 14, pp. 1218–1227, Sep. 2019. DOI: 10.1080/01900692.2019.1668412.

[3] M. Chui, D. Farrell, and K. Jackson, "How government can promote open data and help unleash over $3 trillion in economic value," *Innovation in Local Government: Open Data and Information Technology*, vol. 2, 2014. [Online]. Available: https://www.mckinsey.com/industries/public-and-social-sector/our-insights/how-government-can-promote-open-data.

[4] M. V. Alderete, "Towards measuring the economic impact of open data by innovating and doing business," *International Journal of Innovation and Technology Management*, vol. 17, no. 03, p. 2 050 022, May 2020. DOI: 10.1142/s0219877020500224.

[5] Federal Ministry of the Interior and Community, "Open-Data-Strategie der Bundesregierung," Federal Ministry of the Interior and Community, Federal Strategy Paper, Jul. 2021. [Online]. Available: https://www.bundesregierung.de/resource/blob/997532/1940386/1d269a2ad1b6346fcf60663bdea9c9f8/2021-07-07-open-data-strategie-data.pdf?download=1 (visited on 02/02/2022).

[6] Open Government Partnership, *Open government declaration*, https://www.opengovpartnership.org/process/joining-ogp/open-government-declaration/, Accessed: 2022-10-09, 2011.

[7] Open Government Partnership, *Ogp national members*, https://www.opengovpartnership.org/our-members/, Accessed: 2022-10-09, 2022.

[8] Open Government Partnership, *Ogp local members*, https://www.opengovpartnership.org/ogp-local/, Accessed: 2022-10-09, 2022.

[9] European Parliament and Council, *Directive 2003/98/ec of the european parliament and of the council of 17 november 2003 on the re-use of public sector information*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32003L0098&qid=1665495613387, Accessed: 2022-10-09, 2003.

[10] European Parliament and Council, *Directive (eu) 2019/1024 of the european parliament and of the council of 20 june 2019 on open data and the re-use of public sector information*, https://eur-lex.europa.eu/eli/dir/2019/1024, Accessed: 2022-10-09, 2019.

[11] Group of Eight, *G8 open data charter and technical annex*, https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex, Accessed: 2022-10-09, 2013.

[12] Executive Office of the President, *Open government directive*, https://eur-lex.europa.eu/eli/dir/2019/1024, Accessed: 2022-10-09, 2009.

[13] K.-T. Tai, "Open government research over a decade: A systematic review," *Government Information Quarterly*, vol. 38, no. 2, p. 101 566, Apr. 2021. DOI: 10.1016/j.giq.2021.101566.

[14] J. Attard, F. Orlandi, S. Scerri, and S. Auer, "A systematic review of open government data initiatives," *Government Information Quarterly*, vol. 32, no. 4, pp. 399–418, Oct. 2015. DOI: 10.1016/j.giq.2015.07.006.

[15] F. Gonzalez-Zapata and R. Heeks, "The multiple meanings of open government data: Understanding different stakeholders and their perspectives," *Government Information Quarterly*, vol. 32, no. 4, pp. 441–452, Oct. 2015. DOI: 10.1016/j.giq.2015.09.001.

[16] Y. Gao, M. Janssen, and C. Zhang, "Understanding the evolution of open government data research: Towards open data sustainability and smartness," *International Review of Administrative Sciences*, p. 002 085 232 110 099, Apr. 2021. DOI: 10.1177/00208523211009955.

[17] R. Huang, C. Wang, X. Zhang, D. Wu, and Q. Xie, "Design, develop and evaluate an open government data platform: A user-centred approach," *The Electronic Library*, vol. 37, no. 3, pp. 550–562, Jun. 2019. DOI: 10.1108/el-02-2019-0037.

[18] F. Xiao, "Toward a conceptual model for users' online open government data interaction," in *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, ACM, Mar. 2021. DOI: 10.1145/3406522.3446007.

[19] G. M. Begany and J. R. Gil-Garcia, "Understanding the actual use of open data: Levels of engagement and how they are related," *Telematics and Informatics*, vol. 63, p. 101 673, Oct. 2021. DOI: 10.1016/j.tele.2021.101673.

[20] M. T. Islam, M. S. Talukder, A. Khayer, and A. N. Islam, "Exploring continuance usage intention toward open government data technologies: An integrated approach," *VINE Journal of Information and Knowledge Management Systems*, Jul. 2021. DOI: 10.1108/vjikms-10-2020-0195.

[21] Y. Zhao and B. Fan, "Effect of an agency's resources on the implementation of open government data," *Information & Management*, vol. 58, no. 4, p. 103 465, Jun. 2021. DOI: 10.1016/j.im.2021.103465.

[22] Y. Charalabidis, C. Alexopoulos, and E. Loukis, "A taxonomy of open government data research areas and topics," *Journal of Organizational Computing and Electronic Commerce*, vol. 26, no. 1-2, pp. 41–63, Jan. 2016. DOI: 10.1080/10919392.2015.1124720.

[23] S. Saxena, "Summarizing the decadal literature in open government data (OGD) research: A systematic review," *foresight*, vol. 20, no. 6, pp. 648–664, Nov. 2018. DOI: 10.1108/fs-07-2018-0074.

[24] I. Safarov, A. Meijer, and S. Grimmelikhuijsen, "Utilization of open government data: A systematic literature review of types, conditions, effects and users," *Information Polity*, vol. 22, no. 1, pp. 1–24, May 2017. DOI: 10.3233/ip-160012.

[25] A. Purwanto, A. Zuiderwijk, and M. Janssen, "Citizen engagement with open government data," *International Journal of Electronic Government Research*, vol. 16, no. 3, pp. 1–25, Jul. 2020. DOI: 10.4018/ijegr.2020070101.

[26] M. Janssen, Y. Charalabidis, and A. Zuiderwijk, "Benefits, adoption barriers and myths of open data and open government," *Information Systems Management*, vol. 29, no. 4, pp. 258–268, Sep. 2012. DOI: 10.1080/10580530.2012.716740.

[27] J. Cranefield, O. Robertson, and G. Oliver, "Value in the mash: Exploring the benefits, barriers and enablers of open data apps," in *Proceedings of the European Conference on Information Systems (ECIS)*, Association fo Information Systems, Jun. 2014, ISBN: 978-0-9915567-0-0.

[28] A. Zuiderwijk, R. Shinde, and M. Janssen, "Investigating the attainment of open government data objectives: Is there a mismatch between objectives and results?" *International Review of Administrative Sciences*, vol. 85, no. 4, pp. 645–672, Feb. 2019. DOI: 10.1177/0020852317739115.

[29] A. Francey and T. Mettler, "The effects of open government data: Some stylised facts," *Information Polity*, vol. 26, no. 3, pp. 273–288, Aug. 2021. DOI: 10.3233/ip-200281.

[30] A. Zuiderwijk and M. de Reuver, "Why open government data initiatives fail to achieve their objectives: Categorizing and prioritizing barriers through a global survey," *Transforming Government: People, Process and Policy*, vol. 15, no. 4, pp. 377–395, May 2021. DOI: 10.1108/tg-09-2020-0271.

[31] Hamburger Gesetze und Verordnungsblatt, *Hamburgisches Transparenzgesetz (HmbTG)*, German, https://transparenz.hamburg.de/contentblob/13584338/bea44f37a7a2c739a6208c6bc5a0e149/data/download-gesetzestext-neu-hmbtg.pdf, Legislation, 2012.

[32] I. Susha, Å. Grönlund, and M. Janssen, "Organizational measures to stimulate user engagement with open data," *Transforming Government: People, Process and Policy*, vol. 9, no. 2, pp. 181–206, May 2015. DOI: 10.1108/tg-05-2014-0016.

[33] A. Zuiderwijk, M. Janssen, and Y. K. Dwivedi, "Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology," *Government Information Quarterly*, vol. 32, no. 4, pp. 429–440, Oct. 2015. DOI: 10.1016/j.giq.2015.09.005.

[34] C. Martin, "Barriers to the open government data agenda: Taking a multi-level perspective," *Policy & Internet*, vol. 6, no. 3, pp. 217–240, Sep. 2014. DOI: 10.1002/1944-2866.poi367.

[35] A. Quarati, "Open government data: Usage trends and metadata quality," *Journal of Information Science*, p. 016 555 152 110 277, Oct. 2021. DOI: 10.1177/01655515211027775.

[36] G. M. Begany and E. G. Martin, "Moving towards open government data 2.0 in u.s. health agencies: Engaging data users and promoting use," *Information Polity*, vol. 25, no. 3, pp. 301–322, Sep. 2020. DOI: 10.3233/ip-190169.

[37] B. J. Oates, *Researching Information Systems and Computing*. London, UK: SAGE, 2012.

[38] R. K. Yin, *Case study research and applications*, 6th ed. Thousand Oaks, CA: SAGE Publications, Jan. 2018.

[39] OpenAI, *Whisper*, version n. a. Licensed under: MIT License, OpenAI, 2022. [Online]. Available: https://openai.com/blog/whisper/.

[40] LibreTranslate, *Libretranslate*, version n. a. Licensed under: AGPL-3.0 License, LibreTranslate, 2022. [Online]. Available: https://libretranslate.com/.

[41] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, Jan. 2006. DOI: 10.1191/1478088706qp063oa.

[42] V. Braun, *Thematic analysis : a practical guide*, eng. Los Angeles [etc.]: SAGE, 2021, ISBN: 9781473953246.

[43] Twenty to Nine LLC., *Delve*, version 1422, Licensed under: proprietary license, 2022. [Online]. Available: https://delvetool.com/.

[44] H. Yu and D. G. Robinson, "The new ambiguity of 'open government'," *SSRN Electronic Journal*, 2012. DOI: 10.2139/ssrn.2012489.

[45] E. A. Abu-Shanab, "Reengineering the open government concept: An empirical support for a proposed model," *Government Information Quarterly*, vol. 32, no. 4, pp. 453–463, Oct. 2015. DOI: 10.1016/j.giq.2015.07.002.

[46] B. W. Wirtz, J. C. Weyerer, and M. Rösch, "Open government and citizen participation: An empirical analysis of citizen expectancy towards open government data," *International Review of Administrative Sciences*, vol. 85, no. 3, pp. 566–586, Nov. 2017. DOI: 10.1177/0020852317719996.

[47] K. Y. DAHBI, H. Lamharhar, and D. Chiadmi, "Toward a user-centered approach to enhance data discoverability on open government data portals," in *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)*, IEEE, Oct. 2019. DOI: 10.1109/icds47004.2019.8942309.

[48] K. Hutter, J. Füller, and G. Koch, "Why citizens engage in open government platforms?" In *GI-Jahrestagung*, 2011, p. 223.

[49] A. J. Meijer, D. Curtin, and M. Hillebrandt, "Open government: Connecting vision and voice," *International Review of Administrative Sciences*, vol. 78, no. 1, pp. 10–29, Mar. 2012. DOI: 10.1177/0020852311429533.

[50] O. K. Foundation, *The Open Definition*, https://opendefinition.org/, 2022. (visited on 03/02/2022).

[51] R. Máchová, M. Hub, and M. Lnenicka, "Usability evaluation of open data portals," *Aslib Journal of Information Management*, vol. 70, no. 3, pp. 252–268, May 2018. DOI: 10.1108/ajim-02-2018-0026.

[52] D. Wang, D. Richards, A. A. Bilgin, and C. Chen, "Advancing open government data portals: A comparative usability evaluation study," *Library Hi Tech*, May 2021. DOI: 10.1108/lht-10-2020-0250.

[53] Y. Zhao and B. Fan, "Exploring open government data capacity of government agency: Based on the resource-based theory," *Government Information Quarterly*, vol. 35, no. 1, pp. 1–12, Jan. 2018. DOI: 10.1016/j.giq.2018.01.002.

[54] A. Nikiforova and M. Lnenicka, "A multi-perspective knowledge-driven approach for analysis of the demand side of the open government data portal," *Government Information Quarterly*, vol. 38, no. 4, p. 101 622, Oct. 2021. DOI: 10.1016/j.giq.2021.101622.

[55] R. Máchová and M. Lnenicka, "Evaluating the quality of open data portals on the national level," *Journal of theoretical and applied electronic commerce research*, vol. 12, no. 1, pp. 21–41, 2017. DOI: 10.4067/s0718-18762017000100003.

[56] G. Smith and J. Sandberg, "Barriers to innovating with open government data: Exploring experiences across service phases and user types," *Information Polity*, vol. 23, no. 3, pp. 249–265, Aug. 2018. DOI: 10.3233/ip-170045.

[57] I. Susha, Å. Grönlund, and M. Janssen, "Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries," *Information Polity*, vol. 20, no. 1, pp. 19–34, Jul. 2015. DOI: 10.3233/ip-150353.

[58] K.-S. Yap, "What is good governance?" United Nations Economic ans Social Commission for Asia and the Pacific, report, Jul. 2009. [Online]. Available: https://www.unescap.org/sites/default/d8files/knowledge-products/good-governance.pdf (visited on 03/05/2022).

[59] E. Barry and F. Bannister, "Barriers to open data release: A view from the top," *Information Polity*, vol. 19, no. 1, 2, pp. 129–152, Jun. 2014. DOI: 10.3233/ip-140327.

[60] B. Worthy, "The impact of open data in the uk: Complex, unpredictable, and political," *Public Administration*, vol. 93, no. 3, pp. 788–805, Apr. 2015. DOI: 10.1111/padm.12166.

[61] S. G. Grimmelikhuijsen, S. J. Piotrowski, and G. G. V. Ryzin, "Latent transparency and trust in government: Unexpected findings from two survey experiments," *Government Information Quarterly*, vol. 37, no. 4, p. 101 497, Oct. 2020. DOI: 10.1016/j.giq.2020.101497.

[62] S. S. Dawes, L. Vidiasova, and O. Parkhimovich, "Planning and designing open government data programs: An ecosystem approach," *Government Information Quarterly*, vol. 33, no. 1, pp. 15–27, Jan. 2016. DOI: 10.1016/j.giq.2016.01.003.

[63] E. Lakomaa and J. Kallberg, "Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs," *IEEE Access*, vol. 1, pp. 558–563, 2013. DOI: 10.1109/access.2013.2279164.

[64] M. T. Borzacchiello and M. Craglia, "The impact on innovation of open access to spatial environmental information: A research strategy," *International Journal of Technology Management*, vol. 60, no. 1-2, pp. 114–129, 2012. DOI: 10.1504/IJTM.2012.049109.

[65] C. Boudreau, "Reuse of open data in quebec: From economic development to government transparency," *International Review of Administrative Sciences*, vol. 87, no. 4, pp. 855–869, Jan. 2020. DOI: 10.1177/0020852319884628.

[66] A. Zuiderwijk, M. Janssen, and C. Davis, "Innovation with open data: Essential elements of open data ecosystems," *Information Polity*, vol. 19, no. 1, 2, pp. 17–33, Jun. 2014. DOI: 10.3233/ip-140329.

[67] L. Reggi and S. S. Dawes, "Creating open government data ecosystems: Network relations among governments, user communities, NGOs and the media," *Government Information Quarterly*, vol. 39, no. 2, p. 101 675, Apr. 2022. DOI: 10.1016/j.giq.2022.101675.

[68] T. Davies, "Open data, democracy and public sector reform," Thesis, University of Oxford, 2010, pp. 1–47. [Online]. Available: http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf.

[69] M. Gascó-Hernández, E. G. Martin, L. Reggi, S. Pyo, and L. F. Luna-Reyes, "Promoting the use of open government data: Cases of training and engagement," *Government Information Quarterly*, vol. 35, no. 2, pp. 233–242, Apr. 2018. DOI: 10.1016/j.giq.2018.01.003.

[70] G. Juell-Skielse, A. Hjalmarsson, P. Johannesson, and D. Rudmark, "Is the public motivated to engage in open data innovation?" In *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2014, pp. 277–288. DOI: 10.1007/978-3-662-44426-9_23.

[71] M. Kassen, "Understanding motivations of citizens to reuse open data: Open government data as a philanthropic movement," *Innovation*, vol. 23, no. 1, pp. 44–70, Jun. 2020. DOI: 10.1080/14479338.2020.1738940.

[72] B. W. Wirtz, J. C. Weyerer, and M. Rösch, "Citizen and open government: An empirical analysis of antecedents of open government data," *International Journal of Public Administration*, vol. 41, no. 4, pp. 308–320, 2018. DOI: 10.1080/01900692.2016.1263659.

[73] D. Wang, D. Richards, A. A. Bilgin, and C. Chen, "A predictive model for citizens' utilization of open government data portals," in *Digital Libraries at Times of Massive Societal Transition*, Springer International Publishing, 2020, pp. 159–175. DOI: 10.1007/978-3-030-64452-9_14.

[74] M. S. Talukder, L. Shen, M. F. H. Talukder, and Y. Bao, "Determinants of user acceptance and use of open government data (OGD): An empirical investigation in bangladesh," *Technology in Society*, vol. 56, pp. 147–156, Feb. 2019. DOI: 10.1016/j.techsoc.2018.09.013.

[75] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Management Science*, vol. 46, no. 2, pp. 186–204, Feb. 2000. DOI: 10.1287/mnsc.46.2.186.11926.

[76] Venkatesh, Morris, Davis, and Davis, "User acceptance of information technology: Toward a unified view," *MIS Quarterly*, vol. 27, no. 3, p. 425, 2003. DOI: 10.2307/30036540.

[77] M. Khayyat and F. Bannister, "Towards a model for facilitating and enabling co-creation using open government data," *Information Polity*, vol. 22, no. 4, pp. 211–231, Nov. 2017. DOI: 10.3233/ip-170406.

[78] V. Weerakkody, K. Kapoor, M. E. Balta, Z. Irani, and Y. K. Dwivedi, "Factors influencing user acceptance of public sector big open data," *Production Planning & Control*, vol. 28, no. 11-12, pp. 891–905, Jul. 2017. DOI: 10.1080/09537287.2017.1336802.

[79] A. Purwanto, A. Zuiderwijk, and M. Janssen, "Citizens' motivations for engaging in open data hackathons," in *Electronic Participation*, Springer International Publishing, 2019, pp. 130–141. DOI: 10.1007/978-3-030-27397-2_11.

[80] J. R. Crusoe and K. Ahlin, "Users' activities for using open government data – a process framework," *Transforming Government: People, Process and Policy*, vol. 13, no. 3/4, pp. 213–236, Aug. 2019. DOI: 10.1108/tg-04-2019-0028.

[81] T. Jetzek, M. Avital, and N. Bjorn-Andersen, "Data-driven innovation through open government data," *Journal of theoretical and applied electronic commerce research*, vol. 9, no. 2, pp. 15–16, Aug. 2014. DOI: 10.4067/s0718-18762014000200008.

[82] J. Atenas, L. Havemann, and E. Priego, "Open data as open educational resources: Towards transversal skills and global citizenship," *Open Praxis*, vol. 7, no. 4, p. 377, Oct. 2015. DOI: 10.5944/openpraxis.7.4.233.

[83] E. G. Martin, N. Helbig, and G. S. Birkhead, "Opening health data," *Journal of Public Health Management and Practice*, vol. 21, no. 5, E1–E7, Sep. 2015. DOI: 10.1097/phh.0000000000000127.

[84] O. D. Institute, *Data skills framework*, https://theodi.org/article/data-skills-framework/#1561999285194-8defa99a-39a52584-8cfe, May 2020.

[85] A. W. Li, L. S. Sinnamon, and R. Kopak, "Exploring learning opportunities for students in open data portal use across data literacy levels," *Information and Learning Sciences*, vol. 123, no. 9/10, pp. 601–620, Sep. 2022. DOI: 10.1108/ils-01-2022-0003.

[86] J. C. Prado and M. Á. Marzal, "Incorporating data literacy into information literacy programs: Core competencies and contents," *Libri*, vol. 63, no. 2, Jan. 2013. DOI: 10.1515/libri-2013-0010.

[87] A. F. Tygel and R. Kirsch, "Contributions of paulo freire for a critical data literacy: A popular education approach," *The Journal of Community Informatics*, vol. 12, no. 3, Oct. 2016. DOI: 10.15353/joci.v12i3.3279.

[88] M. J. Antikainen and H. K. Vaataja, "Rewarding in open innovation communities – how to motivate members," *International Journal of Entrepreneurship and Innovation Management*, vol. 11, no. 4, p. 440, 2010. DOI: 10.1504/ijeim.2010.032267.

[89] J. Moodysson, L. Coenen, and B. Asheim, "Explaining spatial patterns of innovation: Analytical and synthetic modes of knowledge creation in the medicon valley life-science cluster," *Environment and Planning A: Economy and Space*, vol. 40, no. 5, pp. 1040–1056, May 2008. DOI: 10.1068/a39110.

[90] *Transparenzranking 2021*, German, https://transparenzranking.de/, 2022.

[91] Bürgerschaft der Freien und Hanstestadt Hamburg, *Drucksache 20/2741, Unterrichtung der Präsidentin der Bürgerschaft*, German, https://www.buergerschaft-hh.de/parldok/dokument/35229/, Notice, 2011.

[92] Transparency International e. V., *Hamburg's Transparency Law to open government more than ever*, German, https://blog.transparency.org/2012/06/25/hamburgs-transparency-law-to-open-government-more-than-ever/index.html, 2012.

[93] Bürgerschaft der Freien und Hanstestadt Hamburg, *Drucksache 20/4466, Erlass eines Hamburgischen Transparenzgesetzes (HmbTG)*, German, https://www.buergerschaft-hh.de/parldok/dokument/37021/erlass_eines_hamburgischen_transparenzgesetzes_hmbtg.pdf, Motion, 2012.

[94] Bürgerschaft der Freien und Hanstestadt Hamburg, *Plenarprotokoll 20/34, Erlass eines Hamburgischen Transparenzgesetzes (HmbTG)*, German, https://www.buergerschaft-hh.de/parldok/dokument/38476/plenarprotokoll_20_34.pdf#page=22, Plenary Minutes, 2012.

[95] *Was braucht ein gutes transparenzgesetz?* German, https://transparenzranking.de/methodik/, 2022.

[96] M. Herr, C. Müller, B. Engewald, A. Piesker, J. Ziekow, and T. Ritter, "Abschlussbericht zur Evaluation des Hamburgischen Transparenzgesetzes," German, Institut für Gesetzesfolgenabschätzung und Evaluation, Report, Jul. 28, 2017. [Online]. Available: https://daten.transparenz.hamburg.de/Dataport.HmbTG.ZS.Webservice.GetRessource100/GetRessource100.svc/38402c6a-df76-4949-b1a4-d903641d2d4d/Akte_T10.pdf.

[97] M. Fishbein and I. Ajzen, *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. 1975, p. 578, ISBN: 9780201020892.

Appendices

Appendix 1 - Overview of TAM and UTAUT constructs

Table 4. *Overview of Core TAM Constructs*

Construct	Definition	Source
PU	"[...] the extent to which a person believes that using the system will enhance his or her job performance"	[75]
PEU	"[...] the extent to which a person believes that using the system will be free of effort"	[75]
SU[1]	a "person's perception that most people who are important to him think he should or should not perform the behavior in question"	[97] in [75]

Table 5. *Overview of Core UTAUT Constructs*

Construct	Definition	Source
PE	"[...] the degree to which and individual believes that using the system will help him or her attain gains in job performance"	[76]
EE	"[...] the degree of ease associated with the use of the system"	[76]
SI	"[...]the degree to which an individual perceives that important others believe he or she should use the new system"	[76]
FC	"[...] the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system"	[76]

[1]Only included in TAM2. See [75].

Appendix 2 - List of Interviewees

Table 6. *List of Interviewees*

No.	Affiliation	Perspective	Occupation	Role in Engagement	Duration
1	Ministry for Culture and Media	Provider	Division Head	Manager, Developer, Decision-maker	>70 min
2	Transparency International Deutschland e.V.	User	Volunteer	Analyst, Promoter, Activist	> 76 min
3	Parlamentwatch e.V.	User	Office Manager	Leader, Manager, Activist	> 35 min
4	Bündnis 90/Die Grünen	User	Member of Parliament	Decision-maker, Contributor	> 42 min
5	Code for Hamburg e. V.	User	Advisor, Volunteer, Entrepreneur	Contributor, Developer, Designer	> 65 min
6	PD – Berater der öffentlichen Hand GmbH	User	City Planer	Analyst, Contributor	> 52 min
7	Senate Chancellery	Provider	Department Head	Manager, Decision-maker, Developer	> 33 min
8	DER SPIEGEL GmbH & Co. KG	User	Data Journalist	Data Analyst, Developer	> 42 min
9	Hamburg University of Applied Sciences	User	University Professor	Manager, Leader, Data Analyst, Developer, Contributor	> 49 min

Appendix 3 - Interview Outline

All interviews were conducted in German, and corrected as specified in section 2.2.1.
The interview durations were approximated to be between 30 and 60 mins. Actual lengths can be viewed from the list of interviewees in table 6 in appendix 2. All interviews were recorded with the consent of the interviewees.
Each topic had a few fixed questions (below) and all follow-up questions had been derived from the emerging conversation in order to explore the topics further.

Introduction

Before we dive in to the three topics, as explained, please introduce yourself briefly, what is your background and what are you currently involved in?

Motivation and Drive

Turning to the first topic – Motivation and drive to engage with Open Government Data in the first place.
Please tell me from your perspective: What was your initial impetus, your initial motivation to engage with Open Government Data?
What about the transparency portal in particular?

Daily Interaction

Turning to the second topic – Daily Business and Interaction with Open Government Data.
Please describe to me how your day-to-day interaction with Open Government Data looks like, what it involves and what do you use the transparency portal for?
Please tell me what skills and tools have been useful or necessary for you to work with Open Government Data in your case?

Pursued and Desired effects, and Potential of OGD

Turning to the last topic – Desired effects, aims and potential of your use of Open Government Data. For you personally: What are the desired effects that you are expecting to see due to your engagement in Open Government data?
What about the desired effects in the context of your organisation (if applicable)?
Where do you see the long-term potential of Open Government Data?

Non-exclusive Licence for Publication and Reproduction of Graduation Thesis

I, John-Philipp Vogt, born 1990-10-25

1. allow the Tallinn University of Technology without any charges my work

   Analysis of Open Government Data Usage in Hamburg

   supervised by Eric Blake Jackson
   (a) to be reproduced for the purpose of conservation and electronic publication, including the digital repository of the Tallinn University of Technology, until the end of the copyrighted time limit;
   (b) to be available to the public through the Tallinn University of Technology online environment, including the digital repository of the Tallinn University of Technology, until the end of the copyrighted time limit.
2. I am aware, that all rights, named in section 1, will remain to the author.
3. I confirm that by allowing the use of this non-exclusive licence, no intellectual rights of third parties will be violated as set in the personal data protection act and other legislation.

(Signed Digitally)

Date: 2022-12-16