

Value creation and barriers in the commercial use of open government data

Frederico Gustavo de Menezes Magalhães Ribeiro Pereira

Dissertation for the Doctor of Philosophy in Digital Media

Faculdade de Engenharia da Universidade do Porto



Supervisor: Dr. Maria Catarina de Almeida Roseira (Faculdade de Economia da
Universidade do Porto)

Co-supervisor: Dr. Sharon L Strover (University of Texas at Austin)

ABSTRACT

The basic premise of open government data is that the release of public sector data in an open and machine readable format may be of significant value to non-governmental actors— citizens, non-profit organizations, and the private sector —who can use it for a variety of political, social, and commercial purposes. Currently, open government data is seen, for instance, as a potential driver of efficiency, increased transparency, citizen participation, and innovation. However, despite the importance of the private sector in generating innovation from this data, the practice and study of the commercial use of open government data is still in its infancy. Accordingly, the goal of this dissertation is to understand how private-sector organizations create value from open government data, and to uncover the main barriers encountered by commercial users in the use of the data.

Drawing on two distinct qualitative studies, we analyze various firms that use open government data to develop products and services, and/or as a source of competitive advantage. Our research aims to advance the extant knowledge theory in the commercial use of open government data, and to offer a set of proposals for entrepreneurs or intrapreneurs on how value can be derived from public datasets in a commercial setting. Simultaneously, we provide government officials with relevant suggestions regarding the aspects that need closer attention in the implementation of open government data strategies, in particular those attempting to foster private sector innovation.

Keywords: open government data; private sector; use barriers

RESUMO

A premissa base dos “dados públicos abertos” é que a publicação dos mesmos pode trazer um valor significativo para a sociedade civil – cidadãos, organizações não governamentais, e sector privado - que pode usá-los para uma variedade de fins políticos, sociais e comerciais. Neste sentido, hoje em dia, é atribuído aos dados públicos abertos o potencial de gerar maior eficiência, o aumento da transparência governamental, e maior participação do cidadão, e um aumento na inovação. No entanto, apesar da importância do sector privado na criação de valor a partir destes dados, a prática e estudo do uso comercial dos dados públicos abertos está ainda numa fase embrionária. Desta forma, o objetivo desta investigação é compreender como as empresas podem criar valor a partir dos dados públicos abertos, e identificar quais as barreiras de utilização dos mesmos.

Com base em dois estudos qualitativos distintos, são analisadas várias empresas que usam os dados públicos abertos para desenvolvimento de produtos e serviços e / ou como uma fonte de vantagem competitiva. Com base nos resultados, sugerem-se um conjunto de propostas práticas direccionadas para empreendedores sobre como criar valor a partir dos dados. São também elaboradas sugestões para os órgãos governamentais sobre quais os aspectos mais relevantes na implementação de estratégias de dados governamentais abertos, em particular aquelas que visam promover a inovação no sector privado.

Palavras-chave: dados públicos abertos; sector privado; barreiras de utilização

ACKNOWLEDGEMENTS

The journey of a Ph.D. is an incredibly challenging, yet profoundly rewarding one. Completing this dissertation would not have been possible without the support and understanding of a great number of people.

I am thankful to Prof. Artur Pimenta Alves and Dr. Carlos Guedes for bringing me onboard of this exciting doctoral program, entrusting me with the responsibility to create something of value. I am also deeply grateful to my supervisors Dr. Catarina Roseira, and Dr. Sharon Strover, for guiding me through these years, mostly in times when I could not see the light at the end of the tunnel. A note to the assistants Marisa Silva in Porto, and Karen Gustafson in Austin, who saved me (and so many other students) countless times from bureaucratic and logistical hurdles.

Throughout my research, I was privileged to work with lots of talented people. Although I cannot name them all, I would like to specially thank Joel Gurin, Laura, Audrey, and Katherine at the Center for Open Data Enterprise, as well as Dharmishta and Catherine Bracy at Code for America. I would also like to express my appreciation to all twelve firm managers that consented to share with me their experience in using open government data.

I would like to thank my family (my mother Ana, my father Miguel, my siblings Ariana, Cláudio, and Tiago, and my brother-in-law Alexandre) and friends for helping me in this odyssey. Special gratitude to my “brother-in-arms”, talented artist, and fellow Ph.D. student João Beira, with whom I had the honor to share this journey closely.

Last, but not least, I would like to express my most sincere recognition to Fundação para a Ciência e Tecnologia for awarding me a full scholarship, without which this work would most likely not have been possible.

TABLE OF CONTENTS

ABSTRACT	ii
RESUMO	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	vi
LIST OF FIGURES	ix
LIST OF TABLES	ix
1 INTRODUCTION	1
1.1 Background context	4
1.2 Research goals and relevance	9
1.3 Research questions	12
1.4 Research design	15
1.5 Contributions	16
1.6 Structure of this dissertation.....	19
PART I (THEORETICAL BACKGROUND)	22
2 THE CONCEPT OF OPEN GOVERNMENT DATA	23
2.1 Open government data as a socio-technical artifact.....	25
2.2 Open government data as a political narrative.....	30
2.2.1 Origins: Different perspectives on access to government information	31
2.2.2 Modern era: The global phenomenon of open government data.....	33
2.3 Conclusions	37
3 OPEN GOVERNMENT DATA: SUPPLY-SIDE	39
3.1 Government as a data provider.....	40
3.2 The value of open government data.....	45
3.3 Conclusions	54
4 OPEN GOVERNMENT DATA: DEMAND-SIDE	56
4.1 Open government data ecosystem.....	57
4.2 External stakeholders of open government data	59

4.2.1	Citizens	60
4.2.2	Non-profits	64
4.2.3	Businesses	68
4.3	Conclusions	78
PART II (EMPIRICAL RESEARCH).....		81
5	CREATING VALUE WITH OPEN GOVERNMENT DATA.....	82
5.1	Background	84
5.2	Methodology	86
5.3	Findings	93
5.3.1	Data to Fact.....	95
5.3.2	Data to Data.....	97
5.3.3	Data to Information.....	99
5.3.4	Data to Interface.....	102
5.3.5	Data to Service	104
5.4	Conclusions	107
6	BARRIERS IN THE COMMERCIAL USE OF OPEN GOVERNMENT DATA	110
6.1	Background context	111
6.2	Research goals and framework	118
6.3	Methodology	120
6.3.1	Data collection.....	121
6.3.2	Data Analysis	123
6.4	Findings	124
6.4.1	Resources	124
6.4.2	Data accessibility	126
6.4.3	Data usability	130
6.4.4	Interaction.....	133
6.4.5	Legislation	136
6.5	Conclusions	138
7	CONCLUSIONS.....	141
7.1	Contribution.....	146
7.1.1	Contribution to knowledge theory	146
7.1.2	Suggestions for firm managers.....	147

7.1.3	Suggestions for public officials.....	149
7.2	Future work.....	152
7.3	Concluding notes.....	152
8	BIBLIOGRAPHY.....	154
	Appendix I (Presentation of firms).....	182
	Appendix II (Interview questions).....	190

LIST OF FIGURES

Figure 1. Open, government, data	6
Figure 2. Publications with the term open data in the title, abstract, or keywords	9
Figure 3. Overview of the main different elements being studied.....	14
Figure 4. data.gov (retrieved May 10, 2016)	44
Figure 5. data.gov.uk (retrieved May 10, 2016).....	44
Figure 6. Frequency of open government data types of use	93
Figure 7. Frequency of “Data to fact” use type by industry.....	96
Figure 8. Frequency analysis of “Data to data” use type by industry	98
Figure 9. Frequency analysis of “Data to information” use type by industry.....	101
Figure 10. Frequency analysis of “Data to interface” use type by industry.....	103
Figure 11. Frequency analysis of “Data to service” use type by industry.....	106
Figure 12. Conceptual Framework.....	119

LIST OF TABLES

Table 1. Comparison of the empirical studies.....	18
Table 2. Summary of the estimated value of open government data	51
Table 3. Example of open government data use by a civic startup	71
Table 4. Example of open government data use by an infomediary firm	72
Table 5. Example of open government data use by an open data service.....	73
Table 6. Examples of new businesses based on open government data	75
Table 7. Examples of businesses that use open government data for competitive advantage	76
Table 8. Excerpt of data sample.....	90
Table 9. Main-themes and sub-themes that emerged from the data.....	94
Table 10. Current literature pertaining to each type of barrier.....	118
Table 11. Distribution of cases according to data use	121
Table 12. List of interviewees.....	122

1 INTRODUCTION

The impressive array of new technologies that have been produced in the last few decades has had a profound effect in society. The current ubiquity of Information and Communication Technologies (ICT's) in most countries around the world has created what has been described as a "knowledge-based society" (David & Foray, 2003). By the same token, the concept of information assumes a central role in the struggle for competitiveness in today's knowledge-economy (Lundvall & Johnson, 1994). The flow of information, and "the Internet in particular, offer unprecedented practical means to access, process, share, combine, organize and reuse vast amounts of information" (Iemma, 2012, p. 3). The magnitude of this change in our lives is almost unimaginable, with the current annual Internet traffic passing the zettabyte (10^{21}) threshold (CISCO, 2016).

The numbers really are astounding. There are over 3 billion Internet users in the world, with mobile phone users accounting for roughly half of the world's population (Williams, 2015). Simultaneously, the number of data-sharing devices (i.e., Internet of Things), which includes machine sensors and consumer-oriented devices such as connected thermostats, light bulbs, refrigerators, and wearable health monitors, is growing exponentially is projected to reach a figure between 50 to 200 billion by 2020 (Sun, 2016). In addition to machine-to-machine communication, every interaction we personally have with any digital technology or smart-device is programmed to generate data. Consequently, colossal amounts of data are being generated at any given moment. According to Lohr (2012), we now live in the "Age of Big Data".

Although the widely-used term “Big Data” is for the most part ambiguous and ill-defined (Boyd & Crawford, 2012), it broadly conveys the notion that manipulating large and complex datasets offers the possibilities of identifying previously impossible levels of insights, granularity of analysis, and relationships between elements in the dataset (Bertot, Lee, Hwang, & Choi, 2012).

TRANSPARENCY

The surge in data creation can also be noticed in the realm of government. Today, most governments are engaged in providing information-based online services through which citizens and officials can make better data-informed decisions, as well as interact and collaborate more efficiently (Kassen, 2013). With the explosion of digital media and the rise of Web 2.0 technologies, which are centered on social interaction and collaboration, massive amounts of data are generated, collected, and stored in repositories such as databases and data warehouses. As a result, modern government data repositories grew significantly in number, use and diversity (Dawes, Pardo, & Cresswell, 2004; Janssen & van den Hoven, 2015).

Despite the technical and political challenges that this situation entails, it simultaneously presents itself as an unparalleled opportunity to improve democratic governance and accelerate innovation. Bertot et al. (2012) posit that the availability of large amounts of data in government has the potential to enable a transformative approach to e-government services, public sector openness and transparency, as well as to redefine the interaction between governments, citizens, and the business sector. Based on the notion that governments collect huge amounts of data and conduct business with millions of citizens, Brown et al. (2011) argue that the public sector is the

most “fertile terrain for change” (p. 5). Conceivable, the growing digital lives of both the citizens and the government offers the opportunity to create unprecedented levels of mutual government–citizen understanding in the years to come (Clarke & Margetts, 2014).

Additionally, along with the evolution towards an information-based world, the importance of government openness has also grown steadily in the last few decades (Clarke & Margetts, 2014). In the last few years in particular, many governments around the world have included the issue of openness in the political agenda, with promises of concrete actions towards making governments transparent, effective and accountable.

By releasing public sector data in an open format (*i.e.* available at no cost and in a machine-readable format), governments can trigger profound changes in the relationship between governmental agencies and their stakeholders. The implications of opening, and subsequently using public data, are manifold (Davies, 2010; T Jetzek, Avital, & Bjørn-Andersen, 2013) contributing namely to higher transparency, more public sector efficiency, and economic growth (Halonen, 2012; Kulk & Loenen, 2012; Longo, 2011). By making data freely available, governments also expect to increase collaboration and citizen participation, such as the development of services by third parties for the benefit of society (Martin, Foulonneau, Turki, & Ihadjadene, 2013).

Overall, it is anticipated that the potential impact of open government data is enormous, as it offers society an opportunity to drive significant social, political and economic change (Jezek, Avital, & Bjorn-Andersen, 2012). Halolen (2012) notes that open government data “has been hailed as one of the most important public policies of

our time, and the potential impacts of sharing such data cooperatively are enormous” (p. 6). In light of this, in the last few years, governments around the world have rapidly enacted open data policies and mechanisms so that both the civil society and the private sector can tap into the data that resides in governmental servers (Zuiderwijk, Helbig, Gil-Garcia, & Janssen, 2014). Consequently, hundreds of open government data initiatives, that include the release of millions of datasets, have already been implemented worldwide (Jetzek, Avital, & Bjorn-Andersen, 2014). This global movement towards more access to government data is broadly called *open government data*.

1.1 Background context

While the concept of open government is not new (Luna-Reyes & Chun, 2012; Peixoto, 2013; Yu & Robinson, 2012), it was not until President Obama’s first day in office that the term “open government” became widely used in public discourse. Broadly, the open government doctrine is based on the pillars of transparency, collaboration, and participation between the government and its stakeholders (Dennis Linders & Wilson, 2011). This implies that “open government now means government where citizens not only have access to information, documents, and proceedings, but can also become participants in a meaningful way” (Lathrop & Ruma, 2010, p. 19).

Many authors have written in detail about open government – invoking citizen participation, collaboration, and government transparency (Downey, 2012; Lathrop & Ruma, 2010; McDermott, 2010; Peled, 2011) – as well as the integration of digital

technologies in the public sector (Fung, Russon Gilman, & Shkabatur, 2013). For example, Harrison et al. note that “while the term open government draws heavily from the collaborative relationships that characterize contemporary technology innovation, it is also tied conceptually to demands for transparency in political governance” (Harrison, Pardo, & Cook, 2012, p. 902). On a similar note, Wirtz & Birkmeyer (2015) describe open government as “a multilateral, political, and social process, which includes in particular transparent, collaborative, and participatory action by government and administration” (p.2).

The phenomenon of open government data has seen a spectacular growth since the announcement of a new era of “transparency and open government” in 2009 (Orszag, 2009). In the era ubiquitous data, the release of data produced or collected by governmental agencies, and its subsequent use without limitations, has thus become central to the paradigm of open government. Based on this notion, the concept of open government data can then be described as the intersection between the concepts of open, government, and data (Figure 1).

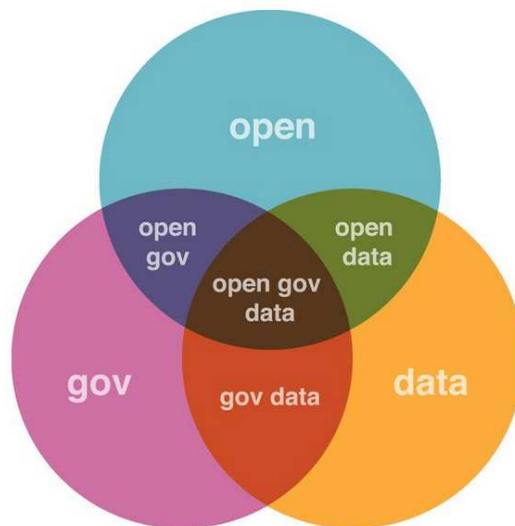


Figure 1. Open, government, data

Source; adapted from Gonzalez-Zapata and Heeks (2014)

Although the concept of information has become central in many contemporary philosophies (and in the practice of our daily lives), there is no “consensus on a single, unified definition of semantic information” (Floridi, 2005, p. 351). For example, Borglund and Engvall (2014) present information as “an interpretation of data – a human interpretation”, and suggest that data “can be described as the unit used to represent information” (p. 167). They add that information can be described as being “different to data in the sense that information has functionality and is useful, which data cannot be” (p. 167). The relevance of this distinction to the study of open government data is grounded in the notion that “as raw data is not meaningful to most end users, open data does not automatically create value” (Cranefield, Robertson, & Oliver, 2014, p. 3).

Consequently, one of the most critical challenges to the phenomenon of open government data is how to manage and make use of the vast amount of public sector data being produced (Halonen, 2012). This difficulty is further exacerbated by the fact that most agencies do not possess the innovativeness to fully exploit the overwhelming amount and diversity of data being generated (Chan, 2013).

In this context, the idea of encouraging external stakeholders to play an increased role in the innovation process has emerged as a promising approach (Klievink, Zuiderwijk, & Janssen, 2014; Tennison, 2015; Yang & Kankanhalli, 2013). Nevertheless, due to a myriad of barriers on the user's side in creating value from the data, the role of third-parties has become critical to the success of open government data initiatives (Robinson, Yu, Zeller, & Felten, 2009). On this topic, Robinson and Yu (2012) argue that "private actors, either nonprofit or commercial, are better suited to deliver government information to citizens and can constantly create and reshape the tools individuals use to find and leverage public data" (p. 161). This means that, by synthesizing and combining data from diverse sources into new products and services, private organizations are able to effectively translate raw data into useful information (Cranefield et al., 2014; DiFranzo et al., 2011). In addition to creating new products and services, private actors can also leverage on the potential of open government data as a source of competitive advantage (Porter & Millar, 1985; Zuiderwijk, Janssen, Poulis, & van de Kaa, 2015).

These circumstances have created a sizeable market opportunity for a wide range of innovative products and services using open government data, *i.e.* the 'open data economy'. It is estimated that open data (from both public and private sources) can

help unlock \$3 trillion to \$5 trillion in economic value annually (Manyika et al., 2013). In the European Union, open government data alone is estimated to generate an annual aggregate direct and indirect economic impact of EUR 40 billion (Ubaldi, 2013). Fundamentally, these economic projections further reinforce the notion of open government data as a promising and valuable socio-technical trend (Misuraca & Viscusi, 2014).

Following the pompous announcements of open data policies in the last few years, the interest in open government data from civil society has surged. Simultaneously, the interest of the academic community in studying the topic of open government data has also risen sharply (see Figure 2). However, despite this, open government data still remains a very recent field of both practice and study. In the first systematic analysis of the extant literature in the topic of open government data, Zuiderwijk et al. (2014) provide a glimpse of the main developments, challenges and barriers currently found in this field. Nonetheless, despite the emerging public attention on this issue, the study of open government data is still in infancy (Jetzek, Avital, & Bjorn-Andersen, 2013; Zuiderwijk et al., 2014).

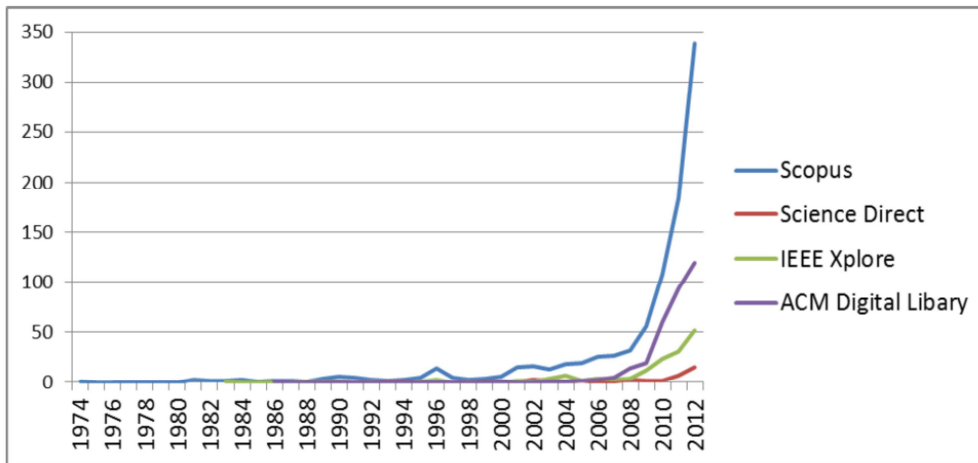


Figure 2. Publications with the term open data in the title, abstract, or keywords

Source: Zuiderwijk et al. (2014)

1.2 Research goals and relevance

The concept of government transparency, regarded as one of the pillars of democracy, is presently under immense transformation. The current historically low levels of trust in government (Walker, 2014), the recent incidents with whistleblowers such as Wikileaks and Edward Snowden, and the emergence of countless government watchdogs have all contributed to the increased level of scrutiny and accountability to which governments are currently held.

In this context, open government data has unsurprisingly taken a more central role in the agendas of various governments worldwide. However, in spite of the remarkable recent progress in this area, plenty of challenges can still be found when implementing open government data initiatives. For example, in opposition to the overall enthusiasm around the phenomenon of open government data, there are examples of government

officials who do not seem as keen on joining the current trend for more openness, imperiling the supply side of the equation. An extreme example has been exposed in Wisconsin, where politicians have attempted to dismantle open data legislation already in place in order to hide government actions from the public (Palmer, 2015). Simultaneously, the main challenge on the demand-side seems to be that the “response from external stakeholders to leverage government data for innovative activities has been lacking” (Yang & Kankanhalli, 2013, p. 644).

This situation muddies our understanding of the extent to which open government data is actually being used (Martin, 2014). Furthermore, it appears that the different stakeholders are not well understood, which can be particularly problematic in the context of research (Gonzalez-Zapata & Heeks, 2014). This dissertation aims to reduce this gap by focusing on the use of open government data by one important stakeholder group, the private sector. Specifically, the main goal of this dissertation is to understand how private-sector organizations create value from open government data, and to uncover the main barriers encountered by commercial users in the use of the data.

Our choice of studying the demand-side of open government data is grounded in the notion that, unlike the case of the supply-side, the user-side remains a rather understudied part of the open government data value-chain (Ferro & Osella, 2012; Kitchin, 2013; Martin, 2014). Jetzek et al. (2012) argue that “there is need for a new approach to explain the generation, capture and measurement of open government data value” (p. 8). More specifically, firms that use open government data are viewed today as a crucial element for stimulating innovation, as well as for delivering

information to citizens (Harrison et al., 2012; Robinson et al., 2009). Accordingly, Kaasenbrood et al. (2015) note that “it is important that private organisations are studied separately from other types of open data users, as they are expected to use open data in different ways” (p. 81). We thus focus specifically on the use case of the private sector due to the growing importance of these organizations in the creation of value through the use of government data.

The goal of this dissertation is well-aligned with the ideas of other authors, such as Dietrich (2015), who posits that “to be successful, open government data programs should not be approached as ‘technical exercises’ but require active engagement and collaborations between different stakeholders” (p. 3). A closer analysis of the use practices undertaken by firms may shed light on important aspects of open government data, including how open government data is used as a value-adding resource, as well as its limitations (Dumpawar, 2015).

With the results and insights produced by our research, we offer a set of proposals for entrepreneurs or intrapreneurs on how value can be derived from datasets in a commercial setting. In concurrence with the notion that a better understanding of the predictors of acceptance and use “can help policy-makers to determine which policy instruments they can use to increase the acceptance and use of open data technologies” (Zuiderwijk, Janssen, & Dwivedi, 2015, p. 429), we also provide government officials with relevant suggestions on which aspects of open government data need closer attention, particularly for open government data programs intended to foster innovation.

1.3 Research questions

Manifestly, the implications of the ubiquity of data in modern world extend far beyond the level of governance. Current mega trends, such as big data, Internet of Things (IoT), or the tremendous progress in areas such as machine learning and data mining are a testament to the current weight of data in every segment of society.

Today, data is a fundamental element in a multitude of businesses. Under these conditions, businesses “themselves are experiencing a transformation encompassing higher incentives towards the valorization of intangible assets” (Iemma, 2012, p. 3). However, extant literature highlights a general gap concerning how data can actually be used to create value for companies (Hartmann, Zaki, Feldmann, & Neely, 2014).

In the specific case of open government data, existing analyses are mainly based on anecdotal evidence, thus suggesting the need for a more rigorous scientific study of the phenomenon. However, the scarce extant academic literature on this topic is mainly of an exploratory nature (Ferro & Osella, 2012, 2013; Foulonneau, Martin, & Turki, 2014). In this context, a systematic analysis of how companies create value using open government data is needed in order to derive important insights on how open government data can be used innovatively to create new businesses, products and services, or to increase competitive advantage. Consequently, the first research question of this dissertation is:

RQ1: *How do for-profit organizations use open government data to create value?*

Considered often to be the raw material of the twenty-first century (Halonen, 2012), data must be located, extracted and refined in order to yield value (Dander, 2013;

Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Therefore, when released in its raw form, i.e. not yet converted into useful information, data is not open to the public in any meaningful sense (Cornford, Wilson, Baines, & Richardson, 2013). In the words of Rath (2012, p. 1), “data does not become ‘open’ until the insights it might potentially contain get revealed, and utilised fruitfully.” In this context, due to the presence of use barriers, the commercial use of open government data still holds untapped potential (Stott, 2014). Based on this notion, researchers have endeavored to study the issue of barriers perceived by users when attempting to convert raw data to meaningful information.

Moreover, it is well-accepted that the various ways in which open government data is used may correspond to different interests in the same data (Klievink et al., 2014; Tong, Irshad, & Ward Revell, 2013; Yu & Robinson, 2012). By acknowledging the plethora of uses through which open government can be exploited, it follows that a dataset can be perfectly useful to one purpose but of no value to another. However, as demonstrated earlier, existing research considers barriers mainly in a generic fashion. This is to say that, despite the importance of this issue, use barriers have not yet been studied in the context of varying uses of open government data. The issue of which barriers are found, and what role they play in different scenarios of data use, is therefore largely unknown. In this context, scholars have suggested that more research should be carried out to study the issue of barriers from the perspective of the user of the data (Martin et al., 2013; Zuiderwijk, Janssen, Choenni, Meijer, & Alibaks, 2012). Accordingly, the second research question of this dissertation is:

RQ2: *What are the barriers encountered by for-profit organizations in the commercial use of open government data, and how does the perception of barriers vary with the scenario of use?*

We argue that together, the two research questions proposed in this dissertation are useful to provide relevant insights into the issue of the commercial use of open government, the main focus of our research.

A synthesized overview of the main different elements being studied in this dissertation is presented below (Figure 3).

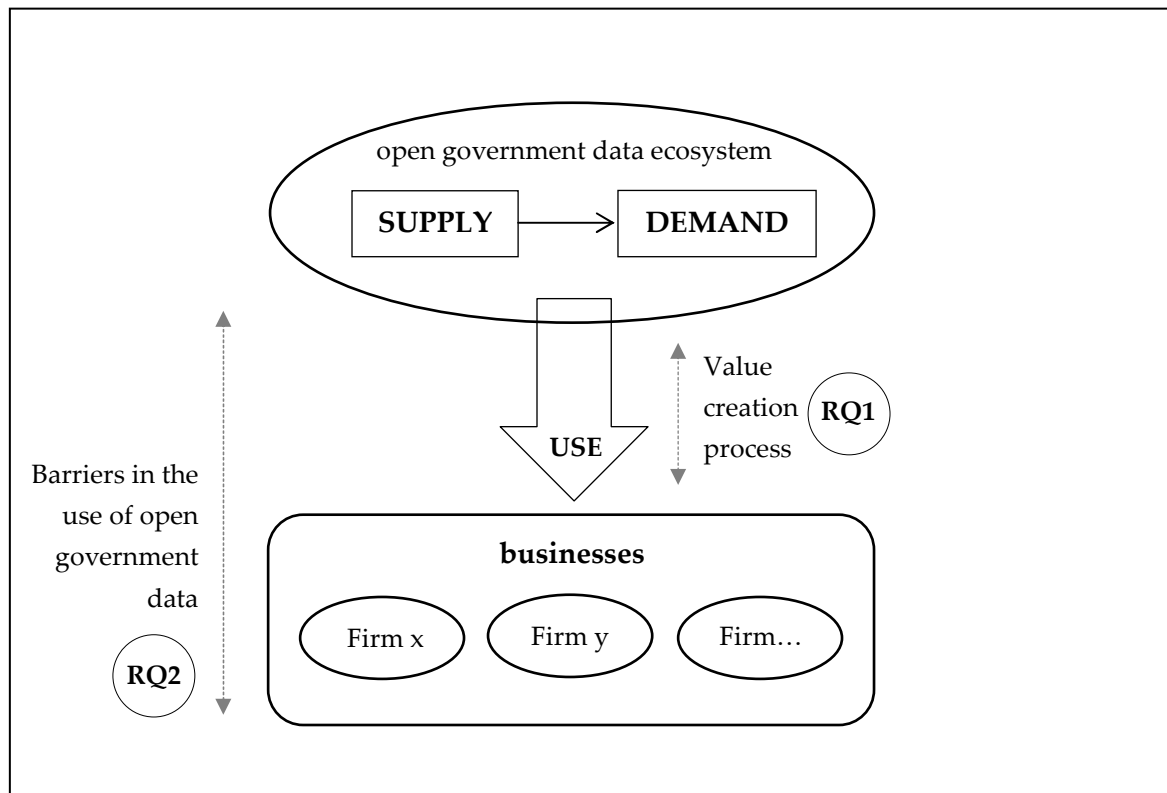


Figure 3. Overview of the main different elements being studied

The different ways in which businesses use open government data is the main focus of research question RQ 1. We answer this question by analyzing how government data is used commercially across different use cases from the private sector (Firm x, Firm y, Firm...). In addition, in RQ2, we investigate the barriers when attempting to derive value from open government data. Broadly, while RQ1 deals mainly with the level of use, RQ2 comprises the stages of discovery, use, and discussion, which encloses the

process of open government data use in its entirety (Zuiderwijk et al., 2012). By considering the process of use in its entirety, we aim to uncover the most prominent barriers that, in one way or another, may hinder the creation of value from open government data.

1.4 Research design

The understanding of the topic of open government data is currently fairly low (Halonen, 2012). In this context, Gray (2014) notes that “empirical research approaches from the social sciences could help to develop a sharper picture and a different perspective on who is using the concept, how they are framing it, and to what end” (p. 21).

In order to address the research questions described above, a qualitative methodology is deployed. Adopting the firm as the unit of analysis throughout this dissertation, our approach consists basically of two distinct, yet complementary, qualitative studies (*Table 1*). The content and sequence of the two studies follow the scope and working process of the research. In accordance, each study tackles a different research question (Chapter 5 refers to RQ1; Chapter 6 refers to RQ2). The specific methodological approach employed in each study is described individually in detail, in each chapter. Basically, chapter 5 employs a qualitative content analysis of 178 firms that use open government data, and chapter 6 draws from the qualitative analysis of interviews with 12 managers of firms using open government data to develop products and services, or for competitive advantage.

1.5 Contributions

This dissertation aims to contribute to a better understanding of the commercial use of open government data by shedding light on the “lack of connection between a reality of data provision and an aspiration of developmental results” (Gonzalez-Zapata & Heeks, 2014, p. 441).

Broadly, the low capacity of society to access and use the information can lead to a lack of interest and low interaction on the demand-side. Consequently, we anticipate that this dissertation is valuable to firm managers who are interested in aligning their firm’s business model towards a more effective use of open government data.

Moreover, our research intends to contribute to an improved awareness of the public sector about the benefits and risks of opening up their data (Janssen, 2011b). We do so by presenting empirical findings of the transformative impact of the releasing government data openly. Grounded in the notion that public institutions “might both enable and constrain the adoption of open data”, depending on how effective they are at supplying high-quality and relevant data to users, we offer insights into the dynamics of open government data use by the private sector. Consequently, this dissertation can be useful when attempting to align government strategies on open data with the needs of the private sector.

Finally, this dissertation makes an academic contribution in the following ways. First, it provides an updated review of the literature in the topic of open government data use, with a particular focus on the commercial use. It also advances the current theoretical

body on the field of open government data by leveraging on extant literature to propose a theoretical framework of the various types of barriers encountered by commercial users.

Study	CREATING VALUE WITH OPEN GOVERNMENT DATA (Chapter 5)	BARRIERS IN THE USE OF OPEN GOVERNMENT DATA (Chapter 6)	Contribution
Question	How do for-profit organizations create value using open government data?	What are the barriers encountered by for-profit organizations in the commercial use of open government data?	In chapter 5, we identify and classify the different types of processes through which commercial value is created from public data.
Purpose	Identify the different ways in which open government data is used by private-sector organizations.	Derive a theoretical framework from extant literature, identify the barriers in the commercial re-use of open government data from the perspective of the user; gain insights on how these barriers hinder the use of open government data by the private sector, and thus limit its potential for innovation.	In chapter 6, we identify and analyze the main barriers encountered by commercial users. We ensure diversity in our cases by including more than one case (firm) for each type of process derived from chapter 5.
Methods	Qualitative thematic analysis of the use of open government data by 178 firms that use open government data as part of their business model.	Case-study approach: interviews with 12 firms that use open government data in different ways.	
Publication	Published in the proceedings of the <i>2014 ICEGOV Conference</i> and awarded the prize of Best Research Paper of the conference.	Published in the proceedings of the <i>2016 ICEGOV Conference</i>	

Table 1. Comparison of the empirical studies

1.6 Structure of this dissertation

This dissertation is composed of two different parts. Part I (chapters 2, 3, and 4) introduces and examines the theoretical background that is relevant to the topic of open government data. In Part II (chapters 5, 6, and 7) we present the empirical work that was executed in order to shed light on the research problem under consideration.

The chapters of this dissertation are arranged in the following way:

We begin part I with a theoretical unpacking of the concept of open government data, which is the fundamental concept in this dissertation. To do so, we rely on the distinction between the socio-technical and the political dimensions of open government data. Accordingly, we first analyze open government data as a socio-technical artifact, and then, supported by a historical contextualization, as a political phenomenon (Chapter 2).

Chapters 3 and 4 are dedicated to explicating the supply and the demand sides of open government data, respectively. In chapter 3, we examine the role of government as the sole creator and supplier of open government data by reviewing the extant literature on this topic. We then focus on the mechanisms that are currently made available for the public to access government information, and on the potential (economic and social) value that may result from the use of open government data. Accordingly, the goal of chapter 4 is to describe the main stakeholders, commercial and non-commercial, that create value from open government data. We begin by introducing the concept of the open government data ecosystem, which is a useful metaphor when studying the demand-side of open government data, and then we individually

introduce the main external stakeholders in the ecosystem. We focus on the commercial use of open government data by private sector organizations, and provide useful real-world examples.

Part II of this dissertation comprises two chapters of empirical research (chapter 5 and 6), and a final chapter (chapter 7) that offers the main conclusions of the research. Each empirical chapter includes a brief, yet important, review of current literature pertaining to the most specific aspects of that particular study. In chapter 5, we investigate the issue of value creation in the context of the commercial use of open government data to propose a classification of the different types of use by commercial users. Differently, chapter 6 examines the issue of barriers in the commercial use of open government data. We determine, from the user's perspective, which barriers are perceived to be the most significant, and what impact those barriers may have on the business. We also highlight the role of barriers in different use scenarios of the data.

This dissertation ends with a concluding chapter (chapter 0) where the results produced by each study, as well as their limitations, are reviewed and consolidated. We begin chapter 7 by revising the main theoretical contributions of this dissertation, and move on to offer a set of proposals and practical insights for entrepreneurs or intrapreneurs on how value can be derived from public datasets. By the same token, we also provide government officials with relevant suggestions on which aspects need closer attention in the implementation of open government data strategies, particularly those attempting to increase or accelerate private sector innovation. We conclude with a brief review of the limitations of the study, and provide suggestions for future work.

PART I (THEORETICAL BACKGROUND)

2 THE CONCEPT OF OPEN GOVERNMENT DATA

The extraordinary progress of information and communication technologies in the last decades has resulted in an unprecedented level of information exchange between humans and machines. Individuals, as well as organizations, are currently impelled by the ubiquity of digital technologies to learn to navigate, and ultimately thrive, in an increasingly technology-centered society.

In spite of the unimaginable potential of a fully connected world, these developments simultaneously challenge some of the most fundamental aspects of our society. A particularly relevant example of this transformation is the process of innovation. What used to be an activity done behind closed doors by an exclusive group of experts is today frequently done openly, often in collaboration with an online community or a large group of people. In this context, the emergence of the practice of “open innovation” has significantly contributed to the idea of an information commons open to all, and has endowed organizations with the capacity to acquire knowledge (i.e., ideas, intellectual capital) from outside to accelerate internal innovation and expand markets (Huizingh, 2011; von Hippel, 2010).

Moreover, the concept of open innovation is now common to a wide range of disciplines, including economics, psychology, sociology, cultural anthropology (Huizingh, 2011), and even government (Kube, Hilgers, Fueller, & Koch, 2013), which illustrates its prominence in various areas of contemporary world. According to Van Veenstra and Van Den Broek (2013), the proliferation of new technologies that grant users the ability to easily collaboratively produce, share, distribute information – e.g. the semantic web, social media – was the main factor leading

to a culture of more openness. Interestingly, the evolution toward more openness manifests itself in the appearance of newborn concepts such as “open innovation, open source, open standards, open web platforms and, predominantly in the public sector, open government and open data” (Van Veenstra & Van Den Broek, 2013, p. 51).

In the particular case of open government data, a common understanding of the concepts and terms used is still missing (Jetzek et al., 2012; Lindman, Rossi, & Tuunainen, 2013). Based on the argument that the concept of open government data rapidly became “too vague” (p. 181), Yu and Robinson (2012) make an important distinction “between the technologies of open data and the politics of open government” (p. 181). They develop this idea as follows:

“The popular term ‘open government data’ is, therefore, deeply ambiguous— it might mean either of two very different things. If ‘open government’ is a phrase that modifies the noun ‘data’, we are talking about politically important disclosures, whether or not they are delivered by computer. On the other hand, if the words ‘open’ and ‘government’ are separate adjectives modifying ‘data’, we are talking about data that is both easily accessed and government related, but that might or might not be politically important.”

The point of departure of our theoretical analysis is driven by the suggestion of Yu and Robinson (2012) to broadly examine the concept of open government data in its socio-technical (Davies & Frank, 2013; Martin, 2014) and political (Denis & Goota, 2014; J. A. Johnson, 2014; Lassinantti, Bergvall-Kareborn, & Stahlbrost, 2014) dimensions. Accordingly, this chapter is composed of two main parts. While in the first part (section 2.1.1) we analyze open government data as a socio-technical artifact, in the second part (section 2.1.2) we delve more into its

political dimension, and provide a historical contextualization of the phenomenon of open government data.

2.1 Open government data as a socio-technical artifact

Like other new-born technological innovations, the recent study of open government data still lacks a coherent and consistent terminology (Charalabidis, Alexopoulos, & Loukis, 2016; Magalhaes, Roseira, & Strover, 2013). Consequently, the terms often employed by experts remain ambiguous (Peixoto, 2013; Sandoval-Almazan, 2015; Yu & Robinson, 2012). This situation makes it harder for various stakeholders – e.g. researchers, policymakers, activists – to articulate clear priorities and make cogent demands (Yu & Robinson, 2012). Additionally, it may also hinder the adoption of open government data by confusing some of the actors who might benefit from it (Tammisto & Lindman, 2012).

For example, the terms “open data” and “open government” data are often used interchangeably in literature (Heimstädt, Saunderson, & Heath, 2014a). However, in this dissertation we make a conceptual distinction between both terms, which is closely related to the central role of data in current ICT’s.

The development of Web 2.0’s interactive social layer in the 2010s introduced the opportunity for an unprecedented level of interaction, and as a result, it changed the dynamics of information sharing between users. Importantly, the social nature of Web 2.0 allows users to, for example, collaboratively produce, share, and distribute content in social network, to join efforts

in collective funding and sourcing initiatives (*crowdsourcing* and *crowdfunding*), and to mash content (e.g., text, image, audio or video) from different sources to generate new content or knowledge (mash-ups). Consequently, the widespread use of Web 2.0 applications, in particular social media and location-based services (Longo, 2011), has resulted in massive amounts of data being generated, collected, and stored. These remarkable new features of web-based technologies represent a shift of incalculable potential.

The impact of the transition from Web 1.0 to Web 2.0 technologies has propagated through all levels of society, blurring the line between open and closed, and ultimately contributing to the rise of “open” as an organizational strategy (Van Veenstra & Van Den Broek, 2013). This new paradigm, one based on openness, largely laid the groundwork for new ideas such as open innovation and open source, and more recently, open data (Clarke & Margetts, 2014).

Open data can be briefly defined as “data that can be freely used, reused and redistributed by anyone”¹, but many organizations have offered other definitions. The Open Data Institute² defines open data as “data that anyone can access, use and share” and adds that “for data to be considered ‘open,’ it must be published in an accessible format, with a license that permits anyone to access, use and share it”³. More comprehensively, Open Knowledge⁴ asserts three conditions⁵ by which data can be considered open:

¹ <http://opendefinition.org/> retrieved Dec 21, 2014

² The Open Data Institute is a non-profit organization, based in London, that aims to catalyze the evolution of open data culture to create economic, environmental, and social value: <http://theodi.org/>

³ <https://theodi.org/open-data-means-business> retrieved Dec 21, 2014

Availability and Access: The data must be available as a whole, and at no more than a reasonable reproduction cost, preferably by downloading over the Internet. The data must also be available in a convenient and modifiable form.

Reuse and Redistribution: The data must be provided under terms that permit reuse and redistribution including the intermixing with other datasets.

Universal Participation: Everyone must be able to use, reuse and redistribute - there should be no discrimination against fields of endeavor or against persons or groups. For example, “non-commercial” restrictions that would prevent “commercial” use or restrictions of use for certain purposes (i.e., only in education) are not allowed.

In yet another approach, the World Wide Foundation Consortium (W3C) ⁶ developed a five star model that has been used widely to assess data readiness. The model attempts to describe the different characteristics of open data according to the following 5 levels:

- Level 1 - Data is visible, licensed for reuse, but requires considerable effort to reuse.
- Level 2 - Data is visible, licensed, and easy to reuse, but not necessarily by all.
- Level 3 - Data is visible and easy to reuse by all (not restricted to using specific software).

⁴ Open Knowledge is a global non-profit network that promotes and shares information at no charge, including both content and data

⁵ <http://opendatahandbook.org/en/what-is-open-data> retrieved Dec 21, 2014

⁶ The World Wide Web Consortium (W3C) is the main international standards organization for the World Wide Web

- Level 4 - Data is visible, easy to use and described in a standard fashion.
- Level 5 - Data is visible, easy to use, described in a standard fashion and meaning is clarified by being linked to a common definition.

In a somewhat approach, Davies (2014c) suggests that open data must be 1) proactively published, 2) machine readable, and 3) legally reusable. Davies' rationale is useful since it provides a more "binary" definition of open data. However, despite efforts to better define (technically and semantically) open data, no definition has yet been able to offer clear criteria to assess whether a dataset is considered to be open or not (Yannoukakou & Araka, 2014).

Moreover, the issue of machine readability of open data is a central one, as it is the basis upon which modern data technologies can read, and process the data. For example, the concept of *Big and Open Linked Data* (BOLD) – a combination of *Linked data*, which relates to connecting structured and machine readable data that can be semantically queried, and *Big data*, which deals with large volumes of data from a variety of sources that need to be processed – relies on the machine-readability of data (Dwivedi et al., 2015; Janssen & van den Hoven, 2015).

Another important aspect of open data is that, by definition, it is not confined to a pre-determined provider. In other words, to be considered *open*, data can be supplied by any given entity – including individuals, and private and public organizations. In fact, the principles of open data have been applied in various domains, such as corporate responsibility⁷, emergency

⁷ The OpenCorporates is a website <https://opencorporates.com/> which shares data on corporate entities as open data under a share-alike license is noteworthy example of the application of open data principles

knowledge management (Yates & Paquette, 2011), crowdsourcing of map data (Hudson-Smith, Batty, Crooks, & Milton, 2008), open data for science (Molloy, 2011).

Despite the expansion of the principles of open data to various areas of application, the largest producer of data is the public sector (Dawes, 2012; Janssen, Charalabidis, & Zuiderwijk, 2012). This is mainly because the public sector produces, collects, and maintains vast amounts of data within the exercise of its public tasks. Then, following the growing pressure for more government openness from both civil society (Lathrop & Ruma, 2010) and the private sector (Bates, 2012) in the last few years, governments began publishing public data online in open format, i.e. open government data.

In this context, the discourse of open data is largely dominated by the notion of open government data (Lathrop & Ruma, 2010).

“While much data is supplied from outside the public sector, compared to many other areas of the economy, the public sector plays an unusually prominent role. In many key areas, a public sector organization may be the only, or one among very few, sources of the particular information it provides (e.g. for geospatial and meteorological information).” (R. Pollock, 2009, p. 2)

Broadly, open government data can be described as open data that is “produced or commissioned by government or government controlled entities”⁸. Complementarily, other aspects that ought to be considered when defining open government data have also been noted. For example, Janssen et al. (2012) define open government data as “non-privacy-restricted and

⁸ “What is Open Government Data?” by Open Knowledge Foundation
<http://opengovernmentdata.org/what/> retrieved May 8, 2016

non-confidential data which is produced with public money and is made available without any restrictions on its usage or distribution” (p. 258). Also, Borglund and Engvall (2014) posit that the main objective of open government data “is not just to make it accessible for everyone but also reusable” (p. 165).

2.2 Open government data as a political narrative

Besides its more technical aspects, the political and participatory facets of open government data are also frequently invoked by scholars and critics. In this perspective, the open government data phenomenon is described by Bauer and Kaltenböck (2011) as a “worldwide movement to open up government/public administration data, information and content to both human and machine-readable non-proprietary formats for re-use by civil society, economy, media and academia as well as by politicians and public administrators” (p. 10). Likewise, other authors describe the noticeable growth of the issue of open government data as a “movement” (Dawes, 2012; J. A. Johnson, 2014; Peixoto, 2013).

The open data movement has flourished in the same timeframe as other open movements such as open source (Lindman & Nyman, 2014), open access, open education, and open knowledge (Davies & Bawa, 2012). Much in the same way, it broadly originates from the foundational principles of a free, equalitarian, and cooperative model of access and redistribution of information (Webber, 2004). Open government data “is not apolitical” (Halonen, 2012, p. 11).

The basic premise of the open government data movement is that such data may be of considerable value to non-governmental actors—civil society organizations, media, private sector, and citizens—who could use it for a variety of political, social, and commercial purposes (Shkabatur, 2012). With the focus on non-personal data, the content of the data may refer to different public domains ranging from health, energy, education, traffic, weather, and public sector budgeting to other kinds of data about policies and inspection (Janssen et al., 2012).

For instance, the Global Positioning System (GPS) data was reserved for military use until 1983. In that year, President Reagan opened America’s Global Positioning System (GPS) data giving all countries access to the GPS data in response to the Soviet Union military action (Brantley, 2015). This historical episode arguably counts as a superlative example of a successful open government data initiative as it gave rise to a plethora of technologies, such as the smart phone, or turn-by-turn navigation, that are only possible because satellite data is made available in the form of open government data. On a similar order of magnitude, the release of meteorological data has led to the creation of a competitive and innovative private weather market at a global scale (Jetzek et al., 2013).

2.2.1 Origins: Different perspectives on access to government information

Historically, governments have commonly restricted access to information resources (Davies, 2010). This is not surprising, given the inherent nature of bureaucracies to be secretive (Weber, 1978). Still, in most modern democratic nations, the “access to government information is an institutional right, which constitutes a precondition for the proper exercise of civic rights

whereon the democratic governance springs” (Yannoukakou & Araka, 2014, p. 334). Hernon (1991) notes that the access to government information “is an inherent right or part of the social contract between government and the governed to hold government accountable to its citizenry”. In this context, civil society movements have continuously battled for more political transparency and access to government information (Bates, 2012).

Broadly, two advocacy groups – “Right to Information” and “Open Government Data” – are commonly credited as the main proponents for greater openness of public information in the last few decades (Access Info Europe & Open Knowledge Foundation, 2010; Davies, 2010; Wirtz & Birkmeyer, 2015). Although there is a significant overlap between them, there are also important differences in the approach employed by each one (Janssen, 2012).

It can be said that the “Right to Information” movement relates closely to democratic ideals. This is centrally supported by “the notion that in order for an individual to be able to freely express ideas, opinions and thoughts, it should be able to freely formulate them, hence to be very well informed” (Yannoukakou & Araka, 2014, p. 332).

Alternatively, as a reaction to public bureaucracy and the information-gathering efforts of governments, the members of the “Open Government Data” movement typically adopt a more technically-oriented approach to datasets that are held in government databases (Access Info Europe & Open Knowledge Foundation, 2010).

On the intersection between the two movements, Halonen (2012) notes that “open data clearly is an integral part of the wider freedom-of-information continuum and should be addressed as

such” (p. 41). Despite their differences, the fundamental principles of Open Government Data and of the Right to Information “are melting together” (Janssen, 2011a, p. 447). In fact, while working together in the demand for more access to government data, these movements have, so far, achieved remarkable results.

“Currently, around 90 countries have adopted freedom of information legislation and 50 more have legislation pending. In addition, freedom of information has been recognized as a constitutional right in more than 30 countries, and it is increasingly considered as a human right” (Janssen, 2012, p. 2).

2.2.2 Modern era: The global phenomenon of open government data

In a rather informative approach, Clarke and Margetts (2014) offer a brief technocratic review of the evolution of government in the Western world in the last 100 years. Roughly corresponding the first half of the 20th century, they point out that the prevailing model was characterized by paper-based systems, formal written files, and command and control in large hierarchical structures. Later, in the 1970s, public sector information began emerging as a policy issue mainly because information and communications technology allowed an easier separation of information from the processes that generated it (Bates, 2012).

Importantly, the period between the 1980s to the early 2000s saw a disaggregation of large-scale departments and increased competition within the public sector, with government IT units being “outsourced and in some cases privatized” (Clarke & Margetts, 2014, p. 395). This

coincided with the increased interest of public bodies in generating revenue from their valuable informational assets (e.g. geospatial mapping data, postal code data) (Bates, 2012).

Finally, during the 2000s, the prevailing model of governance placed digital technologies at the center of bureaucracy, which, in comparison with the earlier decades, contributed to narrow the gap in government-citizen data flow (Clarke & Margetts, 2014). In addition to this, governments themselves began to realize the strong potential of public sector information to generate commercial and social value (Davies, 2010). A testament to this is the enactment of the Directive 2003/98/EC by the European Commission in 2003, which focused on the reuse of Public Sector Information.

Soon after, on August 29, 2005, Hurricane Katrina hit the Louisiana–Mississippi border in United State. With 1,250 casualties and damages estimated at \$125 billion, it became the most destructive hurricane ever to hit the United States (Peled, 2011). Besides the terrible humanitarian losses, Hurricane Katrina also created a negative impact at the governance level: it exposed serious information-sharing flaws at the federal level of the American government (Peled, 2011). In contrast to the prior Administration’s approach to information, President Obama - in his first day in office in 2009 - signed the “Transparency and Open Government” initiative.

Before this initiative, the vast majority of open data was generated by research communities and institutions (Latif, Saeed, Hoefler, Stocker, & Wagner, 2009). However, in light of this initiative, the new administration was committing to “an unprecedented level of openness” and to establish “a system of transparency, public participation, and collaboration” (Obama, 2009).

This initiative marked the beginning of a new era (Luna-Reyes & Chun, 2012; Peixoto, 2013; Yu & Robinson, 2012), one in which the issue open government data was to become visible in the public discourse around the world

A few months later (but after years of mounting pressure from civil society for the release of government data in machine-readable format), the government of the United Kingdom formalized its commitment to publish data online, despite the resistance from a small number of government agencies (Hogge, 2010).⁹

In 2011, in accordance with the rising trend of opening up public data in individual European member states and in countries around the world, the European Commission moved forward to announce an open data strategy. In effect, this established open government data as an important component of the European Commission's vision of a competitive future European Union (Halonen, 2012), which includes the growth of information industries and better government transparency (Kulk & Loenen, 2012).

Following the pioneer commitment (and investment) by the governments of the United States, United Kingdom, and European Union, the issue of open government data picked up much publicity internationally (Zuiderwijk et al., 2014). Open government data became "fashionable" (Janssen, 2012). The launch of the Open Government Partnership (OGP)¹⁰ in 2011 arguably

⁹ in particular the Government's cartographic mapping agency, the Ordnance Survey, which was funded from revenues from the commercial exploitation of public data

¹⁰ An "international platform for domestic reformers committed to making their governments more open, accountable, and responsive to citizens" (retrieved from <http://www.opengovpartnership.org/who-we-are> in May 9, 2016)

constitutes the first milestone in the actualization of open government data as a global phenomenon. Since then, the OGP has grown from eight countries (Brazil, Indonesia, Mexico, Norway, Philippines, South Africa, the United Kingdom, and the United States) to 66 participating countries ¹¹. In the context of developing nations in particular, open government data has been gaining very important ground as a tool for increasing innovation and public accountability (Davies, 2014a; Schalkwyk, Cañares, Chattapadhyay, & Andrason, 2015; Zuiderwijk et al., 2014). In regions or business sectors where corruption is rather problematic, open government data has also been used to deter bribery (Fagan, 2016; Noveck & Gorof, 2013). This has been attempted, for example, by publishing government procurement data (i.e. public contracting) in countries such as Ukraine and Paraguay to make the process of public contracting more transparent ¹². In Mexico too, open government data was used to uncover a shenanigan in which 1,400 teachers had apparently been born on the same day in 1912 (The Economist, 2015).

Eventually, open government data has become the core idea behind a growing cadre of government leaders, civil society actors, and technologists (Davies & Bawa, 2012). However, with ever more governments around the world delineating and implementing open government data strategies, new challenges do arise. For example, considering that the multitude of countries currently launching open government data initiatives are vastly different in their level of development, political structure, and public policy priorities, there is the need to

¹¹ <http://www.opengovpartnership.org/> retrieved Feb 6, 2015

¹² <http://www.open-contracting.org/why-open-contracting/worldwide/> retrieved May 9, 2016

better understand, namely through comparative analyses, the implementation of their strategies in varying contexts (Davies, 2014b; Huijboom & Van den Broek, 2011; Zuiderwijk & Janssen, 2014b).

2.3 Conclusions

In this chapter, we presented some of the most important aspects pertaining to the concept of open government data. We rely on the distinction between the technologies of open data and the politics of open government to unpack the concept of open government data. We do so by first approaching open government data as a socio-technical artifact (section 2.1), and then as a political narrative (section 2.2). Considering the goal of this dissertation, our approach to open government data relates more closely to the perspective of technologies of open data than to the political dimension of open government data. Accordingly, we review some of technical definitions of open government data provided by scholars and practitioners, and make yet another distinction by delimiting open government data only as open data published by government sources. Moreover, we examine the political dimension of open government data to provide the reader with a synthetic account of the political and historical developments (in government as well as in civil society) that contributed to the emergence of this trend. We finalize this chapter with a brief review of some of the most influential political initiatives that granted open government data the condition of a global phenomenon.

In light of the growing informational interdependence between the various stakeholders, “the process in which data are created, published, found, used and discussed, which we refer to as

the open data process, is complex, dynamic and heterogeneous” (Zuiderwijk & Janssen, 2014a, p. 115). Thus, based on the notion that open data should not just be seen as a product, but as an ongoing process (Zuiderwijk et al., 2012), multi-staged models have been proposed to describe open government data as a process (Dander, 2013; Zuiderwijk et al., 2012). Fundamentally, these contributions concur that a major distinction can be done, clearly differentiating between the supply-side (government) and the demand-side (users of the data) (Gurstein, 2011). Accordingly, in the next two chapters (chapters 3 and 4), we address the current literature on the supply-side and demand-side of open government data respectively.

3 OPEN GOVERNMENT DATA: SUPPLY-SIDE

The current ubiquity of digital technology in public administration is heavily transforming the notion of what government is and how it should reach out to its citizens. Today, most governments are engaged in providing information online through which citizens and officials can make better data-informed decisions, as well as interact and collaborate more efficiently with each other (Kassen, 2013).

Understanding the role of government is paramount to any endeavor in the field of open government data. Accordingly, in this chapter, we attempt to provide the reader with some relevant background context regarding the role of government as the sole creator, and supplier, of open government data. To do so, we begin (in section 3.1) by historically reviewing extant literature on the supply-side of open government data, subsequently focusing on the mechanisms that are currently made available for the public to access to government information. Next (in section 3.2), we discuss the potential value generated by open government data, distinguishing between the economic and social value of the data. Based on this rationale, we proceed to review various studies and reports that contributed to the discussion of the value of open government data in the last few years. We conclude by revealing several instances in which the release of open government data has generated unintended negative consequences.

3.1 Government as a data provider

Given its unique role as the primary producer and distributor of data, the weight of government in the overall mechanics of the open government data ecosystem is colossal (Dawes, 2012; Janssen et al., 2012). Accordingly, most of the research in the field of open government data has, so far, looked at the supply side of data (Agrawal, Zhang, & Kettinger, 2014; Evans & Campos, 2008).

Broadly, two different factors may help explain this. First, since the creation of data is the primary stage in the open data process (Janssen & Zuiderwijk, 2014). In other words, the process of providing the data precedes the process of using it. Consequently, it logically follows that the initial focus of analysis is at the government level. The other factor that may contribute to the research focus on the supply-side relates to the fact that the public rhetoric regarding open government data has been, until recently, strongly intertwined with emerging ideas of Open Government - collaboration, participation, and transparency (Cranefield et al., 2014). Consequently, the public discourse on open government data, mainly in the United States and the United Kingdom, has been predominantly geared towards what the government could offer their constituencies rather than the other way around.

In this context, until the early 2010's, the focus was on how to get the data out, i.e. releasing the data that was trapped in government data servers. In accordance, at this point in time, the study of open government was particularly concerned with the overall transition from a "closed to open system" occurring in the public sector (Janssen et al., 2012). As a consequent, supported by a more utilitarian approach towards technology (Evans & Campos, 2008), the study of open

government data delved mainly into government-facing issues, such as government modernization (O'Reilly, 2010), organizational and cultural change in government agencies (Broad, Smith, Duhaney, & Carolan, 2015; Schrier, 2010), the metrics by which to evaluate the success of government open data initiatives (Bertot, McDermott, & Smith, 2012), the liberation of government data (Brito, 2010; Chun, Shulman, Sandoval, & Hovy, 2010), legal and regulatory frameworks (Access Info Europe & Open Knowledge Foundation, 2010; McDermott, 2010), negative attitudes among public officials (Halonen, 2012), and using information technology to create a culture of transparency (Bertot, Jaeger, & Grimes, 2010). In a nutshell, existing research exposed that, despite the laudable efforts of some governments, the road to an open government future was still paved with various hurdles (Schrier, 2010).

With the rapid adoption of open data infrastructures in recent years (Zuiderwijk et al., 2014), research on the supply-side has nevertheless evolved "beyond transparency" (Goldstein & Dyson, 2013). This is to say that, upon the assumption that governments have liberated reasonable amounts of data, some researchers' attention pivoted towards the value of the data being provided (Agrawal et al., 2014). Appropriately, the scope of more recent research has been extended to address aspects that are important to the publication of value from the data, such as data quality (O'Hara, 2012a; Ren & Glissmann, 2012), implementation of open data policies (Davies, 2014b; Misuraca, Mureddu, & Osimo, 2014; Zuiderwijk & Janssen, 2014b), and the "back room" operations of data release (Denis & Goota, 2014). In light of such contributions, it has become apparent that, due to its inherent complexity, an "effective" open government program requires "time and patience to grow" (Peled, 2013, p. 195).

There are mainly four different ways for the public to access to government information (Cuillier & Piotrowski, 2009, p. 442):

1) *proactive dissemination* by agencies through press releases, posting documents online, or providing in a library or depository;

2) *requester release* where citizens and journalists specifically request information from agencies not provided proactively;

3) *leaks* from whistleblowers and others (e.g., Pentagon Papers case); and

4) *open public meetings* where information is discussed and released in a public venue.

In light of the commitments made by various governments to more government transparency and accountability in the last few years, governments are expected to engage in the proactive dissemination of open government data. Moreover, with the advent of electronic government, one-stop government portals have been developed worldwide “as a precondition for more qualitative services by reducing the work overload along with the duplicates and the redundant procedures, and guarantying its unrestricted access by the citizens” (Yannoukakou & Araka, 2014, p. 334). Accordingly, in the framework of their open data policies, most governments are currently involved in increasing the amount and diversity of government data available in public repositories on the Web. These data repositories, known as ‘open data portals’, provide the main point of access to government data that is published in an open format.

The open data portal is a central element in the open government data process as it commonly acts as the main interface between the government and the user. With that in mind, open data

portals typically offer a combination of functionalities for data delivery in order to cater for the needs or expectations of different kinds of users. For the common citizen, who often “cannot perform the essential operations needed to collect, process, merge, and make sense of the data” (Graves & Hendler, 2013, p. 136), portals provide the ability to play with the data through sophisticated visualizations tools (Artigas & Ae Chun, 2013). Alternatively, for users who wish to view the data and subsequently analyze it, such as activists or journalists, portals also present the same data in a tabular form. Finally, for those who intend to query the data in order to develop some form of open data-based innovation, most portals have built API’s (Application Programming Interface) which allows developers to access the data more rapidly and develop applications.

In the United States, for example, President Obama launched data.gov (Figure 4), the central repository of federal open government data, in his first year in office. Four years later, in another executive order entitled “New Default for Government Information” (2013), the President mandated all government agencies to regularly submit datasets to be made available in the portal. In a similar fashion, the United Kingdom launched its open data portal data.gov.uk (Figure 5) in early 2010. Simpson (2011) notes that “in many ways, data.gov.uk can be considered to be the UK’s counterpart of the U.S.’s data.gov site” (p. 51). Following these two pioneering countries, open data portals began emerging in various countries across the globe, not only at the national level but also at the regional and local levels. Today, more than 130,000 datasets reside on data.gov (Patil, 2015), and more than one million datasets have been made available by governments around the world (Graves & Hendler, 2013).



Figure 4. data.gov (retrieved May 10, 2016)

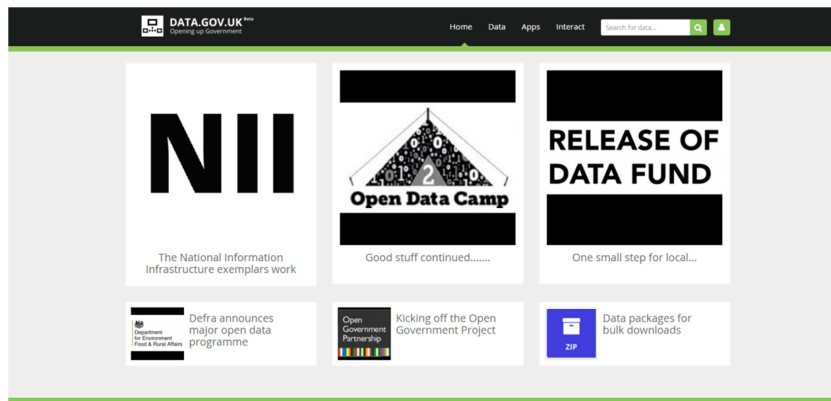


Figure 5. data.gov.uk (retrieved May 10, 2016)

In light of these developments, some research has focused specifically on open data portals. For example, Courmont (2012) analyzes on the politics of legal, economic and technical decision of open data portals, while Braunschweig et al. (2012) examined some of the more technical

aspects of the portals. Other researchers have provided international comparisons of open data strategies by examining open data portals in different countries (Huijboom & Van den Broek, 2011; Petychakis, Vasileiou, Georgis, Mouzakitidis, & Psarras, 2014). For example, based on their comparative analysis of data portals from countries in the European Union, Carvalho and Lafuente (2015) present a useful set of strategies to improve the clarity and usability of data portals.

However, despite impressive progress, open data portals still suffer from several shortcomings (Agrawal et al., 2014; Maali, Cyganiak, & Peristeras, 2010). According to Peled (2013), while the first era of open data portals “suffered from bad design, flawed execution, and adverse consequences” (p. 187), most of these issues have yet to be effectively addressed. In this context, Charalabidis et al. (2014) suggest that governments invest in modern open data portals that are “influenced by the principles of the Web 2.0 paradigm, and oriented towards the elimination of the clear distinction between providers and consumers” (p. 2114).

3.2 The value of open government data

The analysis of open government datasets, often in combination with other data, is increasingly generating new sources of value for society and for the economy (Maali et al., 2010; Stott, 2014; Viscusi, Castelli, & Batini, 2014). However, assessing the value created by the release of open government data is a complex issue.

In response to the criticism that open government data has not yet been able to meet its promise (Horrigan & Rainie, 2015), governments officials are now eager to evaluate the value generated by open government data initiatives. Accordingly, evaluating the value created by open government data has become a central component in current public discourse (UK Cabinet Office, 2012).

Fundamentally, the “key argument for open data is that open access to public sector information is of greater economic, social, and political benefit than exclusive exploitation of this information” (Meijer, Hoog, Twist, Steen, & Scherpenisse, 2014, p. 101). In their review of the state of the art of the open government data, Zuiderwijk et al. (2014) claim that “various studies have confirmed that proactively releasing public and private data in open formats creates considerable benefits for citizens, researchers, companies and other stakeholders” (p. 1). Open government data is anticipated to be a driver of efficiency in both the private and public sectors, increase transparency and citizen participation, and foster innovation (Halonen, 2012; Jetzek et al., 2014; Kulk & Loenen, 2012; Longo, 2011). Moreover, open government data is potentially valuable in a wide range of domains, including health (Iglesias & Boyera, 2015), education (van Brakel, 2014), public safety (Jetzek et al., 2014), food security and agriculture (Jellema, Meijninger, & Addison, 2015), and global development (Carolan, 2015).

Broadly, the value generated by open government data is predominately assumed to be either of economic or of a social nature (Belianska, 2014). This can be explained by the existence of two main ideologies that drive most of the open government data initiatives: the “Re-use of Data” perspective and the “Open Government” perspective (Jetzek et al., 2013). While the latter is

mostly centered on the contribution of social value through the use of open government data, the former is more focused on the economic value (Jetzek et al., 2013; Zuiderwijk & Janssen, 2014b). On this matter, Halonen (2012) cautions that “we must realize the difference between transparency and democracy oriented goals that are usually associated with the freedom-of-information movement and the technology and innovation-oriented goals of the open-data movement” (p. 10).

Most impact studies tend to focus exclusively on the economic gains of open data (Meijer et al., 2014). However, in spite of their relevance and usefulness, quantifying the effect of cost savings or economic growth directly or indirectly derived from open government data is a difficult task (Granickas, 2013). The two main approaches that have been used to measure the economic value of open government are: a) region-wide, and b) sector-wide.

The region-wide approach consists broadly of the analysis of the economic benefit of open government data within a particular region (e.g., global, national, or local). For example, the consulting firm McKinsey predicted the global value of open data to be over \$3 trillion annually, resulting mainly from “improving the efficiency and effectiveness of existing processes” (Manyika et al., 2013, p. 1). It is also anticipated that open data could increase the output of G20 countries by \$13 trillion over five years (Gruen, Houghton, & Tooth, 2014). Other studies have also been undertaken to estimate the economic value of open government data. Most notably, in 2011, the European Commission announced that its open data strategy would correspond to a €40 billion annual boost to the European economy (Kroes, 2011). Similarly, in Australia it was suggested that more vigorous open data policies could add around AU\$ 16

billion annually to the Australian economy (Gruen et al., 2014). Likewise, in Spain, the use of open government data is estimated to generate a total of 4,000 jobs, and €330-550 million annually (Aporta, 2012).

Alternatively to a region-wide approach, sector-wide estimations typically attempt to quantify the cumulative value, or market size, of products and services derived from open government data. The two most compelling sector-wide estimates of the economic value of open government data relate to weather and Global Positioning System (GPS) data.

The public release of weather data from government satellites and ground stations has generated an entire economic sector that gave rise to a plethora of services such as weather forecast, commercial agricultural advising, new insurance options (The White House, 2013), resulting in a market size of billions of dollars/Euros (Cashman, 2014; Jetzek et al., 2013).

Similarly, the decision by the U.S. administration in the 1980s to make GPS available for civilian and commercial access, followed by the European's Galileo system in the early 1990s, led to the emergence of an entire new industry of GPS-powered innovations, ranging from aircraft navigation systems to precision farming to location-based applications. The annual market for global navigation satellite products and services sector is valued at 90 billion dollars in the U.S. (Cashman, 2014) and 124 billion Euros globally (European Commission, 2013). In spite of the difficulty of precisely evaluating the overall economic value of open government data in a given scenario, the order of magnitude of the current estimates provides a forceful illustration of the potential benefit of publishing public data in an open format.

In contrast to the more quantitative approach of the economic impact driven studies, other researchers have attempted to assess the social impact of open government data (Keserú & Chan, 2015; Meng, 2014). The latter is described by Meng (2014) as the “increased inclusion of marginalized groups in society, particularly policy making and access to government services” (p. 1). This perspective is mainly based on “the premise that public resources should be used to increase value, not only in an economic sense but also more broadly in terms of what is valued by citizens and communities” (Jetzek et al., 2014, p. 104).

In this context, Halolen (2012) argues that further study of the social impact of open government data should be encouraged because “the transparency discourse has been gradually reframed to focus mostly on the information market and government performance instead of democratic accountability” (p. 31). On a similar note, Meijer et al. (2014) argues “against sweeping statements about the impact of open data and in favor of studying contextual interactions” (p. 102).

In the last few years, some important contributions to this discussion have pointed out various mechanisms that are critical to the actualization of the creation of social value from open government data. These are, for example, public accountability (Worthy, 2013), citizen participation and empowerment (Kassen, 2013; Luna-Reyes & Chun, 2012; Peixoto, 2013; Sandoval-Almazan, 2012), and best practices to release data and implement tools and policies (Cerrillo-i-Martínez, 2012; Meng, 2014). Essentially, by being able to access information about their government (e.g. budgets, public officials’ salaries, public contracting), civil society advocates can hold governments more accountable for their actions, and support a more

informed public dialogue (Davies, 2013). Furthermore, the release of open government data has also been shown to be potentially beneficial to a range of civic matters, such as advocacy and fundraising (Hall, Shadbolt, Tiropanis, O'Hara, & Davies, 2012), education (A. J. Meijer, 2007; van Brakel, 2014), data-intensive science (Jeffery, Asserson, Houssos, Brasse, & Jörg, 2014), human rights (Beghin, 2014), and international aid (D Linders, 2012; Dennis Linders, 2013).

Below (

Table 2), we present a summary of the main analyses produced regarding the value of open government data, including both the economic and social perspectives.

	Authors	Region/Sector	Estimated value
ECONOMIC VALUE	(Manyika et al., 2013)	Globally	\$3 trillion annually
	(Gruen et al., 2014)	G20 countries	\$13 trillion increase over five years
	(Kroes, 2011)	European Union	€40 billion annually
	(Gruen et al., 2014)	Australia	AU\$ 16 billion annually
	(Aporta, 2012)	Spain	4,000 jobs, and €330-550 million annually
SOCIAL VALUE	(Keserú & Chan, 2015) (Meng, 2014) (Jetzek et al., 2014)	Increased inclusion of marginalized groups	
	(A. Meijer et al., 2014) (Halonen, 2012) (Cerrillo-i-Martínez, 2012)	Government transparency	
	(Worthy, 2013) (Davies, 2013)	Public accountability	
	(Peixoto, 2013) (R Sandoval-Almazan, 2012) (Luna-Reyes & Chun, 2012) (Kassen, 2013)	Citizen participation and empowerment	
	(Beghin, 2014) (Linders, 2012)	Human rights / International aid	

Table 2. Summary of the estimated value of open government data

It is also worth noting that despite the distinction made between economic and social value, there is an evident interdependence between the two concepts. McLeod (2012) observes that, broadly, the open data agenda “takes information access beyond transparency and accountability into the realm of innovation, enterprise, economic growth and return on

investment” (p. 96). Simultaneously, Cerrillo-i-Martínez (2012) points out that even in cases where the main intention is not centered on government transparency, the use of public sector information can also provide greater transparency (p. 770).

Further insights into the interdependence between the economic and social value generated by open government data can be found in the notion that by being actively transparent about their spending (which, due to its importance for public sector scrutiny, typically correlates with social value), governments may simultaneously amass significant savings because they avoid unnecessary spending (Tong et al., 2013; United Nations, 2010). For example, Eaves (2010) illustrates this interdependence by reporting a case in which release of the Annual Information Returns dataset (which includes information on financial and non-financial operations of Canadian charities) allowed activists to uncover that 3.2 billion dollars had been illegally sheltered.

Still, despite the above-mentioned potential of open government data to create value, the increased publication of raw data can also bring about undesirable effects. Meijer (2009) cautions that decontextualized forms of transparency can result in the loss of societal trust. On a similar fashion, Dawes (2012) posits that “instead of providing more transparency, the consumption of this data can actually threaten public trust because the data is removed from shared social experience” (p. 2).

To date, there have been various accounts of unintended consequences, or negative side effects, of opening data (Zuiderwijk et al., 2014; Zuiderwijk & Janssen, 2014c). The release of government data has been shown to, for example, create pressure on governmental agencies to

hide public data (Meijer et al., 2014), exacerbate social injustices (J. A. Johnson, 2014), and decrease the reporting of local crimes (Gibbons, 2011). A well-documented account of the possible negative consequences of open government data relates to the Bhoomi Program, which consisted of digitizing land records in districts surrounding Bangalore (Benjamin, Bhuvanewari, & Rajan, 2007). Ethnographic research-based findings revealed that the “newly available access to land ownership and title information in Bangalore was primarily being put to use by middle and upper income people and by corporations to gain ownership of land from the marginalized and the poor” (Gurstein, 2011, para. 11). This resulted, in practice, in big players being able to use public information to legally challenge and exploit land titles, take advantage of mistakes in documentation, and even identify opportunities and targets for bribery.

In this context, the recent pivot towards more innovation-oriented open data strategies (and the reduction in the attention given to the democratizing potential of the data) in countries such as the United States and the United Kingdom, has been received with some disapproval by civil society advocates (Bates, 2014; J. A. Johnson, 2014). Fundamentally, critics contend that the real agenda of the open government data trend is based on prevailing capitalist interests that enable the private sector to get access to expensively produced data for no cost (Bates, 2012). These idea has also been broadly defended by other authors who have expressed concerns about the privatization of public data (Chignard, 2012; Davies, 2014c; Gurstein, 2011).

3.3 Conclusions

In this chapter, we described in detail the role of government as the sole supplier of open government data. We began by introducing the notion that, until the early 2010's, the focus of the supply-side was on creating a transition from a "closed to open system", which resulted mainly of getting the data out. Accordingly, the extant research was based on a more utilitarian approach toward technology.

However, with the rapid adoption of open data infrastructures in recent years (Zuiderwijk et al., 2014), research on the supply-side has evolved "beyond transparency" (Goldstein & Dyson, 2013). Specifically, more recent research has addressed aspects that are important to the actualization of value from the data. Due to its importance in the open government data process, one of the aspects granted attention by researchers and practitioners, was the issue of open data portals. Accordingly, we briefly examined open data portals and provided examples.

Finally, we analyzed the main contribution regarding the value of open government data by distinguishing between the economic and social value of the data. While in the economic realm, various (region-wide and sector-wide) studies have been able to produce significantly large quantifiable estimates of open government data, the social value of the data has been associated with less easily quantifiable issues such as public accountability, citizen participation and empowerment, and other civic matters. We conclude the chapter by considering instances in which the release of open government data has generated unintended negative consequences.

Despite the many successful examples frequently mentioned in the public discourse, research on the effects of open government data is still rather limited (Craveiro, Machado, & Machado,

2015; Meijer et al., 2014). Scholars have caution that the study and practice of open government data is still in its infancy, and therefore, the purported value of the data still remains largely unclear (Cranefield et al., 2014; Jetzek et al., 2014). This challenge is further amplified by the fact that open government data is often used in different scenarios (i.e. diverse objectives or benefits for different users), which increases complexity (Halonen, 2012; Zuiderwijk & Janssen, 2014b). In the next chapter we examine the demand-side of open government data, thus focusing on the various external stakeholders that use government data in innovative ways, thus creating value from it.

4 OPEN GOVERNMENT DATA: DEMAND-SIDE

With the recent widespread development of information technologies, such as the semantic web and social media, civil society is now much more equipped to collaboratively produce, share, distribute, and innovate with data (Ubaldi, 2013; Van Veenstra & Van Den Broek, 2013). However, considered by some to be the raw material of the 21st century (Halonen, 2012), data must be located, extracted, and refined in order to yield value (Dander, 2013; Fayyad et al., 1996). This implies that “data does not automatically create value” (Cranefield et al., 2014, p. 3). Therefore, when released in its raw form (i.e. not yet converted into linked data or into human-readable form), open government data is not open to the public in any meaningful sense (Cornford et al., 2013; Gitelman & Jackson, 2013).

This chapter describes the main stakeholders that commercially or non-commercially create value from open government data. In a nutshell, we examine the demand-side of open government data. We begin (in section 4.1) by introducing the concept of the open government data ecosystem, which is a useful metaphor when studying the demand-side of open government data. Next (in section 4.2), we individually introduce the main stakeholders that create value from open government data. In consideration of the main goal of this dissertation, we subsequently focus (in section 4.2.3) on the commercial use of open government data by private sector organizations.

4.1 Open government data ecosystem

The first examples of open government data use, either by public transparency activists or profit-seeking businesses, can be traced back to the 1990s (Gray, 2014; J. Woods, 2001). According to Yu and Robinson (2012), the first major project to notoriously explore government data in an open format was created in 1998. This project was openSecrets.org, a website that combined government data with third-party innovation to allow users to search and analyze campaign finance disclosures.

Today, open government data users are drawn from a number of domains and from across the political spectrum (Bates, 2012). In fact, “many stakeholders are involved in open data processes, including open data providers, open data legislators, open data facilitators and many different types of open data users, such as citizens, researchers, journalists, developers, entrepreneurs, archivists or librarians” (Zuiderwijk et al., 2014, p. 2).

Arguably, the most successful endeavor to develop a system-wide framework that describes the dynamics of open government data between the multiple stakeholders was proposed by Harrison et al. (2012). In their influential work, Harrison et al. propose the concept of an “open government data ecosystem.” The underlying idea of this metaphor– the ecosystem – conveys “a sense of the interdependent social systems of actors, organizations, material infrastructures, and symbolic resources that can be created in technology-enabled, information-intensive social systems” (Harrison et al., 2012, p. 900).

The ecosystem analogy “has proved beneficial to various paths of investigation previously” (Heimstädt et al., 2014a, p. 5). In the case of the open government data ecosystem, the central

resource is open government data, and this is in fact what differentiates it from other ecosystem analogies (Heimstädt, Saunderson, & Heath, 2014b). Largely, Harrison et al. (2012) suggest an open government data ecosystem is composed of three main domains: “government policies and practices”; “users, businesses, and civil society”; and “innovators”. Each domain is expected to commit to stimulating innovation and operate as part of an interdependent information-intensive social system.

Other authors have subsequently resorted to the ecosystem metaphor to describe the dynamics between different stakeholders of open government data. For example, Heimstädt et al. (2014b) draw on the literature of business ecosystems to describe open data ecosystems as “organizations and individuals, which generate, share and process datasets mainly within their natural boundaries” (p. 131). In a different approach, Davies (2011) posits that the “successful realization of impacts from open data relies on more than the dataset, involving the mobilization of a wide range of technical, social and political resources, and on interventions beyond dataset supply to support coordination of activity around datasets.” (p. 1) Alternatively, Ubaldi (2013) sums up these dynamics as “open-source governance.” Based on a more collaboration-centered approach, Granickas (2013) introduces the idea of “open government data community” as “an organized network of organizations, institutions, governments, other sectors of society and members of general public with a common agenda to release and reuse open government data and common interest in releasing and reusing open government data (either by creating tools or by using an end product), who collaborate by sharing ideas, information and other resources as well as simply by using open government data products and

services and also recognizing a vital role open data plays in creating these products and services” (p. 6). Immonen et al. (2014) further develop the argument of an ecosystem as follows:

“The actors’ roles in the open data value chain are commonly divided into the data publishing roles and the data end-user’s roles. In addition to data providers and data users, the ecosystem also provides opportunities for other actors that provide tools and other utilities for the ecosystem actors.” (p. 287)

These contributions emphasize the notion, which is well-accepted by now, that stakeholders that are external to the government are central to the sustainability of open government initiatives (Granickas, 2015; Keserú, 2015). A description of the main stakeholders of open government data follows.

4.2 External stakeholders of open government data

Mapping key stakeholders and understanding the dominant incentives is relevant to both open government data research and practice (Granickas, 2015). However, the demand for open government data comprises a complex community of different interests, which makes it particularly challenging to describe (Bates, 2012). Broadly, external stakeholders of open government data may be grouped into 1) citizens, 2) non-profits, and 3) and businesses. We proceed to examine each group in more detail.

4.2.1 *Citizens*

Citizen participation is a critical aspect of the demand-side of open government data (Meng, 2014; Peixoto, 2013; Viscusi et al., 2014). The exploration, analysis, and presentation of public data by civic-minded individuals is essential to the empowerment of citizens to scrutinize government and hold authorities accountable (Cucciniello, Nasi, & Valotti, 2012; Ruppert, 2013). By being able to access data about their government (e.g. budgets, public officials' salaries, public contracting), citizens can hold governments more accountable for their actions and support a more informed public dialogue (Davies, 2013).

In this context, the philosophy of open government data can be seen as a traditional "good governance" principle (Huijboom & Van den Broek, 2011). In this way, "citizens asking, for example, for information on how budgets are spent, should no longer have to trawl through hundreds of printed pages" (Davies, 2013, p. 19).

When considering the use open government data by citizens, one must first acknowledge the case of civic developers. Broadly, a civic developer is a computer and statistics-savvy individual motivated by principles of more government transparency and accountability or captivated by the vast potential of open government data for innovation (Wihbey, 2013). The work produced by civic developers, such as civic-minded websites and mobile applications, is particularly relevant in the overall open government data ecosystem because it is typically centered around the needs of communities, especially those with low/average statistical or computational skills to extract meaning from the data (Gurstein, 2011; Janssen & Zuiderwijk, 2014).

Commonly, civic developers meet in a group setting in order to collaboratively create technological solutions that benefit their community. These meetings, often called *hackathons*, can also be seen as a form of civic engagement (Johnson & Robinson, 2014). Regular *hackathons*, occurring weekly or monthly, generally consist of volunteering citizen group meetings where members of the local community tackle civic issues relevant to them. Alternatively, *hackathons* can also be initiated by some organization, typically a government agency, in the form of “challenge competitions, where participants compete to provide the best idea or develop the most useful application with the available data” (Yang & Kankanhalli, 2013, p. 644). In this case, *hackathons* are set up with the goal of stimulating the use of open government data for innovation (Bertot et al., 2012; Kube et al., 2013). In addition to generating civic solutions, *hackathons* also provide suitable conditions for entrepreneurs and small businesses to connect with venture capitalists and showcase their ideas for applications of open government data (Vasa & Tamilselvam, 2014).

The growth of the open government data movement in the last few years has been remarkable, captivating a large number of civic-minded individuals¹³. Despite the recent proliferation of developer groups around the world that are dedicated to open government data, very little is still known about this trend (Kuk & Davies, 2011).

Additionally, the use of open government data by citizens “who synthesize content by combining data and information from diverse sources in new ways” (Cranefield et al., 2014, p.

¹³ For instance, the 2017 Open Data Day (an annual celebration of open data, where groups from around the world create local events to use open data in their communities) has 346 registered events around the globe on the same day.

3) can also be found in other areas beyond civic developers. Individuals, such as journalists, and researchers are often credited with the development of innovative applications using open government data (Chan, 2013), or may use the data as part of an investigative process (Aitamurto, Sirkkunen, & Lehtonen, 2011).

For instance, open government data is currently a critical component of modern day journalism¹⁴, where it is often used as an instrument for investigation or reporting (Aitamurto et al., 2011; Appelgren & Nygren, 2014). So much so that Rogers argues that “data journalism has its roots in publicly available data” (Rogers, 2012, para 7). In a complementary perspective, Johnson (2014) suggests that investigative and data journalism can provide alternatives to direct participation in the context of an “information justice movement”.

The publication of public sector data in a machine-readable format and free-of-cost is also of immense benefit to academic research (Arzberger et al., 2004; Buteau, Larquemin, & Mukhopadhyay, 2015; Whitmore, 2014). It is worth noting that the use of open government data by academic researchers should not, however, be confused with open access, i.e., unrestricted online access to publicly funded research or any peer-reviewed scholarly research (OECD, 2012). By having access to open government data available, researchers can have “easy online access to vast amounts of data without the need for repeated data requests, transcription of data from print to electronic formats, and other tasks that would limit user interest and data usefulness” (Whitmore, 2014, p. 623).

¹⁴ Yankova (2016) provides many examples of how open government data can further enrich the work of journalists.

The issue of citizen participation in the realm of open government data is a growing topic of interest (Kalampokis, Hausenblas, & Tarabanis, 2011; Yu & Robinson, 2012). Linders & Wilson (2011) distinguishes citizen participation mechanisms into two discrete forms: 1) collecting opinions (citizen engagement), and 2) collecting ideas and solutions, or crowdsourcing (Linders & Wilson, 2011).

A relevant example of citizen engagement can be found in participatory budgeting initiatives around the world, which has been widely discussed in literature and is now instituted in at least 1,500 cities worldwide (Baiocchi & Ganuza, 2014). In effect, research focusing on the impact of participatory budgeting on municipal expenditures suggests that allowing the public to vote on the allocation of public funding can result in increased political accountability and responsiveness (Edwards, 2014).

Additionally, citizen collaboration in the realm of open government data can be found, for example, in the context of disaster relief, where crowdsourcing activities of data collection have been provided tremendous value (Jetzek et al., 2013; Linders, 2012).

However, when considering the sweeping potential of open government data to leverage public participation, Peixoto (2013) cautions that:

“... to date, mechanisms of participation related to open data largely have been limited to ad hoc events in which technologists and interested parties collaborate on software related projects (such as hackathons and competitions) [...] They do not, however, replace participatory institutions

designed to leverage the dispersed knowledge of citizens to shape decisions that affect their lives."

(p. 208)

On a similar note, Meijer et al. (2012) assert that "open government is much too important to leave it to the 'techies': scientists and practitioners with backgrounds in law, economics, political science and public administration should also get involved to build sound connections between vision and voice" (p. 26). According to this perspective, non-computational skills (e.g. understanding user needs, product design, marketing) should be embraced substantially when developing community-driven solutions based on open government data (Kuk & Davies, 2011).

4.2.2 *Non-profits*

Upon acknowledging the conceptual and practical difference between *data* and *information*, Gurstein (2011) differentiates "between the opportunity for the digitally enabled activity presented by information and communications technology (ICT) access, from the actual realization of those opportunities in the form of *effective use*" (para. 4). In order to make an effective use of open government data, Gurstein (2011) maintains that one should have "access to the bandwidth, the computers, the software, and the organizational structures, or "to the financial or other resources" (para. 23). Also, other factors such as the lack of context in data, insufficient data literacy, and negative attitudes toward public officials have been suggested as potential obstacles in the adoption of open government data by citizens (Halonen, 2012; Zuiderwijk et al., 2012).

As a result, more likely than not, a common citizen is unable to explore the array of datasets available. Essentially, the complexity and overwhelming amount of data make it difficult for users to find useful patterns, trends, rules, and correlations in the data (Zurada & Karwowski, 2011). Besides this, most agencies do not possess the strength of innovation to respond to this challenge effectively (Chan, 2013). In this context, the emergence of external intermediaries that help turn raw data into actionable information is a logical development (Tennison, 2015; Yang & Kankanhalli, 2013; Yu & Robinson, 2012).

An intermediary can be defined as “any sort of actor that bridges two or more different perspectives, establishing communication mechanisms that let them interact fluently and in understandable ways” (González-Zapata & Heeks, 2015, p. 4). Importantly, the notion of intermediation took a more central role in the early days of the Internet when the exchange of digital information skyrocketed (Rose, 1999). In the specific domain of open government data, intermediaries are also a crucial aspect of the open government data ecosystem (Schalkwyk et al., 2015) through their “acts of sanitising, organising, compiling, formatting, and documenting available open government data” (Chattapadhyay, 2014, p. 7).

When in the context of a non-profit endeavor, for example in the form of a Civil Society Organization (CSO) or a Non-Governmental Organization (NGO), open government data intermediaries strive to bridge the barriers that separate public sector data producers from civil society data consumers (González-Zapata & Heeks, 2015). Several noteworthy examples of non-profit organizations have recently emerged with the goal of helping citizens extract the most knowledge from public data and make governments accountable (Janssen et al., 2012; Longo,

2011; Mayer-Schönberger, 2012; Robinson et al., 2009). Motivated by the potential of open government data to change lives, these organizations hope to move beyond the rhetoric of transparency towards direct action and problem solving at the community level (Goldstein & Dyson, 2013). Chattapadhyay (2014) describes the role of an open government data intermediary as follows:

“... the intermediary organisation would help local communities to engage and transact with various service delivery agencies (public and private) more effectively – and also in the context of information – where the intermediary organization would focus on improving the access to and use of information for its intended audiences.” (p. 4)

An interesting example of this kind of approach is presented by Capital Area Food Bank (CAFB), a non-profit organization based in Washington D.C. that tackles the issue of hunger in the area. While working with more than 400 local community groups, including many churches and supermarkets, CAFB is using open government data to collate a heat map that shows where food insecurity is at its highest points. In this way, CAFB is able to monitor and predict donation levels so it can match supply and demand more effectively (Hargrave, 2016).

However, although it is considered a vital element in the open government data ecosystem, intermediaries has received limited attention, mostly when compared to other stakeholders (Davies, 2014a). Still, a few studies in this area are worth mentioning. For example, Van Schalkwyk et al. (2014) study the use of open government data by university planners, and conclude that intermediaries are found to play several important roles in the ecosystem, such as increasing the accessibility and utility of data, and democratizing the use of the data. Similarly,

González-Zapata and Heeks (2015) study the role of open government data intermediaries in Latin America. Using a somewhat different approach, Mercado-Lara and Gil-Garcia (2014) examine the specific case of AidData ¹⁵. Their study suggests that variables such as organizational structures and processes, institutional arrangements, or contextual factors, may influence, for instance, the methodologies and the final products offered by AidData. Alternatively, Chattapadhyay (2014) examines the actual practices by which non-profit intermediaries access and share open government data. In attempt to narrow the gap in empirically-based research on open data intermediaries, Schalkwyk et al. (2015) contend that “consideration needs to be given to the presence of multiple intermediaries in an open data ecosystem, each of whom may possess different forms of capital to enable the use and unlock the potential impact of open data” (p. 2).

Alternatively, intermediaries can also take the role of facilitators, thereby assisting citizens in extracting knowledge from public data. In this case, although there is not an intermediation of the access or use of the data *per se*, these organizations can fund, support, or organize the development of solutions that use open government data for the benefit of society. In alignment with the dissemination of the fundamental principles of the open government paradigm, several examples of civil society organizations of this kind have emerged recently. In the US, the non-profit organization Code for America was founded in 2009 to partner with state, county, and city governments to redesign public services by deploying temporarily teams of developers

¹⁵ AidData is a research and innovation lab that seeks to make development finance more transparent, accountable, and effective by tracking funding for development around the world.
<http://www.aiddata.org/>

in the actual government to collaborate in civic solutions. Code for America has been a pioneer in this space by providing the means for *hacktivists* as well as for entrepreneurs to develop solutions based on open government data. Other examples of open government data facilitators that have been instrumental in making government and politics more accountable and transparent are the Sunlight Foundation, an American non-profit organization devoted to making the government more transparent and accountable through data, tools, policy and journalism, and the World Wide Web Foundation, which has sponsored various activities related to the research and practice of open government data.

4.2.3 *Businesses*

The commercial exploitation of government data by the private sector can be traced back to the 1990's (J. Woods, 2001). However, the large expansion of the commercial use of open government data occurred mainly during the Government 2.0 era, where private stakeholders were "given a greater role in managing information" (Borglund & Engvall, 2014, p. 166). On this, Dietrich (2015) notes the following:

"Before the digital age, traditional intermediaries of PSI [Public Sector Information] used to be a relatively small group of experts and specialists in science, specialized consultancy agencies, law firm and the media. Although these experts remain relevant, the rise of the digital age, the increased public access to the Internet and other communication technologies as well as the increase availability of information and data opens a space for new intermediaries." (p.6)

Today, it is not uncommon for firms to leverage on their know-how and infrastructure to make use of open government data in innovative ways (Dawes, 2012; Howard, 2013; Longo, 2011). In this context, it is anticipated that the commercial use of open government data has the potential to boost innovation, and to create jobs and wealth (Manyika et al., 2013; Tinholt, 2012). Ultimately, it can be said that open government data offers the possibility for “a new start for economic reform and productivity growth” (Gruen et al., 2014, p. 2). With this in mind, various governments around the world have recently committed to the promotion of entrepreneurship, innovation, and scientific discovery to generate the development services based on open government data.

The commercial use of open government data by the private sector is a central concept in this dissertation. Accordingly, in the following section, we analyze the main types of commercial users of open government data (section 4.2.3.1), as well as their motivations for using the data (section 4.2.3.2).

4.2.3.1 Types of commercial users

In spite of the rising interest in open government data, a common understanding of the main concepts and terms is still missing (Jetzek et al., 2012; Lindman et al., 2013). In the realm of the commercial use in particular, the terms referring to private-sector organizations that use open government data are often used interchangeably in the literature. These can broadly be distinguished into three distinct groups: *civic startups*, *open data services*, and *infomediaries* (Magalhaes et al., 2013). A brief description of each group follows.

Open government data is said to be “shaping the future of civic life” (Voss, 2015, para. 15). Importantly, this notion is strongly connected with the concept of “civic innovation”, which broadly refers to the development of technology aiming to improve governments transparency and accountability, and citizen-government (Dietrich, 2015; Patel, Sotsky, Gourley, & Houghton, 2013). When developed in a commercial context, civic innovation is often credited as the output of *civic startups*.

Civic startups can be described as private sector organizations that adopt of a civic-oriented mission to create and sustain value that contributes to a more efficient open government ecosystem, and ultimately to the public good (Living Cities, 2012) (see Table 3 for an example of open government data use by a civic startup). The organization Code for America, an influential actor in the open government ecosystem, defines *civic startups* as “having one of three models: 1) they provide services on top of open data; 2) they bring modern web technologies directly to governments; 3) they change the way citizens ask, get, or need services from government” (Tran, 2012). Accordingly, some products and services developed by *civic startups* make use of open government data. In integration with open government data, these products and services can also take advantage of other data-centered technologies (e.g. social media, participatory design, co-creation, and crowdsourcing) (Goldstein & Dyson, 2013). The issue of *civic startups*, and in particular how these firms create value, is still poorly understood (Living Cities, 2012).

Type of user: Civic startup

Company Name: AreaVibe

Company Description: AreaVibe works to provide users with information about every city in the United States.

Use of open government data: It uses various datasets related to the quality of life in different cities (e.g. crime rates, demographics, weather) so that users can find the best places to live.

Table 3. Example of open government data use by a civic startup

Other authors have adopted the term “infomediaries” (information + intermediaries) when referring to private sector organizations that make use of open government datasets. This term is not unique to the open government data discourse, and in fact, may take different meanings depending on the context.

The concept of *infomediaries* was first proposed by Hagel and Rayport (1997) following their prediction that consumers were going “to take ownership of information about themselves and demand value in exchange for it” (p. 53). Originally, the term was introduced to convey the idea of an intermediary agent acting in the internet service industry. The term was subsequently used by other authors in the field (Song & Zahedi, 2002). However, with the sharp increase in the amount and diversity of available data, the role of *infomediaries* has grown more prominent. Consequently, since its initial use, the concept has been adopted in other areas, such as library and information science (Elizabeth Gould & Gomez, 2010), knowledge management (Lamb, 2001), and political science (Latham, 2003).

In one of the earliest uses of the term in the context of access to public information, in the year 2003, the United Nations (2003) referred to “infomediaries” that “synthesize, translate, simplify and direct information on behalf of others” (p. 4). Acting as an information-based liaison between the producer and user of data, *infomediaries* are commonly associated with the development of commercial solutions that attempt to translate raw data into meaningful information for end-users (Howard, 2012; Rufus Pollock, 2011) (see Table 4 for an example of open government data use by an infomediary firm). Fundamentally, *infomediaries* develop products or services based on open government data (O’Hara, 2012b). An interesting example in the literature is found in the work by Gould & Gomez (2010) who make use of the concept of *infomediaries* to study how libraries, telecenters and cybercafés help support public access to information and communication technologies in developing countries.

<p>Type of user: Infomediary</p> <p>Company Name: FlightAware</p> <p>Description: FlightAware offers flight tracking services for both private and commercial air traffic.</p> <p>Use of open government data: It uses open data from the Federal Aviation Administration of the U.S. government on aeronautical and general aviation information to deliver user-friendly information on flight tracking.</p>
--

Table 4. Example of open government data use by an infomediary firm

Finally, other authors have introduced the notion of “open data services” in the context of the commercial use of open government data. For example, Tammisto & Lindman (2012) classify open data according to its degree of openness (internal vs. external) and technical format. Accordingly, *open data services* describe the use of open data external to the firms’ boundaries (e.g. open government data), as well as internal open data. While external open data typically corresponds to open government data (see Table 5 for an example of open government data use by an open data service), in the case of internal open data it corresponds to data that is made accessible only to a restricted group of people (e.g. employees of the organization, networks, and communities). Differently from *civic startups* and *infomediaries*, the concept of *open data services* is not centered on the notion of government data, but of any kind of open data. Consequently, it becomes useful to the study of open data both at an intra-organizational and an inter-organizational levels (Lindman et al., 2013; Tammisto & Lindman, 2012).

<p>Type of user: Open data service</p> <p>Company Name: StreetCred Software</p> <p>Company Description: StreetCred Software helps law enforcement agencies manage their arrest warrants, eliminate warrant backlogs and improve efficiency while increasing officer safety.</p> <p>Use of open government data: It uses open data from courts, as well as warrant data that is available only to law enforcement officers, to help agencies understand their fugitive and gang populations and relationships, so that they can measure effectiveness and make improvements.</p>

Table 5. Example of open government data use by an open data service

4.2.3.2 Motivations for open government data use

It is widely accepted that the combination of the increasing availability of open government data together with other factors, such as the “inadequacy of existing services, [and] the availability of new techniques to process data” (Stott, 2014, p. 8), gives private stakeholders a larger role in the management of the data-information (Borglund & Engvall, 2014; Chan, 2013; Vickery, 2011). Accordingly, a wide range of “opportunities for entrepreneurial researchers to use government databases” (OECD, 2012, para. 4) has culminated in a multitude of commercial applications using open government data.

Plenty of examples of products and services that creatively exploit public datasets in some shape or form have emerged. Examples include large digital conglomerates such as Amazon and Google, who provide free access to public datasets through user-friendly interactive platforms, and smaller businesses, such as real estate companies that use open government data to help home buyers make better decisions (Park & Vanroekel, 2013).

Largely, the commercial use of open government data can be fundamentally distinguished into two types of use: to develop products or services; and to increase competitive advantage (Gurin, 2014; Zuiderwijk, Janssen, Poulis, et al., 2015). The first type of use can be basically described as the development of products or services based on open government data, which are then marketed to third parties (Aporta, 2012; Immonen et al., 2014; Stott, 2014). (See examples of this type of use in Table 6). The market space comprising these firms is sometimes referred to as the “open data economy” (Howard, 2013).

(Example 1)

Company Name: CommoPrices

Company Description: CommoPrices enables users to track the price of any commodity on a single portal

Use of open government data: In order to collect easily accessible and affordable market data regarding commodity price references, CommoPrices fetches data from French Customs. This allows the company to build a single portal integrating the widest range of commodities.

(Example 2)

Company Name: Openlaws

Company Description: Openlaws aims to help companies to achieve legal compliance and create better access to justice for citizens by providing personalized legal information via the online platform www.openlaws.com.

Use of open government data: Openlaws connects to open government data sources from various countries in the European Union in order to gather extensive information on legislation and case law applicable to each country or the European Union area.

Table 6. Examples of new businesses based on open government data

The second type of use relates the use of open government data, sometimes in combination with other data sources, to achieve or sustain competitive advantage (Gurin, 2014; Hartmann et al., 2014). (See examples of this type of use Table 7) In this case, open government data – from a wide range of domains such as health, energy, education, traffic, weather, public sector budgeting – is used as a complementary data source to improve existing processes or strategies, thereby increasing competitive advantage (Zuiderwijk, Janssen, Poulis, et al., 2015). Considering

that achieving and sustaining competitive advantage is an significant component of a firm's managerial activities and cross-functional decision making (D'Aveni, 1994), open government data may also represent an operational and strategic resource. However, Zuiderwijk et al. (2015) cautions that the actual creation of competitive advantage from open data "might not be simple" (p79).

(Example 1)

Company Name: Zizabi

Company Description: Zizabi is a home shopping marketplace in Portugal, focused on giving home buyers, sellers and renters the information they need to make better decisions about where to live.

Use of open government data: Open government data allows Zizabi to provide users with key additional information that goes beyond square footage and heating details, such as nearby school locations, crime statistics, and transit routes.

(Example 2)

Company Name: Arup

Company Description: Arup provides structural engineering consultancy services, offering building design, economics and planning, infrastructure design, management consulting, and specialist technical services.

Use of open government data: It takes advantage of open government data, such as data about traffic, planning, natural hazards and other topics, to deliver more efficient services and help mitigate against risks, for example of natural disasters.

Table 7. Examples of businesses that use open government data for competitive advantage

The commercial use of open government data by the private sector is by no means limited to online businesses. In fact, there are plenty of examples of firms that operate traditional brick-and-mortar businesses that are currently tapping into the data released by governments. For example, the retail giant Wal-Mart uses weather forecast data to predict sales and improve business intelligence. Specifically, Wal-Mart “regularly uses weather data to increase stocks of a long list of products in stores that lie in the path of hurricanes and other inclement weather patterns” (Bitner, 2014, para. 18). Another interesting example relates to how Coca-Cola used public data in a marketing strategy:

“As a creative example, Coca Cola’s Share a Coke campaign in 2013 replaced the Coca Cola logo on its bottles with 250 different first names in several European countries. The massive soda company relied solely on government name data to select the names to use in each country’s campaign. The campaign was enormously successful as customers got a kick out of buying a Coke bottle with their own name or with the names of loved ones on it. Had the names been uncommon, the campaign would have been a flop.” (Bitner, 2014, para. 14)

In an attempt to make sense of the exploitation of government data by the private sector, some researchers have endeavored to collect, and subsequently analyze, as many examples of use as possible. For example, the Open Data 500 study ¹⁶ conducted by The GovLab at the New York University in 2014 was the first comprehensive study of U.S. companies that use open government data in their business. Its findings are particularly important as it provides a

¹⁶ <http://www.opendata500.com/>

snapshot of the various new businesses, as well as the products and services that are developed based on open government data (Verhulst & Caplan, 2015). In a similar fashion, the “Open Data Impact Map” project, coordinated by the Center for Open Data Enterprise¹⁷ is a searchable, centralized database of open government data use cases from around the world. It is a collaborative effort that draws on previous studies, as well as examples from an international network of regional supporters, an online survey, and thorough research.

4.3 Conclusions

In this chapter we mapped the demand-side of open government data, which comprises the external stakeholders that (commercially or non-commercially) create value from open government data.

We explored the concept of the open government data ecosystem as an interdependent social system of stakeholders of open government data, which lends a collaboration-centered view of the demand-side of open government data. We then introduce individually the main groups of external stakeholders – *Citizens*, *Non-profits*, and *Private sector*.

Regarding the use by citizens, we examine participation mechanisms in the realm of open government data, paying closer attention to the use of data by civic developers, journalists, and academic researchers. We also note the perspective defended by some critics that the use of open government data should not be confined to computationally-skilled individuals.

¹⁷ <http://www.opendataenterprise.org/map.html>

Next, we review the emergence of civil society organizations that use open government data mainly to ameliorate 1) the difficulty of most citizens in exploring the overwhelming amount of data available, and 2) the lack of innovative capacity of government agencies. We also contemplate the anticipated growing importance of this stakeholder group in the process of value creation for other stakeholders in the ecosystem.

Finally, we focus on the private sector users of open government data, which comprises fundamentally firms that leverage on their know-how and infrastructure to make use of open government data in innovative ways. Since the commercial use of open government data is a central idea in this dissertation, we expand on the analysis of this group, and note the lack of a common understanding of the main concepts and terms referring to commercial users in the open government data ecosystem. Specifically, we do so by defining three different concepts found in literature: *civic startups*, *infomediaries*, and *open data services*. Additionally, we make the distinction that businesses use open government data mainly to 1) develop new products or services, and 2) to increase its competitive advantage. We also note that the commercial use of open government data by the private sector is by no means limited to online businesses, and provide various examples of firms that operate traditional brick-and-mortar businesses who tap into the data released by governments.

In Part I of this dissertation (chapters 2, 3, and 4) we examined the extant literature on open government data in order to provide the reader with a broad, yet thorough, understanding of the theoretical scope of this topic. We now proceed to Part II (chapters 5, 6, and 7), where we

present and discuss the empirical work that was executed, as well as the findings regarding the commercial use of open government.

PART II (EMPIRICAL RESEARCH)

5 CREATING VALUE WITH OPEN GOVERNMENT DATA

Current mega trends, such as big data, the Internet of Things, or the tremendous progress in areas such as machine learning and data mining are a testament to the current weight of data in every segment of society. Under these conditions, businesses “themselves are experiencing a transformation encompassing higher incentives towards the valorization of intangible assets” (Iemma, 2012, p. 3).

Simultaneously, in the realm of open government data, it is well accepted that private-sector innovation is vital to the overall development of the ecosystem (Stott, 2014). Harrison et al. (2012) anticipate that “citizens’ expectations, interests in, and appetite for government services and data—what they think is possible, normal, and desirable—will be engendered by what they experience in their interactions with non-governmental and private organizations” (p. 4).

However, the extant literature highlights a general gap concerning how data can actually create value for companies (Hartmann et al., 2014). Also, critics have drawn attention to the limited empirical data available (Cranefield et al., 2014) coupled with “very little systematic and structured research” in the study of the commercial use of open government data (Zuiderwijk & Janssen, 2014b, p. 17).

Accordingly, in this chapter we investigate the different ways in which companies create value from open government data. We attempt to shed light on this issue by analyzing the use of open government data by 178 firms. Grounded in a systematic research approach and solid empirical data, the findings in this chapter address this gap, and offer an important descriptive, yet focused, insight into the practice of the commercial use of open government data.

The findings from this chapter aim to provide both entrepreneurs and intrapreneurs with a view of how open government data can be used innovatively to create new products and services, or to create competitive advantage.

Faced with criticism that open government data has not yet been able to meet its promise (Bates, 2012; Gurstein, 2011), governments are currently under pressure to justify the investments that have been made in open government data initiatives. Consequently, our research also intends to provide government officials with empirical evidence regarding the potential of open government data to generate private sector innovation. This seems particularly important given the overall difficulty among public servants in fully understanding the actual benefits of open government data (Crane, 2013).

This chapter is structured as follows. We begin by introducing the role of the business model concept in the extant literature on the commercial use of open government data. Next, we describe in detail the methodological approach undertaken in order to uncover the different commercial uses or applications of open government data that emerge from the data analyzed in our sample. Upon defining the dimensions of analysis that are adopted, we present our findings, and discuss the various types of commercial use. We conclude by reviewing the contributions and the practical implications of our study, as well as its limitations, and provide recommendations for future research.

5.1 Background

Despite the impressive progress of the concept of business model in the last few years (Lambert, 2006), there is still much discussion on what the term actually means (Zott, Amit, & Massa, 2010). The business model concept seems to have originated with the rise of the Internet, and the subsequent need to better analyze e-business models (Osterwalder, 2004; Steininger, Huntgeburth, & Veit, 2011). Today, however, the *business model* is a strategic conceptual tool used widely (Onetti, Zucchella, Jones, & McDougall-Covin, 2012) that seeks to explain generally both value creation and value capture (Zott et al., 2010).

In this context, the extant literature on the commercial use of open government data has mainly resorted to the business model concept to study the issue of value creation (Ahmadi-Zeleti, Ojo, & Curry, 2014). For example, Gurin (2014) separates the commercial use of open government data into two different models: “*Better Business Through Open Data*” and “*Open Data Pure Plays*”. Similarly, Perricos et al. (2012) suggest five functional business model archetypes in the open data marketplace: “*suppliers*”, “*aggregators*”, “*developers*”, and “*enrichers*”.

In the academic context on open government data, the existing literature that employs the business model concept is still rather scarce and mainly of exploratory nature (Ferro & Osella, 2012, 2013; Foulonneau et al., 2014). For example, the research by Ferro and Osella (2013) “aspires to shed light on the mechanisms allowing profit-oriented value creation based on public datasets” (p. 1). Accordingly, drawing from interviews with 13 firms that commercialize products and services based on open government data, these authors identify eight business models archetypes in the commercial use of open: “*Premium Product*”, “*FreemiumProduct /*

Service”, “Open Source”, “Infrastructural Razor & Blades”, “Demand-Oriented Platform”, “Supply-Oriented Platform”, “Free as Branded Advertising”, and “White-Label Development”. Similarly, Kaasenbrood et al. (2015) make use of the business model concept to study the factors influencing the adoption of open government data in six private sector organizations in the Netherlands. Using a different approach, Janssen & Zuiderwijk (2014) analyze 12 services based on open government data and social media data to identify 6 types of business models: “single-purpose apps”, “interactive apps”, “information aggregators”, “comparison models”, “open data repositories”, and “service platforms”. Finally, Ahmadi-Zeleti et al. (2014) examine a number of studies on the topic of open government data business models to derive 5 main categories: “Freemium”, “Premium”, “Cost Saving category”, “Indirect Benefit”, and the “The Razor-Blade”.

Considering their exploratory nature, these studies can be viewed as stepping stone for richer analyses of the commercial use of open government data. However, they have been limited in responding to the need for more empirical research in this topic, typically relying on small-sized samples.

In light of the general gap concerning how value is created from open government data (Ferro & Osella, 2013; Janssen & Zuiderwijk, 2014), it seems therefore relevant to undertake a more thorough and systematic analysis of the issue of value creation in the commercial use of open government data. By examining the various ways through which businesses make use of open government data, we attempt to offer a solid descriptive foundation for researchers and practitioners.

5.2 Methodology

In this study we follow a qualitative approach in order to “arrive at an understanding of a particular phenomenon from the perspective of those experiencing it” (Vaismoradi, Turunen, & Bondas, 2013, p. 398). We employ the method of thematic analysis as a means to find themes in the data (Boyatzis, 1998).

Thematic analysis should “be seen as a foundational method for qualitative analysis” (Braun & Clarke, 2008, p. 78), and has been described as very similar to Grounded Theory in terms of its procedure for coding from data (Braun & Clarke, 2008). However, differences exist, namely in the timing of the data collection and analysis processes. For example, Grounded Theory becomes unsuitable for researchers who seek to compare two separate sets of data that are gathered at different times (Alhojailan & Ibrahim, 2012). Also, in thematic analysis, the “concept of theme is critical to the accurate interpretation of qualitative data” (DeSantis & Ugarriza, 2000, p. 77). Thematic analysis allows the researcher to systematically mine large amounts of textual data, make inferences, and identify common shared properties concerning the phenomenon under investigation (Holsti, 1968; Krippendorff, 2004; Vaismoradi et al., 2013). Braun & Clarke (2008) indicate that “through its theoretical freedom, thematic analysis provides a flexible and useful research tool, which can potentially provide a rich and detailed, yet complex, account of data”. However, themes are not the exact words, but the exact meanings implied and inferred from words, behaviors, and events (DeSantis & Ugarriza, 2000). Consequently, in contrast to content analysis, thematic analysis incorporates both manifest and latent aspects (Vaismoradi et al., 2013).

Even though some of the phases of thematic analysis are similar to other qualitative research, and therefore not unique, Braun and Clarke (2008) provide a useful outline of the six phases of thematic analysis:

1. Familiarizing yourself with your data: Transcribing data (if necessary), reading and re-reading the data, noting down initial ideas
2. Generating initial codes: Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code
3. Searching for themes: Collating codes into potential themes, gathering all data relevant to each potential theme
4. Reviewing themes: Checking if the themes work in relation to the coded extracts and the entire data set, generating a thematic 'map' of the analysis
5. Defining and naming themes: Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
6. Producing the report: The final opportunity for analysis. Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.

Following the rationale loosely prescribed by thematic analysis, we analyze a dataset that includes information about the use of open government data by for-profit organizations in the

United States. The data was collected in the context of the *Open Data Impact Map*, a project run by the non-profit organization Center for Open Data Enterprise based in Washington, D.C. At its core, the *Open Data Impact Map* aims to demonstrate the value of open government data and to provide a basis for further analysis of the impact of open data globally. The data was collected from January to November in 2015 using various sources and techniques: exhaustive desktop research, which included examining various sources of information (e.g. press releases, interviews, news articles) online, which was complemented with a survey that was sent by email to firms that seemed to be a good fit for the map. The use of various sources is expected to enrich the sampling data and achieve more robust results (Steininger, Trenz, & Veit, 2013).

The dataset, which included nonprofit and for-profit organizations around the world that use open government data, was in the file format .csv, and consisted of the following fields: *Region, Country, Organization Name, Organization Type, Industry, Organization Description, URL, City, State/Region, Founding Year, Size, Data Use, and Entry Based On*. In preparation for our analysis, we proceeded to narrow down the sample so that it contained only relevant information for our study. Accordingly, we included only for-profit organizations located in the United States. Moreover, we kept only the firms whose information was collected using the same method, which we choose to be desktop research because this method of data collection produced significantly more results than the survey method. This process rendered 178 firms that were considered suitable for analysis.

Next, we reduced our sample to the most relevant fields: *Organization Name, Organization Type, Industry, Organization Description, and Data Use* (see Table 8 for an excerpt of data sample).

Lastly, in order to further improve the manageability of the sample, the options in the *Industry* field were collated from 19 to 9, as follows: 1) *Agriculture, Energy, Environment, and Weather* were grouped into *Energy and Environment*; 2) *Business and legal services, Finance and investment, Research and consulting, and Insurance* were grouped into *Business and legal services*; 3) *Consumer services, Housing/real estate, Media and communications* were grouped into *Consumer*; 4) *Security and public safety, and Logistics and security* were grouped into *Logistics and security*; 5) *Data/information technology, and Geospatial/mapping* were grouped into *Data/information technology*; 6) *Education*, 7) *Governance*, 8) *Healthcare*, and 9) *Scientific research* were not changed.

Organization Name	Industry	Organization Description	Data use
360 Yield Center	Energy and Environment	Taking a 360-degree view of key yield-limiting variables - 360 Yield Center develops strategies for better-performing crops and better on-farm profits.	Uses weather data and agriculture data to offer new tools that allow you to measure and supply the right amount of N when the plant is ready to use it.
5PSolutions	Data/information technology	5PSolutions are artisans of mobile platforms.	Uses open data to develop apps such as maps showing UK and US embassies around the world.
Abt Associates	Business and legal services	Abt Associates is a mission-driven organization conducting research and program implementation in the fields of health - social and environmental policy - and international development.	Uses open data to generate research and recommendations on how to tackle social issues (e.g. environment, education, food security).
Accela	Governance	Accela improves citizen engagement by making it easier to do government business - driving civic innovation and improving transparency - accuracy and accountability - providing civic engagement solutions for government.	The Accela Civic Platform uses open government data to build cloud services that automate and streamline civic processes
AccuWeather	Energy and Environment	AccuWeather provides local forecasts for everywhere in the United States and over two million locations worldwide.	Uses open weather data to provide weather forecast services.
Addepar	Business and legal services	Addepar streamlines the increasing complexities of modern investment management into one seamless platform: comprehensive data aggregation - unmatched data integrity - flexible analytics - intuitive reporting - limitless customization and interactive portals.	Aggregates open government data to provide flexible analytics, intuitive reporting, and limitless customization and interactive portals
AgSquared	Energy and Environment	AgSquared online farm planning and management service helps farmers track valuable information about their farm operations to help build a complete picture of a farm's productivity - profitability and sustainability.	the AgSquared software integrates U.S. soil and weather data to help farmers manage their farm.
Aidin	Healthcare	Aidin helps hospitals discharge their patients to the best available nursing homes - home health agencies - and other post-acute care providers.	Uses healthcare open data to recommend the most appropriate providers.

Table 8. Excerpt of data sample

In light of the preference for clear and concise guidelines in the process of thematic analysis, thus avoiding the “anything goes” critique (Braun & Clarke, 2008), we derive our initial coding frame from Davies (2010), who studies the implications of the use of open government data in democratic change and public sector reform. Although Davies does not focus specifically on the commercial use intermediation of open government data, he suggests that further research could be conducted to test the adequacy of the proposed model on a larger sample of open government data use cases.

Davies argues that there are mainly five distinct, yet non-mutually exclusive, processes of open government data use:

- *Data to fact* – individuals may seek out specific facts in a newly open dataset. These facts may support their engagement in civic or bureaucratic processes, or in business planning.
- *Data to data* – sharing derived data (either simply an original dataset in a new format, or data that is augmented, combined with other data, or manipulated in some way. A whole dataset may be shared, an API onto a dataset created, or an interface that makes it easy to download subsets of a large dataset.
- *Data to information* – creating a static representation and interpretation of one or more data sources. Leading to visualizations, blog posts, infographics and written reports.
- *Data to Interface* – creating a means to interactively access and explore one or more datasets. For example, creating a searchable mapping mash-up, or providing a tool to browse a large dataset and crowd source feedback or scrutiny.

- *Data to service* – where open government data plays a ‘behind the scenes’ role in making some online or offline service function. For example, the use of boundary data to route messages reporting potholes to the responsible authority.

Importantly, the first four listed processes correspond to new products or services that are based on open government data, while the fifth listed theme, *data to service*, refers to the application of open government data to achieve or sustain competitive advantage.

Upon setting the main dimensions of analysis, or codes, as 1) *data to fact*, 2) *data to information*, 3) *data to interface*, 4) *data to data*, and 5) *data to service*, we began the process of coding. Codes were collated into potential sub-themes by moving back and forth and comparing within-group similarities, and simultaneously looking for patterns to allow sub-themes to flow from the data (Hsieh & Shannon, 2005; Kondracki, Wellman, & Amundson, 2002). A “miscellaneous” folder was created for each theme so that cases that did not fit directly the emerging sub-theme could be considered later in the process of data analysis. At this stage, plenty of notes were taken in order to assist in “the continuous movement between an empirical world and a model world” (Dubois & Gadde, 2002, p. 554), which is to say, between the data and the coding frame.

Braun & Clarke (2008) contend that “a theme captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set” (p. 82). Thus, the researcher’s judgment is necessary to determine what a theme is. At the end of the coding process, the cases included in “miscellaneous” folder were examined in order to assess their fit to the coding frame.

We proceed to describe the findings of our analysis in detail.

5.3 Findings

Primarily, our results seem to provide support the categorization suggested by Davies (2010) in the case of the commercial use of open government data. No other main-themes were identified in our analysis. This is to say that all five processes (*data to fact*, *data to data*, *data to information*, *data to interface*, and *data to service*), corresponding to the main-themes in our analysis, were found to be recurrent in the data (see Figure 6). This outcome is not negligible because it gives strength to the idea suggested by Davies that, regardless of the stakeholder in question, the manifold uses of open government data can fundamentally be classified by these five processes.

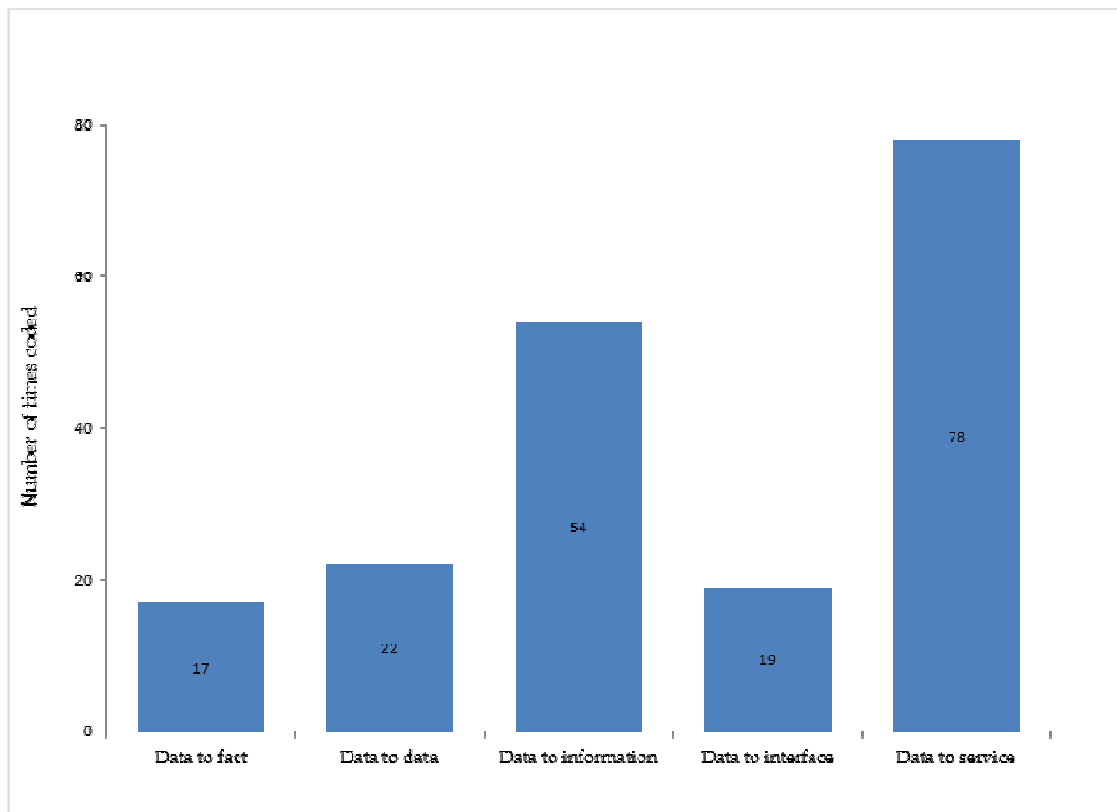


Figure 6. Frequency of open government data types of use

In order to provide a richer account of how open government data can be used to create value in a for-profit setting, we further decompose each process (main-theme) into sub-divisions (sub-themes). Our

findings suggest twelve distinct types of use: *Advocacy, Consultancy, Data refining, Data structuring, Single purpose apps, Interactive apps, Data platforms, Open data portals, Business intelligence, Process optimization, Product/Service improvement, and Research and Development* (see Table 9). The above-listed sub-themes are not mutually exclusive, whereby the same company may be engaged in more than one process simultaneously (Davies, 2010).

	MAIN THEMES	SUB-THEMES	
New products or services built on open government data	<i>Data to fact</i>	Advocacy	
		Consultancy	
	<i>Data to data</i>	Data refining	
		Data structuring	
	<i>Data to information</i>	Single purpose apps	
		Interactive apps	
	<i>Data to Interface</i>	Data platforms	
		Open data portals	
	Open government data for competitive advantage	<i>Data to service</i>	Business intelligence
			Process optimization
Product/Service improvement			
Research and Development			

Table 9. Main-themes and sub-themes that emerged from the data

A more detailed discussion of our findings for each category, including the analysis of the main industries represented in each type of use, is provided in the following sections.

5.3.1 *Data to Fact*

The open government data movement has typically been mainly defined by two streams or ideologies: the “re-use of data” perspective and the “open government” perspective (Jetzek et al., 2013). However, the “discourse is increasingly citing both social and economic reasons for opening data, and the principles of supplying data for open government and re-use are converging” (Jetzek et al., 2013, p. 5).

In the category of *data to fact*, we find two distinct types of use by firms. First, we identify that firms may employ a civic-centered approach to business in order to *advocate* for the advance of open government data. On the user end, this can be translated to, for example, the delivery of services aimed at improving access to data for citizen participation.

Our findings seem to be aligned with the idea of convergence between the two streams of open government data because the use of data for advocacy is often associated with the principles of ‘open government’ – transparency, collaboration, participation – and simultaneously with the profit-oriented tenets of the private sector. Accordingly, our analysis suggests that firms using open government data for advocacy commonly embody a civic component in its mission, which is mainly driven by the goal of promoting the openness of government and the engagement with local communities (e.g. citizen empowerment, universal access to government data).

Additionally, in the theme of *data to fact*, we identify another type of use. In this case, open government data appears to be a potential source of value creation when providing consulting services around open government data. This can mean advising organizations on the legal and technical aspects involved in publishing or using the data, or helping firms leverage their existing resources to create value from open government data.

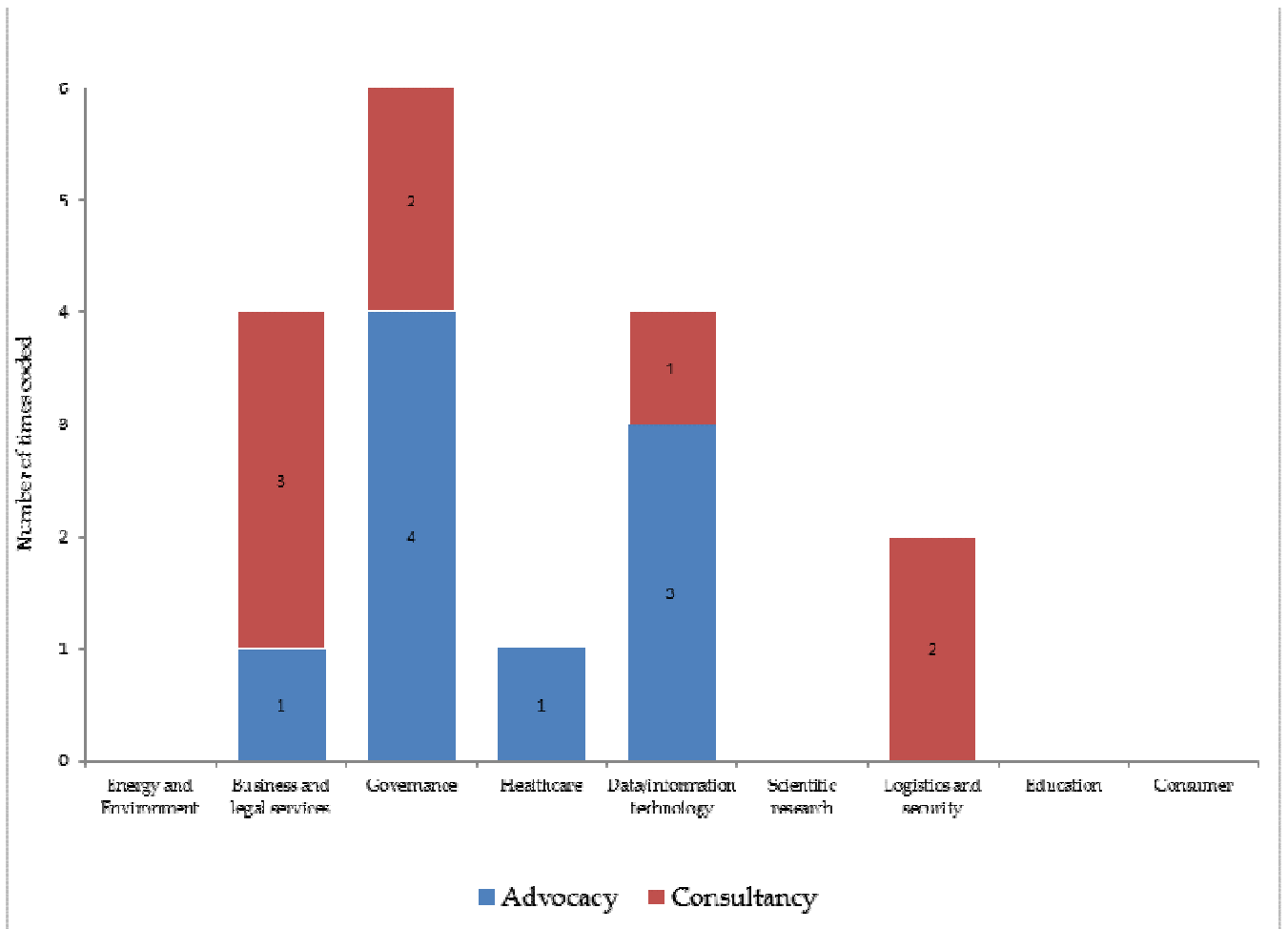


Figure 7. Frequency of “Data to fact” use type by industry

When considering the most prominent industries found in the category of *data to fact* (see **Erro! A origem da referência não foi encontrada.**), it can be observed that firms can be found mainly in the industries of business and legal services, data/IT sector, and governance (i.e. service providers to government). While the aspect of open government data advocacy seems to be carried out mostly by IT companies or by firms operating in the realm of governance, the consultancy work around open government data is more evident in business and legal services. The sector of logistics and security is also noticeable in the category of *data to fact*. In light of the remarkable progress of data-driven services in the transportation industry, the sector of logistics and security is mainly represented by firms that

provide consultancy services around open transportation data, a valuable resource for planning trips more efficiently, to other stakeholders in the ecosystem.

5.3.2 *Data to Data*

The main category of *data to data* broadly comprises the re-packaging of open government data in a format that more is suitable for other users to consume.

The first sub-theme that we introduce in this category is based on the notion of *data refining*. In this case, the data is initially collected (e.g. scrapped, downloaded) from a government source, and is then improved (e.g. cleaned, refined), mainly in terms of its data quality, for more usability. In order to offer a more competitive package of data, services devoted to data refining may also mix data from other government sources. Using open government data in this way is frequently associated with the needs of specific market segments (e.g. business, financial, and legal services) who are interested in higher quality data. The refining of the data can also include, for example, the formatting of the data to conform to industry standards (e.g. open data certificates) for commercial users in a particular industry.

Another recurring theme that emerges in the category of *data to data* relates the concept of *data structuring*. In this case, data is collected or extracted from a large number of government sources, and subsequently published in a repository online. The core notion in this category is that by indexing and structuring large amounts of government data, these services can supply the data in a structure that is more suitable to the needs of users. Commonly, firms pertaining to this category do not only provide the data but also the interface. The latter, essentially, acts as a central database where open government datasets that can be accessed, searched, and downloaded for further analysis. Like in the case of *data*

refining, firms that are engaged in *data structuring* often employ a targeted approach to the market by focusing on commercial users that find it particularly valuable to have access to high-quality data through an efficient access point (e.g. financial institutions, management consulting firms).

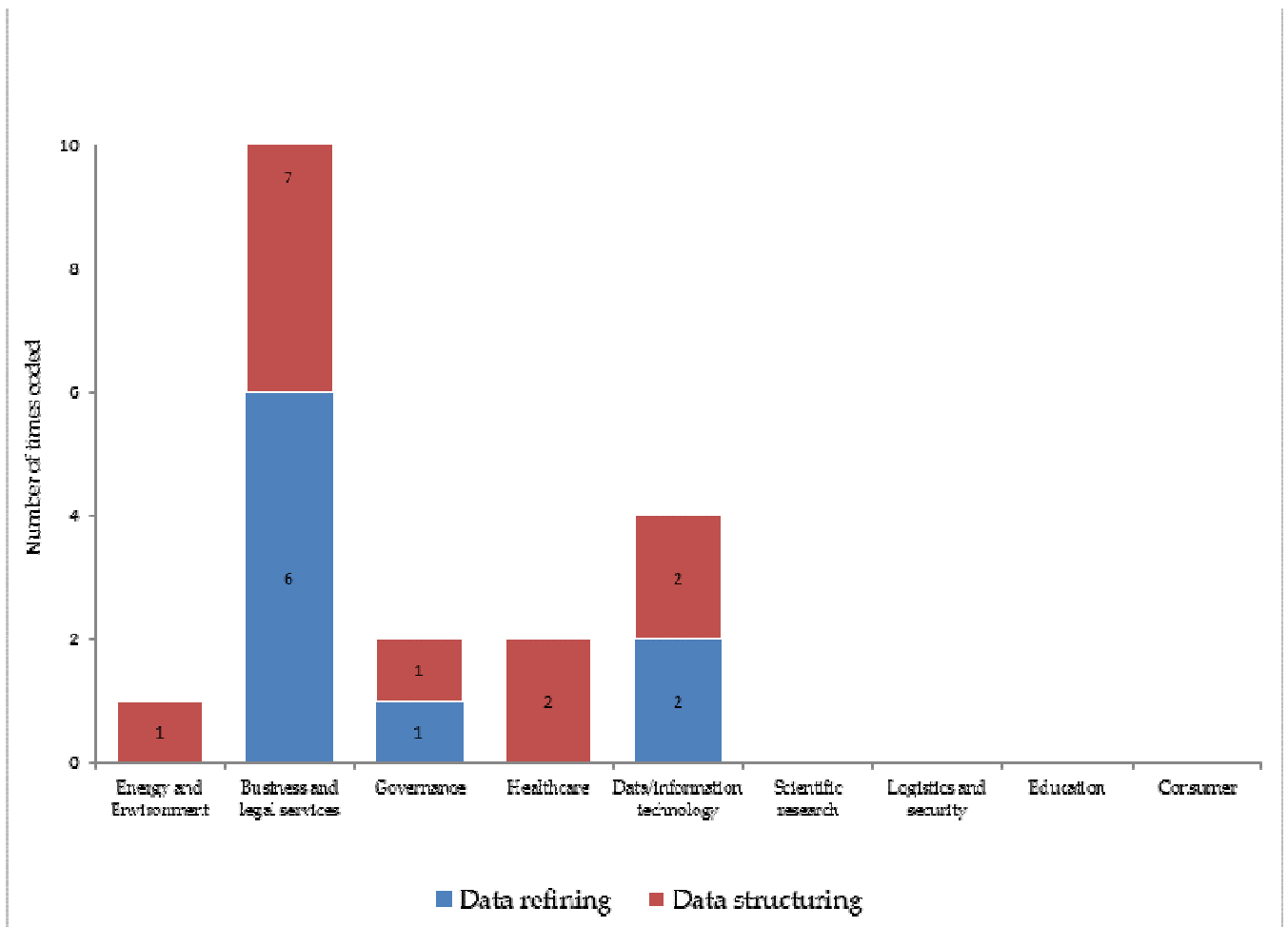


Figure 8. Frequency analysis of "Data to data" use type by industry

Somewhat identically to the main category of *data to fact*, firms in the theme of *data to data* (see Figure 8) can be found mainly in business and legal services, and to a lesser extent in data/IT companies.

Moreover, the sub-categories of *data refining* and *data restructuring* seem to be equitably represented in

each of these two industries, thereby conveniently offering the two main aspects of the *data to data* process (data refining and data restructuring).

5.3.3 *Data to Information*

As raw data is not meaningful to most end users, open government data does not automatically create value (Cranefield et al., 2014). Consequently, the emergence of solutions, such as apps that provide meaningful or actionable open government data to average users, has grown significantly (Longo, 2011). Mainly by converting raw data to actionable information, these solutions can, for example, transform complex raw datasets into intuitive visualizations that have the power to directly provide valuable insights to users. Our findings seem to reinforce the idea that converting *data to information* is, indeed, one of the main ways through which companies create value from open government data.

In the context of apps that use open government data, Janssen and Zuiderwijk (2014) introduce the concept of “single-purpose apps” as apps that typically offer a single function by processing open government data and presenting it visually in an intuitive fashion. Moreover, Janssen and Zuiderwijk propose the concept of “interactive apps” as apps that, alternatively to “single-purpose apps”, integrate more features or functions, and allow the user to combine open data with other sources of data, such as other public datasets, or user generated content.

As part of the larger category of *data to information*, we identify in our analysis the sub-theme of *single-purpose apps*. We find that firms engage in this type of data use to develop apps that offer effective visualizations of large or complex datasets in various domains, like the weather, crimes reported, patents, restaurant inspections, and government contracting. Typically, *single-purpose apps* are based on one type of data (e.g. weather data, patent data), and sometimes from just one government agency.

Consequently, this type of use can be viewed as a well-directed and actionable use of open government data.

In addition to *single-purpose apps*, we also identify the development of *interactive apps* as another possible type of use of open government data. *Interactive apps* are distinct from *single-purpose apps* in that they offer an enhanced level of dialogue with the user, and frequently also aggregate data from multiple sources. Accordingly, firms using data in this way frequently develop apps called “mash-ups” (Cranefield et al., 2014; DiFranzo et al., 2011). These apps integrate open government data with user data, user-generated content, or data from third-party sources. An example of an *interactive app* is a transportation planner app that aggregates information from the user (e.g. location), as well as data from various transportation modalities and companies, to provide the user with practical recommendations regarding the user’s travel options. Another example often encountered is the case of comparison apps, such as school or neighborhood finder apps that incorporate different indicators, which are first processed, and then presented to the user.

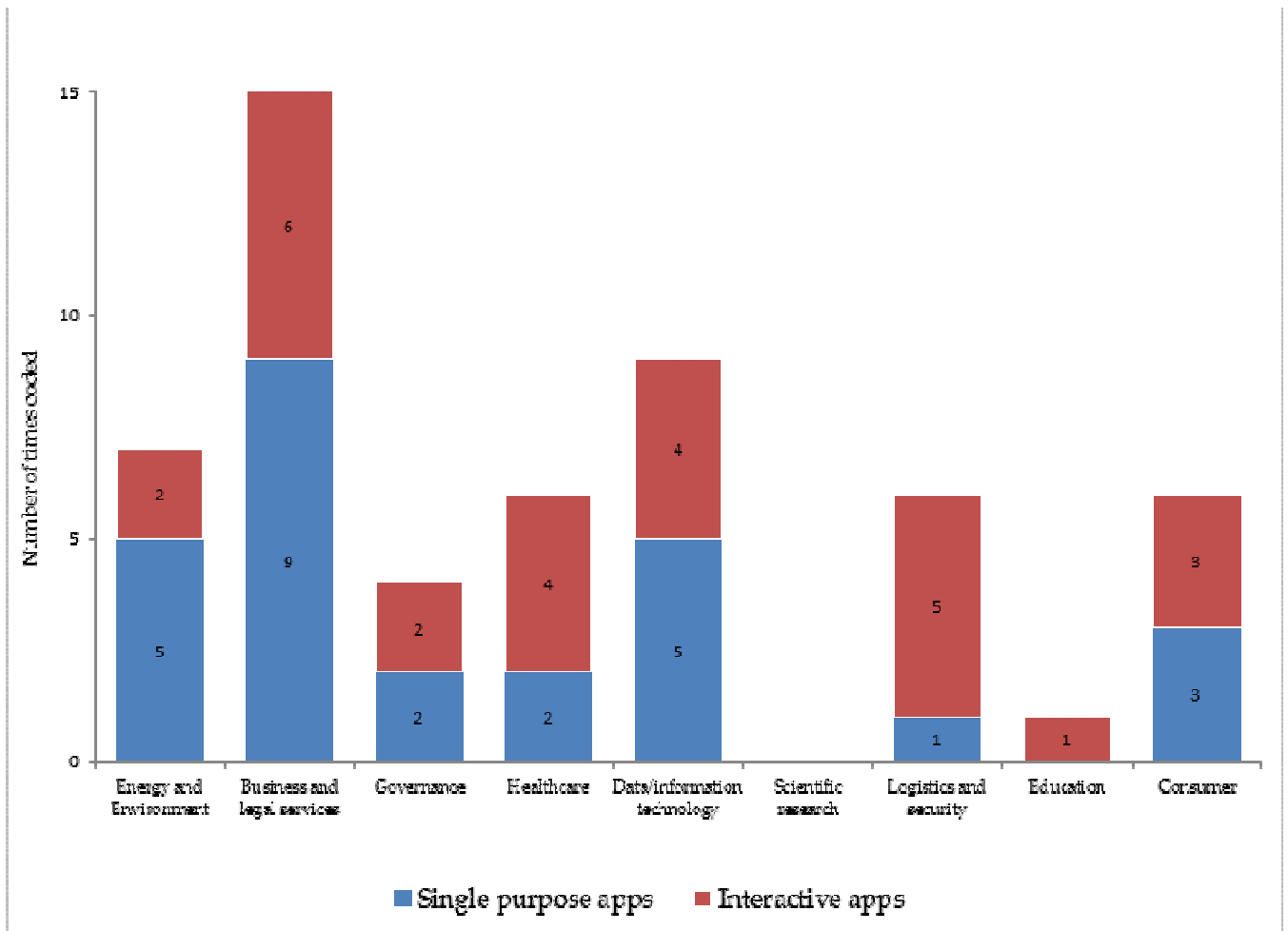


Figure 9. Frequency analysis of "Data to information" use type by industry

When compared to the previous types of use that have been examined (*data to fact*, and *data to data*), the category of *data to information* displays a more equitable distribution across different industries. Still, the sectors of business and legal services, and data/IT remain as most prominent. Contributing to the importance of the former (business and legal services) is, for example, the development of software for legal professionals that relies on open patent data. Alternatively, the latter (data/IT industry) seems to be mainly populated by applications that use open geospatial data to provide visualizations and maps based on open government datasets. The sector of energy and environment is also notable in the category of *data to information*. Mainly, this can be attributed to the development of applications that

provide weather forecast services based on data provided by meteorological agencies from government.

5.3.4 *Data to Interface*

The potential value of open government data is often unlocked by combining public datasets with each other and/or additional data in novel ways, thus creating new sources of value (Maali et al., 2010; Stott, 2014). Consequently, the combination of the increasing availability of data with new techniques to process data ascribe a greater role to stakeholders in the management of their informational resources (Borglund & Engvall, 2014). In our analysis, we identify the process of converting *data to interface* as a possible type of use of open government data in the private sector. Firms that resort to this process are typically committed to creating platforms where users can interactively access and explore (e.g. search, clean, process, and visualize) datasets.

Our analysis suggests that the category of *data to interface* can be broadly decomposed into two sub-themes: the development of 1) *open data portals*, and 2) *data platforms*. While the former is generally more concerned with creating value to the supply-side of open government data (e.g. government agencies), the latter targets largely the demand-side (e.g. citizens, researchers, other businesses).

In light of the tremendous growth of open data portals that serve as data repositories online, it is not uncommon in modern open data portals to find advanced features such as APIs, data visualization, and real-time monitoring. In this context, various firms have endeavored to help public organizations publish their data online. These emerging services, mainly from the sector IT, rely of government data being published openly online to develop *open data portals*. These services can include, for example, the planning, design, or technical development of open data portals using advanced Web technologies.

In addition to *open data portals*, our analysis also suggests a sub-theme related to the development of *data platforms*. In comparison with *open data portals*, *data platforms* typically imply a higher level of interactivity, thereby offering users the ability to effectively explore open government datasets. As part of their enhanced interactivity, *data platforms* offer users the option to overlay other types of data (e.g. proprietary data, social media data, and personal data). By being able to upload, and subsequently combine open government data with other types of data, users are more likely to be able to get new insights from datasets.

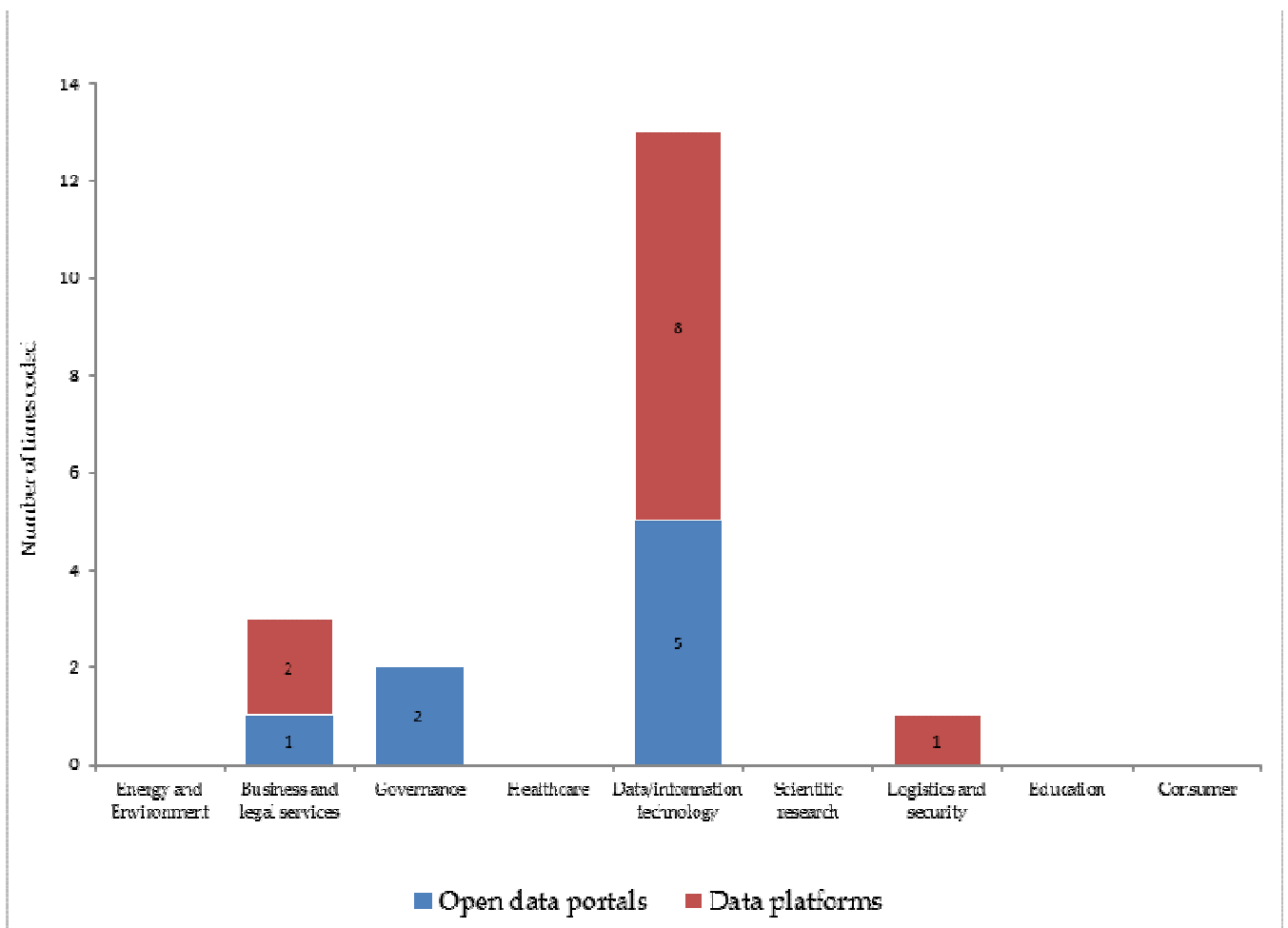


Figure 10. Frequency analysis of “Data to interface” use type by industry

The distribution of firms across industries in the category of *data to interface* (see Figure 10) is strongly represented by the data/IT sector. This is not surprising since the development of open data portals and data platforms is, as expected, mainly undertaken by technological firms.

5.3.5 *Data to Service*

The types open government data use described so far in this chapter (sections 5.3.1 to 5.3.4) relate to the use of open government data for the creation of new products or services. However, our findings also suggest that open government data is exploited by firms with the purpose of creating competitive advantage (Zuiderwijk, Janssen, Poulis, et al., 2015). Using open government for competitive advantage (e.g. to improve existing products or services, or increase organizational efficiency) can be viewed as working “behind the scenes” (Davies, 2010, p. 3) and forms, therefore, the core idea in the process of *data to service*. In our analysis, we identify four distinct sub-types in the *data to service* category. These are: *business intelligence*, *process optimization*, *product/service improvement*, and *research and development*.

As part of their managerial activities and cross-functional decision making, firms constantly strive to find and build competitive advantages (D’Aveni, 1994). In this context, open government data can be used as a valuable resource to assist in business analysis, and in this way improve *business intelligence*. By combining public datasets with proprietary data, firms are therefore able to gain important insights, and understand patterns occurring in the market. By using open government data as supplementary information for various aspects of business management, such as conducting market analysis of a specific demographic area, firm managers can produce relevant output information such as sector benchmarks, or inform risk management models.

Furthermore, open government data seems to potentially add value in another area of business. Solaimani and Bouwman (2012) posit that “the creation of a service is enabled by a ‘strategic-level’ business model, while the implementation or execution of the service is described by ‘operational-level’ business processes” (p. 655). Accordingly, at the operation level, our findings suggest that, when integrated as a supplementary type of data, open government data has the potential to optimize internal business processes. Examples of open government data use for the purpose of internal optimization can be found in domains such as industrial processes, or logistics.

Additionally, open government data appears to be used frequently in the *improvement of existing products or services*. In this case, the use of open government data creates value by endowing existing products or services with advantageous new features. For example, demographic data is used abundantly in real estate in order to enrich the information provided to home buyers in realty online platforms. Another example is the use of open government data as demonstration content or contextual information in a variety of services. Common examples are the use of geographical data to be presented in geo-spatial tools, or economic data that is used as contextual information in investment platforms.

Finally, as a result of our analysis, we posit that open government data can be used by firms in activities of *research and development*. In this case, the data is used to aid in the process of exploration or investigation. This use of open government data can be found mainly in activities of research, or at the base of the innovation life cycle, where data is used in the context of applied research for new product development. An example of this is found in the green energy industry, where public environmental data is used in the process of research and development (R&D) of new solar panel components. Open government data can also be valuable in other aspects of the R&D, such as the use of open patent data to assess the feasibility of developing a new technology.

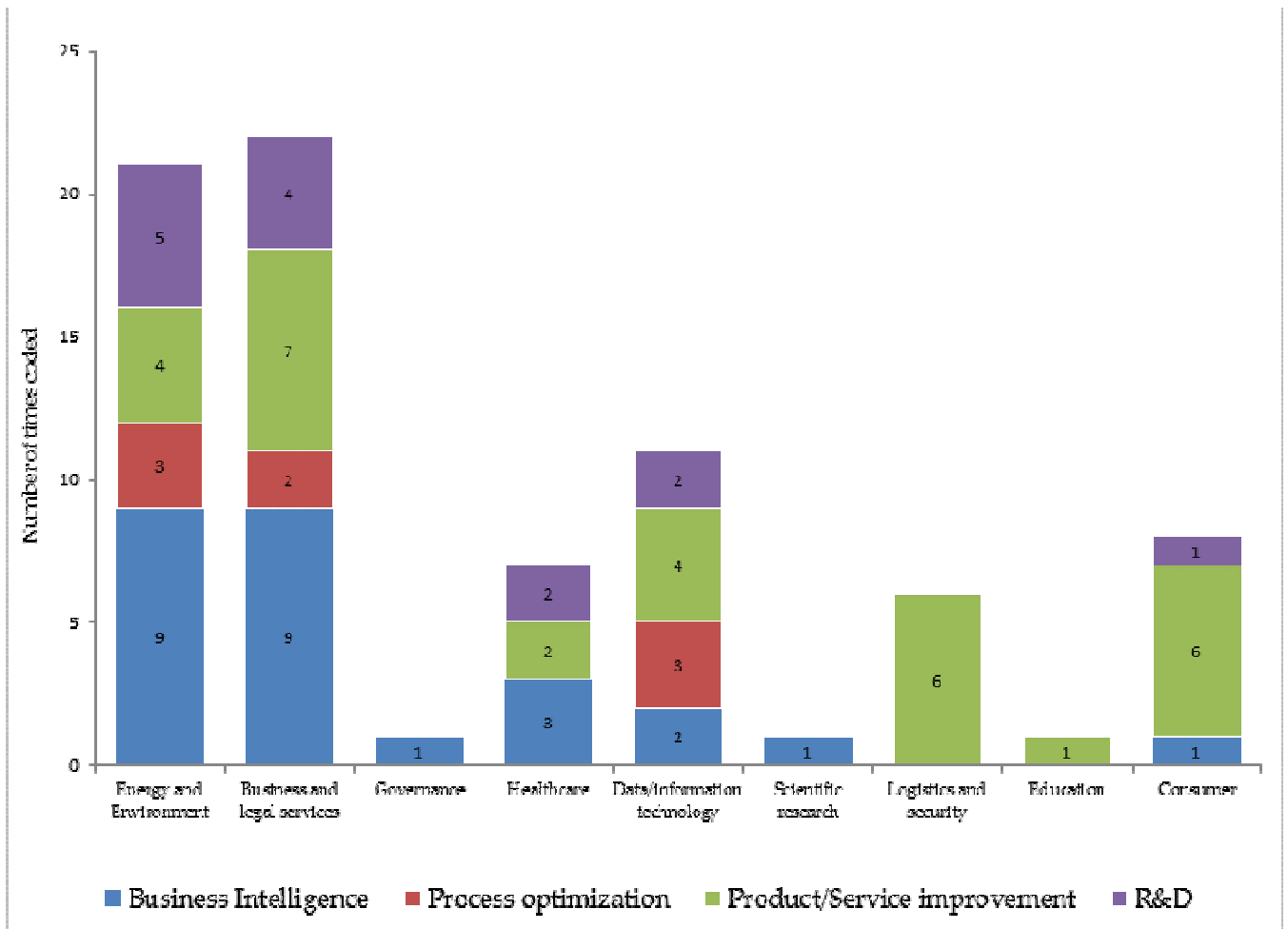


Figure 11. Frequency analysis of "Data to service" use type by industry

When compared to the previous main categories (that refer to the use of open government data to develop new products and services), the distribution of industries in *data to service* (see Figure 11) seems to be most equitable. Still, of most prominence are the business and legal services, as well as the sector of energy and environment. In the group of business and legal services, open government data seems to be used frequently to increase the competitive advantage of financial and consulting services by, for example, determining trends and economic forecasts. Alternatively, in the group of energy and environment, open government data related to, for example, historical climate data, seems to be used

widely for business intelligence and process optimization of firms operating in the fields of renewable energy sustainability and agriculture.

5.4 Conclusions

The increased openness of government data offers the potential to transform the interaction between governments, citizens, and the business sector (Gorham, Jaeger, Sarin, & Choi, 2014). In particular, government officials seem especially excited with the prospect of generating wealth as a result of innovative uses of public data (Gray, 2014). However, despite the emerging public notoriety of this issue, the practice and study of open government data is still in its infancy, in part due to relatively limited empirical data available (Cranefield et al., 2014; Jetzek et al., 2014).

Based on the analysis of the use of open government data by 178 firms in the U.S., we expand the current theoretical perspective that lists five distinct processes through which open government data can be used (*data to fact, data to information, data to interface, data to data, and data to service*). In order to provide a richer account of how open government data can create value commercially, we further decompose this classification into twelve distinct types of use: *advocacy, consultancy, data refining, data structuring, single purpose apps, interactive apps, data platforms, open data portals, business intelligence, process optimization, product/service improvement, and research and development* (Table 9). All twelve subtypes relate to the use of the data for both the creation of new products and services, and for competitive advantage.

Our analysis also confirms that the different types of use are not mutually exclusive, whereby, the same company can be engaged in more than one type of use simultaneously. Through the analysis of the frequency distribution across different industries for each type of use, we find that, overall, open

government data is more prominently used to develop new products and services in the context of business and legal services, and in the data/IT sector. While the former is more pronounced in the *data to data* category, the former takes the central role in the category of *data to interface*. Simultaneously, the use of open government data in the process of *data to fact* seems to be mostly represented by firms providing services to the public sector (governance). Finally, in the most represented theme in our analysis, the main category of *data to service*) the distribution of industries appear to be rather equitable. These findings seem to support the notion that open government data is well established as a value-adding resource for competitive advantage in the private sector.

Overall, the manifold uses of open government data that emerge from our analysis provide a compelling view of the diversity of applications of public data in the private sector. In light of the current need to justify the investments made by governments, our findings seem to strengthen the notion that releasing open government data in an open format can positively impact private-sector innovation.

Our work contributes to academic literature in the topic of open government data in following manner. First, we provide a comprehensive and systematic analysis of the commercial use of open government data by the private sector based on robust empirical data, thus contributing to a domain that has lacked such detail (DiFranzo et al., 2011; Jetzek et al., 2014). When compared to other relevant studies, which mainly focus on products or services exclusively based on open government data (Ahmadi-Zeleti et al., 2014; Ferro & Osella, 2012, 2013), our research expands the current debate on the commercial use of open government data by broadening the scope of analysis of the topic. This is done by including in our analysis firms that create value from open government data in any shape or form.

We also anticipate that our findings, as well as the various real-world examples provided, may be useful to both entrepreneurs and intrapreneurs who are interested in deriving value from public data. Additionally, our research intends to provide government officials with empirical evidence regarding the potential of open government data to generate innovation and competitive advantage.

Still, our study suffers from several limitations. First, the data that was considered for analysis was gathered from various sources online through desk research. Although the reliability offered by this approach is somewhat limited, we attempted to minimize this limitation by having the same researcher conducting the desktop research, and applying the same procedures for all cases (Kaasenbrood et al., 2015). Moreover, the available information in the dataset under analysis was generally short (averaging 1-2 sentences), which offered only a superficial view of how the data is actually used by companies. However, this challenge was considered acceptable in view of the exploratory nature of this study.

6 BARRIERS IN THE COMMERCIAL USE OF OPEN GOVERNMENT DATA

Most prior studies investigating the issue of barriers have been conducted from the perspective of government agencies (supply-side), with little focus on external stakeholders (demand-side). Moreover, it is well-accepted that the various ways in which open government data is used may represent different interests in the same data (Klievink et al., 2014; Tong et al., 2013; Yu & Robinson, 2012). By acknowledging the plethora of uses through which open government can be exploited, it follows that a dataset can be perfectly useful to one purpose but of no value to another.

However, as demonstrated earlier, existing research mainly considers barriers in a generic fashion. This is to say that, despite the importance and challenging nature of this issue, use barriers have not yet been studied in the context of varying uses of open government data. Grounded in the notion that private sector users form a critical stakeholder in the ecosystem, the goal of this chapter is to analyze the barriers encountered by businesses in the commercial use of open government data. To do so, we derive a conceptual framework from relevant literature, and subsequently carry out interviews with firms that use open government data.

Our work contributes to the theoretical knowledge on open government data by developing a framework comprising the barriers in the commercial use of the data. Additionally, we derive useful insights for public officials regarding the implementation of open government data strategies, namely ones that aim to foster innovation.

The structure of this chapter is as follows. First, we introduce and review the issue of use barriers in open government data, paying particular attention to barriers in the context of the commercial use. Next, we describe in detail the methodological approach that was employed in the study. We then discuss our findings by providing an analysis of which barriers play a role, and in what way, in

different scenarios (*i.e.* types of use). We conclude by elaborating on the contribution of our work, and by providing suggestions for future research.

6.1 Background context

Broadly, barriers in the context of open government data can be described as any impediment, challenge, risk, or limitation that is present at any stage of the open data process (Martin et al., 2013; Zuiderwijk et al., 2012). Also, barriers can be mainly distinguished between “implementation barriers” and “barriers to use”, relating to the supply side and the demand side respectively (Martin, 2014). In particular, the presence of use barriers may result in costly challenges, thus limiting the implementation or the use of open government data (Martin et al., 2013; Zuiderwijk et al., 2012).

Despite the newness of the trend of open government data, a considerable number of studies addressing the issue of implementation barriers as encountered by data suppliers has been published (Martin, 2014). These include, for example, barriers related to the lack of organizational support (Sayogo & Pardo, 2012; Van Veenstra & Van Den Broek, 2013), unstructured approach to data publication (Ren & Glissmann, 2012), resistance to change (Janssen et al., 2012), deficiency of applicable legal or policy frameworks (Dawes, 2010; Kulk & Loenen, 2012), and excessive costs (Cranefield et al., 2014), and other organizational hurdles inside government agencies (Davies & Frank, 2013; Denis & Goota, 2014; Helbig, Cresswell, Burke, & Luna-Reyes, 2012).

With respect to “barriers to use”, it must first be noted that the adoption of open government data depends primarily on whether the user has the necessary technologic resources to access and use the data (Gurstein, 2011). Not having access to the required resources, such as a computer device (Bertot et

al., 2010), statistical software (Zuiderwijk, Janssen, Poulis, et al., 2015), or visualization tools (Carvalho & Lafuente, 2015) can be considered, potentially, the first impediment on the user-side.

Upon having access to the data, users can be, for example, partially or fully unable to explore the array of datasets available due to the lack of necessary computational or statistical skills (Martin et al., 2013; Tong et al., 2013) or domain knowledge (Martin et al., 2013). In a commercial setting, this can be manifested as lack of adequate processing power, or insufficient server space to digest very large datasets, which can impede the use of open government data.

Simultaneously, in this data-rush era, “new intermediaries need new skill-sets, including media literacy, but also code literacy and data literacy (Dietrich, 2015). Consequently, finding the necessary talent, either inside or outside the organization, who can endow the company with a statistical skillset that is able to respond to the competitiveness of the commercial market can be also turn out to be a challenge for firms.

The barriers in the use of open government data can “take many forms” (Whitmore, 2014, p. 623). For instance, despite having the required resources to explore open government data, some firms may find it challenging to access public datasets. DiFranzo et al. (2011) note that government datasets are typically published using a variety of approaches, ranging from web services with APIs (Application Programming Interface) to downloadable files.

However, extant literature indicates that users may have difficulty accessing the data. For example, users can find it challenging to find the data online (Cranefield et al., 2014; Kaasenbrood et al., 2015) due to issue such as data fragmentation (Zuiderwijk et al., 2012), or under-developed search features (Vasa & Tamilselvam, 2014). In a similar fashion, it has been suggested that there are challenges

regarding the retrieval or extraction of the data, namely unsuitable file formats or interfaces (Zuiderwijk et al., 2012), and the inexistence of an API (Vasa & Tamilselvam, 2014).

Furthermore, existing literature has brought up use barriers that relate to issues of data usability, such as data quality (Cranefield et al., 2014; Peled, 2013) and metadata (Lemieux, 2014). The discussion of data quality is “as old as data itself”, however “the changing nature and increasing volume of data has exacerbated the problem” (Sadiq, Yeganeh, & Indulska, 2011, p. 153). In the particular realm of open government data, data quality plays a critical role in the ability to efficiently extract value from the data (Ahmadi-Zeleti, 2015; Detlor, Hupfer, Ruhi, & Zhao, 2013). Consequently, the lack of quality data in the form of, for example, low data accuracy (Martin et al., 2013), inconsistency between different releases of the same data (Ahmadi-Zeleti, 2015; Maali et al., 2010), or insufficient data completeness (Janssen et al., 2012) has the potential to hinder the use of open government data. Since the issue of information quality is a critical aspect to any organization (Lee, Strong, Kahn, & Wang, 2002), the quality of government datasets can play a decisive role in outcome of any endeavor of commercial use of open government data (Behkamal, Kahani, Bagheri, & Jeremic, 2014).

Similarly, besides data quality, cases of inadequate or insufficient metadata have also been suggested as a challenge (Eberius, Braunschweig, Thiele, & Lehner, 2012; Whitmore, 2014). Accordingly, it is assumed that metadata can also play an important role in the process of deriving value from open government data (Eberius et al., 2012). In this context, Kerschberg (2011) contends that “metadata matters”. Along these lines, other authors have noted problematic issues broadly related to metadata, such as the lack of a single format to describe datasets (Janssen, 2011b; Martin et al., 2013), incomplete or inexistent metadata (Martin et al., 2013; Vasa & Tamilselvam, 2014), and no interoperability between datasets (Zuiderwijk & Janssen, 2014a).

Additionally, literature points out the aspect of government-user interaction as a possible barrier in the use of open government data (Zuiderwijk et al., 2012). Based on the intrinsic interdependence among different stakeholders in the ecosystem (Albano & Reinhard, 2014), it has been shown that insufficient government-user interaction, such the absence of effective update and feedback mechanisms, can be frustrating for open government data users. Dawes and Helbig (2010) provide a compelling description of this problem:

“There is a notable absence of update and feedback mechanisms in the typical data flow which nearly always goes in only one direction – from the data source to a requester. Many users who obtain data from municipal, county, and state sources often find and correct errors as they use it, but these users are neither expected nor allowed to return data corrections, enhancements, or other improvements to the data sources. Consequently one clear overall benefit of use, data improvement for everyone, is never realized” (p. 56)

Despite this, it is anticipated that the widespread of Web 2.0 technologies can be used to improve government-user interaction (Alexopoulos, Loukis, & Charalabidis, 2014; Charalabidis et al., 2014). This can be achieved by, for example, including modern features in open data portals, such the ability to comment and rate datasets, or allow users to efficiently communicate with the data provider or with other users of the same data.

Finally, due to the fact that the open government data phenomenon is still a very recent field, the legal dimension of open government data – ranging from the enactment of new open data policies and laws, to the licensing of government datasets – is currently under great transformation¹⁸. In addition,

¹⁸ A recent incident in Canada where the postal corporation has sued a business, for what it alleges to be an unauthorized use of its copyrighted postal code data, illustrates this point poignantly. This case is said to have “important ramifications for all businesses in Canada” (King, 2013, para. 1).

because of its wide reach, legal matters pertaining to the use of open government data have the potential to unlock very important challenges regarding the use of the data, and should therefore be given significant attention (Cerrillo-i-Martínez, 2012; Fitzgerald, Hooper, & Cook, 2013).

Tong et al. (2013) define “legislative barriers” as “certain acts of legislation and other regulations reduce the usability of public sector information datasets by consumers” (p. 24). For example, in the context of legislative barriers, literature has brought forward the issue of privacy as a potential problem when making use of open government data (Cranefield et al., 2014; Tong et al., 2013). Privacy-related issues in the context of open government data should matter to both suppliers and users of the data (Simpson, 2011), however “complying with privacy legislation appears to be seen as involving complexity and risk” (Cranefield et al., 2014, p. 9). Most notably, there are concerns about the risk of deanonymization of datasets, which can conceivably be accomplished through advanced data analysis techniques (Ohm, 2010; Sweeney, 1997).

In addition to the question of privacy, issues regarding the licensing of the datasets, which include for example unknown, heterogeneous, and the absence of licenses across datasets, have also been raised (Foulonneau et al., 2014; Janssen, 2011b; Martin, 2014).

Considering the meaningful body of literature on the issue of barriers in open government data, and in light of its relevance for the success of any open government data initiative, some scholars have attempted to conceptually organize barriers (Martin et al., 2013). To do this, scholars have mainly employed an empirical approach (Tong et al., 2013). For example, through the analysis of survey responses measuring the perceptions of U.K. users, Martin (2013) suggests 33 barriers relating to both social and technological factors. Alternatively, Janssen et al. (2012) interview suppliers and users of open government data, and subsequently classify barriers in six main groups: institutional, complexity

of the tasks, use and participation, legislation, quality of information and techniques. In another study, Cranefield et al. (2014) interview various users of open government data to identify twelve barriers, ranging from change-related issues to problems relating to sustainability. Finally, Zuiderwijk et al. (2012) provide a robust qualitative and quantitative research analysis of the barriers by listing 118 impediments (in 10 categories) of socio-technical impediments from the perspective of the user. However, despite the importance of these studies, they have, for the most part, failed to contemplate a clear differentiation between different users of the data. Instead, while focusing on the issue of use barriers, these studies have mainly approached users of open government data rather generically.

Still, a few studies do approach open government data stakeholder groups individually. For example, some researchers have looked into how open government data developers create or reshape services from the release of public datasets, and the conditions required to move from data release to service innovation (Kuk & Davies, 2011; Vasa & Tamilselvam, 2014). Regarding the case of the commercial use of open government data, Kaasenbrood et al. (2015) utilize the business model concept to explore the factors influencing the adoption of open government data by private companies in the Netherlands. Their findings reveal that the content and source of the data needs to be clear, that a usable open data license must be present, and that the continuity of data needs to be ensured through updates. However, in their study, “none of the investigated organisations is heavily dependent on the use of open government data” (p. 86). This presents a significant limitation when studying the issue of use barriers because firms that develop products and services based on open government data represent a central stakeholder of the ecosystem.

It appears that literature has, so far, mainly treated the issue of use barriers generically regarding how the data is used. We argue that this presents a significant limitation, since using open government data

in different ways may result in different interests, and therefore, imply a different perception regarding the barriers of use (Albano & Reinhard, 2014; Klievink et al., 2014; Tong et al., 2013). While introducing the notion of open government data that is “fit for use”, Dawes and Helbig (2010) contend that data should not only be “intrinsically good, but also contextually appropriate for the task, clearly represented, and accessible to users”, whereby it follows that “the same information may be fit for some uses, but completely inappropriate for others that have different temporal, security, granularity, or other requirements” (p. 57).

In light of the extant literature addressing the issue of barriers in the commercial use of open government data, we thus compile published studies according to the main types of barriers identified *resources, data accessibility, data usability, interaction, and legislation* (Table 10).

TYPE OF BARRIER	SOURCE
Resources <ul style="list-style-type: none"> • Technology • Skills 	(Bertot et al., 2010) (Gurstein, 2011) (Martin et al., 2013) (Tong et al., 2013) (Dietrich, 2015) (Zuiderwijk, Janssen, Poulis, et al., 2015)
Data accessibility <ul style="list-style-type: none"> • Data availability • Data retrieval 	(DiFranzo et al., 2011) (Cranefield et al., 2014) (Vasa & Tamilselvam, 2014) (Zuiderwijk et al., 2012) (Kaasenbrood et al., 2015)
Data usability <ul style="list-style-type: none"> • Data quality • Metadata 	(Maali et al., 2010) (Janssen et al., 2012) (Braunschweig et al., 2012) (Peled, 2013)

	(Detlor et al., 2013) (Martin et al., 2013) (Behkamal et al., 2014) (Vasa & Tamilselvam, 2014) (Zuiderwijk & Janssen, 2014a) (Cranefield et al., 2014) (Ahmadi-Zeleti, 2015)
Interaction <ul style="list-style-type: none"> • With government • With other users 	(Dawes & Helbig, 2010) (Zuiderwijk et al., 2012) (Albano & Reinhard, 2014) (Alexopoulos et al., 2014) (Charalabidis et al., 2014)
Legislation <ul style="list-style-type: none"> • Privacy • Licensing 	(Ohm, 2010) (Simpson, 2011) (Janssen, 2011b) (Cerrillo-i-Martínez, 2012) (Fitzgerald et al., 2013) (Tong et al., 2013) (Cranefield et al., 2014) (Martin, 2014) (Foulonneau et al., 2014)

Table 10. Current literature pertaining to each type of barrier

6.2 Research goals and framework

Although the “barriers relating to the demand for and use of open government data are perceived as prevalent” (Martin, 2014, p. 236), most prior studies have been conducted from the perspective of government agencies, with little focus on the user’s perspective (Yang & Kankanhalli, 2013; Zuiderwijk et al., 2012). Accordingly, in this study (chapter 6), we aim to shed light on the issue of use barriers of open government data from the commercial user’s perspective. We endeavor to identify the barriers encountered by for-profit organizations that use open government data, and investigate how the

perception of barriers varies with the scenario of use. In this way, we aim to produce useful insights that can be used for the alignment of government information policies and the interests of commercial innovators of open government data.

Based on our review of literature presented in the previous section (6.1), we proceed to derive a conceptual framework that comprises the various barriers encountered by firms when commercially using open government data (Figure 12).

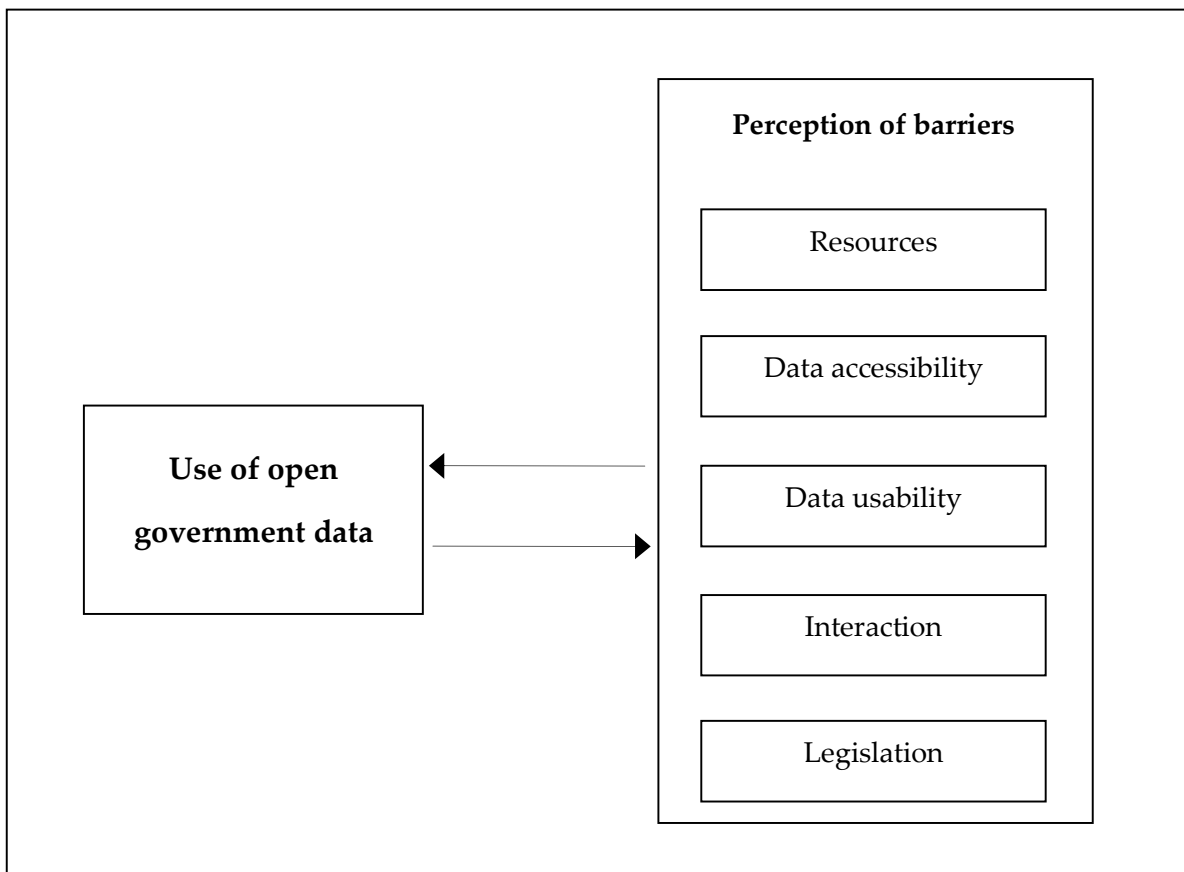


Figure 12. Conceptual Framework

Miles and Huberman (1994) define a conceptual framework as a visual or written product that “explains, either graphically or in narrative form, the main things to be studied—the key factors, concepts, or variables—and the presumed relationships among” (p. 18). Importantly, the most important thing about a conceptual framework is that “it is primarily [...] a tentative theory of the phenomena that you are investigating” (Maxwell, 2013, p. 39).

In the case of the framework that we propose, it should be noted that the dimensions and sub-dimensions of our framework are non-exclusive. As noted by Janssen et al. (2012), the “barriers identified are often interrelated and do not stand alone”. Thus, the same user may experience more than one type of barrier simultaneously. By defining our framework in this way, we ensure that our categories are both “all-inclusive and mutually exclusive” (Gorden, 1992, p. 1).

6.3 Methodology

The choice of a qualitative approach seems appropriate since qualitative research is expected to provide researchers with a rich description of a complex phenomenon (Sofaer, 1999). In this study, we draw on interviews with managers at firms that use open government data (see Appendix I for presentation of firms).

The firms from which managers were interviewed were chosen from the sample used in the previous study (chapter 5) in order to leverage on our existing knowledge of the use of open government data by these firms. Additionally, we made use of the categorization proposed by Davies (2010), which was introduced in the previous chapter, to select firms that display distinct uses of the data to ensure diversity in our study (see Table 11).

	Data to fact	Data to data	Data to information	Data to interface	Data to service
Amazon Web Services	x			x	
Aunt Bertha					x
Billguard			x		
Calcbench		x			
Civic Insight			x	x	
Fastcase		x	x		
Relationship Science					x
TrialX		x			
TuvaLabs	x			x	
Zebu Compliance		x			
Zillow			x		
Beyond Value			x		

Table 11. Distribution of cases according to data use

6.3.1 Data collection

A set of questions, based on our framework, was elaborated for a semi-structured interview format (see Appendix II). Considering that this is still an emerging field, the choice of semi-structured interviews for the method of data collection was chosen to allow for the “elaboration of information that is important to participants but may not have previously been thought of as pertinent by the research team” (Gill, Stewart, Treasure, & Chadwick, 2008, p. 291).

Based on the notion that qualitative semi-structured interviewing seeks to explore meaning and perceptions to gain a better understanding of a particular issue (DiCicco-Bloom & Crabtree, 2006), the goal of our interviews was primarily to get insights into the main barriers in the use of open government data from the user’s perspective. In this context, Burns (1997) argues that, in semi-

structured interviews, a respondent should be approached more as an “informer rather than a respondent” (p. 372).

Interviews were conducted with key decision-makers at the companies (Table 12) through phone or Skype between the months of April and September of 2015. The majority of the interviews (9 out of 12) were carried out with CEOs in order to obtain a more robust, holistic, view of the various challenges, (both technical and non-technical), that can be found in the use of the data. In the remaining cases (C1, C7, and C11), which refer to larger companies, interviews were undertaken with managers who were directly involved in the use of open government data, and were therefore better positioned to provide a rich account of their hands-on experience with public datasets. All interviews were audio recorded, and subsequently transcribed *verbatim*.

Code	Company	Position
C1	Amazon Web Services	Open Data Technical Business Development Manager
C2	Aunt Bertha	CEO
C3	Billguard	CEO
C4	Calcbench	CEO
C5	Civic Insight	CEO
C6	Fastcase	CEO
C7	Relationship Science	Senior Director of Data & Content
C8	TrialX	CEO
C9	TuvaLabs	CEO
C10	Zebu Compliance	CEO
C11	Zillow	Director of Government Relations and Public Affairs
C12	Beyond Value	CEO

Table 12. List of interviewees

6.3.2 *Data Analysis*

Typically, qualitative content analysis research relies on the inductive approach to interpret meaning from the content of text data (Hsieh & Shannon, 2005). However, “generating concepts or variables from theory or previous studies is also very useful for qualitative research” (Zhang & Wildemuth, 2009, p. 2). Unlike other methods, such as grounded theory or thematic analysis, that largely seek to find patterns that emerge from the data, deductive qualitative analysis consists broadly of deriving preliminary codes from theory, which can be tested subsequently by the data (Elo et al., 2014).

In acknowledgement of the number and diversity of prior studies that address the topic of barriers in the use of open government data, we choose to employ the method of deductive qualitative analysis. In this way, the dimensions that compose our framework are viewed as hypotheses (arranged in a categorization form) to be tested by the analysis of the interview data (Marshall & Rossman, 1999).

In conformity with the methods employed in deductive content analysis, following the transcribing of the interviews, the interview data was imported into NVivo version 10, and subsequently coded. All data was reviewed for content, and coded for correspondence with the categories (Elo et al., 2014). Coding was performed by assigning passages of the interview transcripts to the dimensions of our framework. In order to avoid fragmentation, the unit of analysis employed for coding was the sentence. However, surrounding data was included whenever it was relevant to our analysis (Bryman, 2004). ~

In the next section we present the results that emerge from our analysis.

6.4 Findings

In this section, we offer a discussion and provide insight into the barriers perceived by users in the commercial use open government data. Below, we provide a detailed account of each type of barrier encountered by our interviewees. Additionally, in order to go beyond the identification of barriers, we look for patterns across categories (scenarios of use) to assess the extent to which the perception of barriers varies with different uses of the data. This level of analysis is included in the discussion of each barrier.

6.4.1 Resources

Upon analyzing the interviews, we find that the issue of technical resources (e.g. IT infrastructure, hardware and software) was largely not viewed as a barrier in the use of open government data. Interviewees indicated that, in spite of the inevitable challenges with technology when running modern businesses, no special adjustments were made in order to derive value from open government data.

It was noted by some managers that, mostly in the case of smaller-sized firms, the recent emergence of cloud services has empowered businesses with the capacity to easily work with very large amounts of data (i.e. host and process datasets) without the need to make hefty investments (C1, C4). It appears therefore that any issues related to the confidentiality and security of data in cloud services have largely been addressed, which has contributed to firms (as well as their customers) being generally comfortable about using these services, instead of developing their own infrastructure.

"I don't think we could have done this 10 years ago, we just simply couldn't have because we would have had my servers, or have a server farm that we would have to administer. Instead, we use Microsoft. Amazon Web services and Microsoft and these big companies have these programs that they can stick you

in and give you cheap access to computer power. I'm sitting in one of these co-working open spaces, and I work in a Mac laptop. Everything is virtual." (C4)

In relation to the issue of human resources (e.g. skills and competences), the analysis of our interview data shows that, despite the growing offer of academic programs regarding data science, firms still find it challenging to recruit or maintaining talent that is skilled to work with open government data.

Importantly, interviewees commented on two issues that firms are compelled to deal with when allocating their human resources to explore open government data commercially. The first issue relates to the concept of context when working with the data. Since not *"every IT company can simply just download some of these things [...] and then just right use it"*, firms who are interested in creating value from open government data are required to first *"understanding the context, understanding the problem"* (C8). Although the task of computer programming are fundamentally identical to any other field of application, *"what it means and how to operate it [open government datasets], requires special knowledge"* (C8). Consequently, human resources commonly have to go through some kind of skill training in order to understand the *"the complexity of the problem"* (C7) they are working on.

Basically, *"you're taking developers who are pretty good with databases and all of that, and you are trying to marry to logic, that's not straightforward"* (C10). It was suggest that this can be achieved by, for example, employing a team approach, arranging groups formed by software developers who can collaborate with senior managers or other people who are domain experts or are familiar with the field of application. Although the training of human resources can be costly when using open government data, it may simultaneously enable firms to do a lot of the data operations in-house, potentially cutting down the costs in external services.

Besides the aspect of context, the second issue noted by interviewees when considering the barriers related to human resources, was that the work with open government data is typically *“very behind the scenes in a lot of ways”* (C5), mostly when compared with other new and trendy Web technologies. This can be a hurdle for firms that are in the process of recruiting.

“It’s just not that exciting for some people to work on this kind of problem. We are not a social media dashboard, we are not a retail platform, or anything like that...” (C4)

“In order to acquire the right people, mostly when competing against giants like Google or Microsoft, firms using open government data should therefore strive to “find a good mix of someone who has got the right values and understand the space, what work and what doesn’t work in this space, and also has the chops.” (C5)

When analyzing whether the issue of resources is perceived as a barrier across different uses of the data, we find that interviewees denote a consistency in their opinions. This is to say that, in all interviews the issue of technical resources (e.g. IT infrastructure, hardware and software) was largely not viewed as a barrier in the use of open government data. Simultaneously, the ability of firms to recruit skilled human resources is considered somewhat challenging, although not a problematic, by all interviewees.

6.4.2 *Data accessibility*

The issue of how data is accessed, namely how it is found and retrieved by users, has been suggested by literature as a potential barrier in the use of open government data (DiFranzo et al., 2011; Vasa & Tamilselvam, 2014). Our analysis seems to confirm this idea, since most firm managers that we

interviewed implied that companies frequently encounter difficulties in accessing the open government datasets that they need.

In particular, the issue of the *discoverability* of the data seems to be considered to be “*a real issue*” (C1). Even though it is generally acknowledged that the situation has improved greatly in the last few years, some managers expressed their concern regarding the fragmentation of public datasets across different websites. The fragmentation of data seems to be a recurring issue mostly for data at the local or regional level.

“We have had actually more trouble with state level data, so whenever we want to aggregate across different states – for instance, we tried to find information on state level political donation data – and we ended licensing that from a vendor or we would have had to go to each state’s individual website, download, and translate each state’s framework...” (C7)

In addition to the challenge of finding datasets across different cities or regions, two other factors were identified in our analysis as potentially problematic. It was mentioned that finding published raw data can be difficult. Conversely to summary data, which is data that has been aggregated at some level, raw data can also be useful in some cases. For example, raw data may be more suitable when the goal is to learn concepts in statistics or data analytics, instead of just looking for patterns.

Still regarding the issue of finding the data, one interviewee also remarked feeling “*discriminated*” when local governments, in contradiction to the principles of open data, resist providing the data at no cost because the data requested is to be used for for-profit activities.

“... you can call a city directly and ask for the data and a lot of times, again this is all over the board, some places right now have evolved significantly and say “yes, absolutely, we’re an open data city and

you can go to the FTP and download data you need right here, that kind of thing”, so that’s one end of the spectrum. The other end of the spectrum is “what are you going to do with our data? We will only give it to nonprofits and not to private companies. Or we will give it to nonprofits for free, if you are a private company we will charge you anywhere from \$100 to process the data on to the disk to \$40,000 for the same data that in other places would be given away for \$100.” (C12)

Furthermore, our analysis indicates that business managers generally do not consider the aspect of data retrieval to be a barrier in the use of open government data. Typically, there are two main ways through which firms retrieve open government data: file download, or API. The file download option has the advantage of allowing users to save the data files in their machine, thereby avoiding potential technical issues, such as weak network connection. In this way, developers can pull any data independently and quickly play with it, which can often be *“far more convenient”* (C2).

However, automatically or manually downloading files from public server also presents drawbacks. For example, it was noted that *“there are a lot of them [files], and they change all the time. They change the data formats, or move around the web, so it’s difficult to keep track of all the different ones”* (C6). In spite of its efficacy and robustness, downloading files can be resource intensive for firms. In one of the cases, *“it’s a daily thing to stay up and track what’s available”* (C10) to keep up with file changes online. Because of the limitations in the file download option, some managers expressed their preference to retrieve open government data through an API, particularly when working with dynamic data or specific data queries.

“... if you’re working with documents, or a collection of documents or collection of statistics, that would be somewhat static, just give people the files. If it is real-time data or real-time sensor data and things like that, the only effective way to give people the data would be through an API or some kind of feed.” (C1)

In addition to data feeds, APIs can also be of benefit to firms that use open government data that is updated frequently. In this specific case, the option of an API seems more appropriate than downloading the data file.

"We just get the entire data and that is an application, and then later on we are querying the API to maybe get the specific to keep our data up to date." (C7)

Still, the option of retrieving data through an API can present challenges too. For instance, the development of an API is technically not trivial, and potentially too expensive to most agencies. Furthermore, since an API is a software product, it requires regular maintenance. One interviewee (C8) provided important additional information regarding the use of APIs for open government data:

"If it was just a simple API call, and the data was not right or there were some errors in the data, a lot of those errors propagate much faster across systems. In a way it helps that you have to spend some time setting things up and looking at it and identifying issues, because that's when you realize that there are issues in the data itself. It's not the problem of accessing it, but a problem of data integrity, data errors, data incompleteness, and all those issues. In a way, this manual process helps know what the extent of the problem of the data itself is, so that way you have caveats when you build systems that just might be missing, or might be incomplete." (C8)

Despite the advantages and drawbacks of both options for retrieving open government data (i.e. file download vs. API), our analysis suggests that firms are generally technically capable of exploiting either option. This suggests that possibly more research is needed to further understand which data retrieval options provided in government data portals (e.g. download, API) can respond efficiently to the needs of commercial users.

“It is typically what’s available from the provider. We are technically equipped to handle either, but whatever is the technical format that they have, we typically just go with that”. (C7)

With regards to the perception of barriers at the level of *data accessibility* across different uses of the data, our analysis suggests that *discoverability* seems to be perceived as a problem particularly in the cases where open government data is used to develop products and services that aggregate data from multiple public sources. Due to the fact that many sources of data must be managed, firms risk running into issues of scalability or fragmentation of the data. Alternatively, when open government data is used for competitive advantage, firms generally do not seem to see *discoverability* as a barrier because they are more likely to resort to a single, or very few sources of data.

6.4.3 *Data usability*

Data usability, which we introduce as the extent to which open government data can be actually transformed, relates broadly to aspects of data quality and metadata. Being a multidimensional (Otto & Aier, 2013) and application-dependent concept (Sattler, 2009), the issue of data quality is not straightforward. Due to this conceptually fuzziness around the notion of data quality, we purposely allowed a rather open interpretation of the concept of data quality when inquiring firm managers.

From the analysis of our interview data, the quality of the data provided is consistently viewed as paramount to extracting value from open government data. Even though governments have done a very good job of improving datasets in the last few years, interviewees claim that the quality of the data is still, for the most part, deficient. While some agencies do provide high-quality data, plenty of datasets available online seem to suffer from issues of timeliness and accuracy in the data.

In addition, the data usability of available open government data is often limited by incomplete or the absence of metadata, resulting in insufficient descriptive and structural information about the data being provided. Sometimes *“the answer is really buried deep in the definitions of how this stuff all works together”* (C4), however, *“metadata is usually very little or nonexistent for most datasets, and you don’t really know what certain attributes mean or what the units are.”* (C9)

The impact of low-quality datasets in the context of the private sector can be severe. Due to the various challenges resulting from poor data quality in public datasets, firms can be forced to allocate significant resources to tasks such as cleaning and normalizing the data.

“... it’s a hugely expensive and time-consuming process, and that process is very necessary if you are going to end up with high-quality output.” (C6)

In this context, in order to cope with the limitations in data quality, firms commonly develop internal and/or external mechanisms of validation and error-checking. Internal mechanisms correspond to, for example, cross-checking public data with other proprietary data in order to fix the dataset, and therefore minimize the chance of data inaccuracies. Conversely, external mechanisms are broadly based on the idea of including customers in the process of data cleaning or verification by encouraging them to provide feedback about the data back to the firm.

In addition to mechanisms of data validation and error-checking, some firms have also shaped communication with their clients as a way of conveying the lack of reliability entailed in low-quality data, thereby managing the expectations of their customers. This means, for example, establishing that the firm is not responsible for inaccuracies or inconsistencies in government datasets.

“This is not like a chronic problem for us, but it is something that we are very sensitive to. When there is issues with the data, we have to be clear that it is not necessarily our fault.” (C5)

It appears that using a combination of the two mechanisms (data validation and a careful communication strategy regarding the attributes and accountability of the data) can grant firms an advantage in using open government data. For example, interviewee C7 provides a brief account of how these mechanisms can be implemented.

“So we have had to design some collection processes for verification surrounding the fact that the data is out of date, and make sure that we are communicating that in the right way to clients because we don’t really want to have flags all over our system that say “haven’t confirmed this for a while” because that would make us look weird. But we still need to make sure that the expectation is not that that data is immediately up to date.” (C7)

At this point, it is also important to note that, due to the complexity of the issue of data quality, effective solutions are not expected to come from individual endeavors, but instead from collaborative efforts among stakeholders. In this context, the notion of a dynamic ecosystem is central to any efforts towards the improvement of usability of open government data. Interviewee C5 describes this as follows:

“We have to be careful to not blame the city either. It’s one of these things it’s not about passing the blame, but it’s also like you can’t absorb all of it either. [...] We apologize to the customers, to the users, say that we are working on it as soon as possible, we are looking into it, we also go back to the city, we propose ways in which they can solve it, we try and help out the city. (C5)

Finally, firm managers also expressed their desire to see significant progress in the standardization of datasets. Importantly, it was hinted that the lack of standardization of public datasets can ultimately cripple business development, deterring business expansion. Below we transcribe two instances from our interviews where managers describe the importance of data standardization at national and international levels, respectively, for their business.

"... if we can't scale the open data across the country it's not something that we can use." (C11)

"What we want is more countries to adopt the same format that the US has adopted. [...] You can get your hands on it. I can't do the same in the UK, or Australia, or Japan, or Germany, I can't. And if we could that would be tremendous, that would be a real game changer for us." (C4)

When analyzing whether the issue of *data usability* is perceived as a barrier across different uses of the data, we find that that was the case for every interviewee. This suggests that, regardless of the type of use, the quality of the data and metadata are perceived to be critical in the commercial use of open government data.

6.4.4 Interaction

With respect to the interaction with the provider of data, most interviewees confirmed that they already currently interact with governments, albeit mostly at a technical level. Still, interviewees expressed their desire to interact with the government agencies beyond the technical level, turning it into a broad inclusive *"conversation that needs to happen"* (C11). This may include, for example, providing feedback about the data to the government agency, such as report errors in the data, request a new dataset, contribute to the development of open data standards, or showcase data uses.

“... if we had some dialogue or some interaction with some of the engineers or the people behind the data initiatives... What they are thinking? How they get data out? Where they think their role end and where they expect developers should pick up? Maybe at least having that dialogue at some points, maybe once every six months, or once in the year, we can know what they are thinking about, what are the new datasets they are looking at, what are the limitations of those datasets, how we would step in and pick up from there. I think that dialogue might help.” (C8)

Firm managers anticipate that a rich interaction with government agencies should be highly beneficial to both the supply and demand sides of the ecosystem. From the point of view of managers, whenever new policies are made, or new datasets are released, the firms that already hold contact with government agencies will be in an advantageous position by being informed first of these developments.

Despite their desire to interact more closely with government agencies, firms often find difficulties in actually achieving this due to, for instance, the continuous high turnover in public administration.

“Yes, and because of lots of turnover and people changing, you never expect the person you talked to last is still there. There are few of those people in a few agencies but you just pretty much start over with issue every time.” (C10)

In addition, building and maintaining relationships with governments can also be costly due to the resources that are required to manage personal connections. Beyond a certain point, which may differ for every case, firms may run into issues scalability when attempting to develop relations with a large number of public servants at government agencies.

With regards to the interaction with other stakeholders in the open government data ecosystem, such as other users of the same datasets, our analysis did not yield straightforward results. On the one hand, the idea of interacting with other users of the data is generally viewed as positive development for the open government data community. In this way, the notion of the open government data ecosystem as a collaborative environment, with plenty of interaction between different users, is reinforced.

“I think that there is a lot of collaboration, surprisingly. I think that there is a general interest to make this industry viable and successful, so if we start being cutthroat right now, I think it’s a little early. All of us are still too young, there’s still a lot of space, there is still a lot of work to do, so if we start pointing their guns at each other we are just going to destroy ourselves and won’t actually be able to grow in the space.” (C5)

“... being able to collaborate with like-minded partners in order to open up data we see as serving a common good, which is certainly helping us to achieve our mission.” (C11)

However, on the other hand, the interaction with other users may not be, at this point, a priority mainly because the market space is still rather embryonic.

Our analysis suggests that firms that use open government data as a value-adding resource for competitive advantage (unlike those firms that use open government data as a core resource), do not see the development of relationships with government and/or other users of the data as a priority. We hypothesize that the difference between these cases may be related to the degree to which open government data is central in the firm’s business model. Essentially, while firms whose business model is more centered on open government data tend to value these relationships, other firms, to whom open government data is a supplementary resource, do not attribute the same importance to interacting with other users.

6.4.5 Legislation

Finally, in what relates to the legal dimension of open government data, we inquire firm managers about whether, and how, the aspects of licensing and privacy have been perceived as a barrier in the use of the data. Although, in general, the issue of legislation was not viewed by managers as a problem *per se*, it was added, nonetheless, that there is much room for improvement.

For example, managers noted that in order to fully comply with the licensing terms and conditions of open datasets, managers are often required to spend more time examining the license content, reaching out to government agencies to understand what the license implies, and procuring legal services to efficiently deal with license standardization issues.

Some interviewees provided interesting accounts of situations in which the licensing of datasets hindered the use of open government data, such as instances when local/municipal or state/regional governments claiming the ownership of the data. The issue of ownership of the datasets has been considered in literature by Conradie and Choenni (2012) as being contrary to the principles of open governance, potentially resulting in a significant barrier of use.

It was also mentioned in the interviews that the lack of license standardization across different cities or states can also be considered a problem since it can lead to coverage gaps in service development, which sometimes, is reflected on the user end.

With regards to the issue of privacy when using open government data, firm managers do not, for the most part, consider it to be impeditive. Still, an interesting account of how privacy can become, somewhat unexpectedly, an issue was offered:

“... one of the biggest challenges of privacy is the death of practical obscurity. What I mean by this is that in the past that has been a lot of private information in civic data but that has not been a problem because very few people had access to. For example, there are public Form 990 filings from the IRS for nonprofits. And they’ve always been public but no one has really asked for them, no one has really needed to use them. Well, now there’s a lot of people that use them. Publicresource.org has collected a lot of these public forms 990s, and they contain a lot of things like Social Security numbers. You’re not supposed to put this kind of private information into your filing, but in the past it has not matter because it was practically obscure, no one ever would come across it. Well, now people do, and so that privacy becomes an extra problem.” (C6)

Despite this account, the issue privacy was not viewed as a problem when using open government data mainly because users assume that datasets were thoroughly vetted by government agencies, and therefore, should be, in principle, liberated from privacy issues.

Nonetheless, due to the central importance of the issue of privacy in modern data-centered businesses, some firms make redundant efforts to ensure that all data is *“completely stripped of PII [Personally identifiable information], this is 100% anonymous data”* (C3), and therefore virtually free from privacy related risks. It was also disclosed by managers that some firms make it explicit in the user agreement that, in general, whatever use (or misuse) of the data, which may constitute an infringement of privacy legislation in place, is of the user’s responsibility.

In light of the data collected in the interviews, we note that respondents provided an agreeing perspective regarding the legal aspects of open government data under consideration. Despite the complexity and sensitivity of this matter, firm managers that use open government data in different ways seem to agree that the legal aspects pertaining to open government data use, such as issues of licensing and privacy, do not constitute real barriers to the use of the data.

6.5 Conclusions

Open government data is a “very large topic” (Ferro & Osella, 2013, p. 1). Due to the diversity of applications of public datasets, the study of open government data can take various perspectives (Gray, 2014). Although it remains largely unknown, the issue of barriers in the use of open government data is currently a rather important aspect to both scholars and practitioners (Whitmore, 2014).

The goal of this study (chapter 6) was to uncover the main barriers encountered by firms in the use of open government data and its role in different scenarios of use. To do so, we derive a conceptual framework from relevant literature, and subsequently conduct interviews with 12 firms that use open government data. The selection of cases was chosen to ensure diversity by including cases that present distinct uses of open government data.

By employing a deductive content analysis to our interview data, we found the presence of barriers mainly in the dimensions of *data accessibility*, *data usability*, and *interaction*. We also shed light on the impact of these barriers at the business level. For example, our analysis suggests that datasets with low data quality can be significantly costly for firms, and that the lack of standardization of datasets has the potential to hinder business development. With regards to interaction, our findings suggest that firms using open government data are predisposed to expand the interaction with government agencies from a technical discussion into a rich dialogue.

With respect to the perception of the same barriers across different types of use, our analysis does not show a significant variation. In other words, we find that interviewees are reasonably aligned in their view of what constitutes a barrier, mostly in the dimensions of *resources*, *data usability*, and *legislation*. In the case of *data accessibility*, our analysis suggests that *discoverability* is more likely to become a barrier in the cases where open government data is used to develop products and services than when data is

used for competitive advantage. Similarly, at the level of *interaction*, firms whose business model is more centered on open government data appear to be more open to developing relationships with government agencies and other users of the data.

Our work aims to extend existing understanding of the commercial use of open government data as a whole by examining in detail the main barriers encountered by managers. We also provide insights to both government officials and firm managers on which aspects need closer attention when respectively supplying and using open government data commercially.

In this study, we ensure construct validity by including various past studies and theoretical contributions to the issue of barriers in the use of open government data. Nonetheless, our study suffers from several limitations. Firstly, the open government data ecosystem is a very new arena in constant transformation, which may, to some degree, challenge the consistency of our result. We attempted to minimize this limitation by having the same researcher developing the semi-structured interview, and applying the same procedures for all cases, which increased reliability (M. Woods, 2011). Furthermore, it is unclear to what extent the findings are indeed general to open government data as a whole, or whether they should be contextualized by different characteristics, such as business sectors, locations, etc. Thus, we attempted to increase external validity by selecting cases that present different uses, or processes, when using government data.

Considering the scarce understanding of this emergent field, various opportunities for further research arise from our findings. For example, our approach could be reproduced in specific contexts within the realm of the commercial use of open government data (e.g. sectors, locations, etc.) in order to understand how findings may vary. In a similar fashion, researchers could also examine the role of socio-cultural factors in the commercial use of the data by studying and comparing findings from

different regions or countries. Furthermore, it would be useful to extend this study by investigating whether the barriers that were identified in this study are being counterbalanced with actionable policies and guideline. Finally, it would be rather valuable to comparatively explore which barriers are predominant, and what is their impact, in other groups of open government data users (e.g. citizens, non-profits, civic software developers). Studying other groups of open government data users can lend essential information for the enactment of policies and legislation that account more equitably for differences between stakeholders.

In the next and concluding chapter, we review and consolidate the findings offered in each chapter of this dissertation, as well the limitations of each study. We conclude by examining the main theoretical contributions of our work, and by providing suggestions on what mechanisms can be explored by both governments and firms in order to create a more effective commercial use of open government data.

7 CONCLUSIONS

The implications of opening government data are manifold, including for example more public scrutiny, higher public service efficiency, and increased innovation. In light of this, in the last few years, governments around the world have rapidly set forward concrete measures to promote both the supply and the demand of open government data.

Despite the effort by governments, the different stakeholders in the open government data ecosystem, as well as their respective perspectives, are not well understood (Gonzalez-Zapata & Heeks, 2014). In this context, the “response from external stakeholders to leverage government data for innovative activities has been lacking” (Yang & Kankanhalli, 2013, p. 644). Partially, this is attributed to the presence of an extensive range of socio-technical barriers on the user side that appears to be preventing open government data strategies from achieving greater impact (Martin, 2014).

In the specific case of the commercial use of open government data, a critical part of the ecosystem, the presence of barriers has the potential to significantly hinder private-sector innovation (Kaasenbrood et al., 2015). However, very little is still known about how open government data is used by firms, and which factors obstruct the use commercial use of open government data. Accordingly, the main the goal of this dissertation is to understand how private-sector organizations create value from open government data, and to uncover the main barriers encountered by commercial users in the use of the data.

Our research is composed by two main studies. In the first study (chapter 5), we focus on the issue of value creation by examining the different ways in which companies use open government data. It is widely accepted that, broadly, firms use open government data to 1) develop new products and services, and 2) achieve or sustain competitive advantage. However, the scarce extant academic

literature on this topic is mainly of an exploratory nature, leaving the issue of value creation in the commercial use of open government data largely unknown.

In order to tackle the first research question of this dissertation (*how do for-profit organizations use open government data to create value?*) we conduct a thematic analysis of 178 firms in U.S. that use open government data. Grounded in a systematic research approach and solid empirical data, the findings in this study offer a glimpse into how firms create value from open government data. We identify twelve distinct types of commercial use: *advocacy, consultancy, data refining, data structuring, single purpose apps, interactive apps, data platforms, open data portals, business intelligence, process optimization, product/service improvement, and research and development.*

Our analysis also confirms that the different types of use are not mutually exclusive, whereby the same company is engaged in more than one type of use simultaneously. Through the analysis of the frequency distribution across different industries for each type of use, we find that, overall, open government data is more prominently used to develop new products and services in the context of business and legal services, as well as in the data/IT sector. The most equitable distribution of industries is found in the most represented category in our analysis (*data to service*), where open government data is used as a value-adding resource for competitive advantage. Our findings provide primarily an important descriptive, yet focused, insight into the practice of commercial use of open government data.

Still, our study suffers from several limitations. First, the data that was considered for analysis was gathered from various sources online through desk research. Although the reliability offered by this approach is somewhat limited, we attempted to minimize this limitation by having the same researcher conducting the desktop research, and applying the same procedures for all cases (Kaasenbrood et al.,

2015). Moreover, the available information in the dataset under analysis was generally short (averaging 1-2 sentences), which offered only a superficial view of how the data is actually used by companies. However, this challenge was considered acceptable in view of the exploratory nature of this study.

In the second study of this dissertation (chapter 6), we focus on the aspect of barriers encountered by commercial users in the use of open government data. Considering that the presence of use barriers on the user's side contributes to a decreased use of open government, the issue of barriers becomes a critical one. Moreover, by acknowledging the plethora of uses through which open government can be exploited, it follows that a dataset can be perfectly useful to one purpose but of no value to another.

Accordingly, by taking advantage of approaching the issue of barriers for a specific stakeholder group, instead of a generic approach, we undertake the second research question of this dissertation (*what are the barriers encountered by for-profit organizations in the commercial use of open government data, and how does the perception of barriers vary with the scenario of use?*). This approach allows us to focus on what works, and conversely, what can be improved for the specific stakeholder group under consideration (i.e. commercial users). Grounded in the data produced by interviews with managers of 12 firms that use open government data, we found the presence of barriers to be prominent in the dimensions of *data accessibility*, *data usability*, and *interaction*. We also discuss the impact of these barriers at the business level. The issues of low data-quality and lack of standardization of data appeared to be the most problematic for firms.

Through the analysis of the data collected in our interviews, we find that firm managers are reasonably aligned in their view of what constitutes a barrier in different scenarios of use. The implications of these findings are not trivial. For example, based on the notion that the barriers in the commercial use of open government data are rather common to different scenarios of use in the private sector,

governments can potentially devise strategies and initiatives that aim to ameliorate the main challenges currently encountered by the private sector as a whole.

Also, based on our findings, our understanding is that barriers should also not be seen atomistically. Our research hints that an overlap often occurs between different barriers. An example of this can be found in the discussion around the issue *data quality*, which despite its central role in the public discourse around the use of open government data, is still fuzzy. While sometimes data quality refers to a technical characteristic of the dataset, other times it is broadly related to how relevant the data is to the user. In this context, we posit that the study of barriers in the use of open government data requires the contribution of every stakeholder in the ecosystem in order to arrive at more solid definitions.

In this study, we ensure construct validity by including various past studies and theoretical contributions to the issue of barriers in the use of open government data. Nonetheless, our study suffers from several limitations. Firstly, the open government data ecosystem is a very new arena in constant transformation, which may, to some degree, challenge the consistency of our result. We attempted to minimize this limitation by having the same researcher developing the semi-structured interview, and applying the same procedures for all cases, which increased reliability (M. Woods, 2011). Furthermore, it is unclear to what extent the findings are indeed general to open government data as a whole, or whether they should be contextualized by different characteristics, such as business sectors, locations, etc. Thus, we attempted to increase external validity by selecting cases that present different uses, or processes, when using government data.

In this concluding chapter (chapter 7), the results and insights produced by each study are summarized and discussed in terms of their implications. We begin by describing some general conclusions that can be extracted from this dissertation. Next, we examine the main contributions of this dissertation to

knowledge theory, and subsequently offer a set of proposals and practical insights to both firm managers and public officials regarding the creation of value from the commercial use of open government data.

With respect to more general conclusions that can be extracted from this dissertation, first and foremost, we note that the practice of open government data has greatly improved in recent years. This perception is clear throughout the analysis conducted. In both empirical studies of this dissertation, it becomes rather apparent that local and federal governments have strived to increase the relevance and quality of published open government data. In chapter 5, this is illustrated by the multitude of applications that open government enjoys in the private sector. This notion is further reinforced by our interviewees in chapter 6.

Our analysis also hints that the use open government data is coupled with a sense wanting to create public value, or as it noted by one of our interviewees: "It's about the common good." In both studies presented in this dissertation, our findings suggest that deriving value from open government data is often associated with the notion of civic innovation. Importantly, this finding potentially lends an optimistic perspective to the harmonization of the social and the economic value created from the use of open government data.

Nonetheless, there is also a broad acknowledgement that the open government data ecosystem is a very recent endeavor, at times almost showing dynamics of an embryonic stage. In this context, basic concepts such as the quality of open government datasets, or the obligations of government in the supply of data, are still rather unsettled.

7.1 Contribution

Our research attempts to advance the existing understanding of the commercial use of open government data by shedding some light on the “lack of connection between a reality of data provision and an aspiration of developmental results” (Gonzalez-Zapata & Heeks, 2014, p. 441). Furthermore, our work intends to be useful to government officials in the implementation of open government data strategies, namely ones that aim to foster innovation.

In this section, we review the main contributions of this dissertation to knowledge theory, and subsequently offer a set of proposals and practical insights regarding how value can be derived from datasets in a for-profit setting.

7.1.1 *Contribution to knowledge theory*

This dissertation contributes to knowledge theory in multiple ways:

- First, it provides an updated review of the literature in the topic of open government data, with a particular focus on its commercial use. Despite the growth of the theoretical body on the issue of open government data, this is, to our knowledge, a novel contribution to literature.
- It expands existing theory on the value creation of open government data by focusing on the context of commercial use. Supported by the unpacking of the various processes listed in the literature, it also provides a richer account of how firms use open government data to create value.
- It provides a comprehensive and systematic analysis of the use of open government data by the private sector based on robust empirical data, which has pointed by the research community as a gap in extant literature (Jetzek et al., 2014; Wood, 2011).

- It extends the concept of intermediaries beyond the notion of “re-use” (Vickery, 2011) to offer a holistic view of the for-profit organizations that derive value from the exploitation of open government data.
- It proposes a theoretical framework of barriers perceived by users when using open government data commercially. No framework of use barriers for the specific case of commercial users had yet been proposed in literature. Grounded in the notion that different stakeholders may have distinct interests in using the data, we anticipate that this framework may be useful to future research on the commercial use of open government data or when studying the issue of use barriers.

7.1.2 *Suggestions for firm managers*

Based on our findings, we provide firm managers who are interested in adjusting their firm’s business model towards a more effective use of public data with the follow suggestions:

- The adoption of the most adequate method for retrieving open government data is not a straightforward issue. Although this discussion can sometimes be dictated by an *either/or* approach (API vs. file download), we note that configurations that include the two modes is not only possible, but sometimes even preferable. Thus, we encourage firm managers to evaluate the pros and cons of each data retrieval method that is made available to them, as to find the most appropriate one for their needs.
- In order to effectively cope with issues of low data quality, firms should develop mechanisms of data validation. The implementation of mechanisms for data validation has the potential to ameliorate the impact of datasets with inaccurate, incomplete, or outdated data. These

mechanisms can be internal (e.g. cross-checking public data with other datasets), or external to the firm's boundaries (e.g. collaborating with customers in the process of data cleaning or verification).

- Due to the complexity of the legal issues pertaining to the use of open government data, we suggest that firms communicate, in a clear language, and in as much as detail as possible, the legal conditions under which they are providing the data. When possible, products or services should specify in the user agreement which party is liable in the case of poor quality datasets, or misuse of the data.
- Open government data is not the most exciting topic in mainstream society, particularly when compared to other new technologies, such as social media, or machine learning. This can turn out to be a complicated hurdle for firms that engage in the process of recruiting talent. Accordingly, we suggest that managers give priority to applicants that strike a balance between feeling identified with the values and ideas of openness (e.g. open government and open data), and having the technical skills required to work with complex data problems.
- Finally, our analysis suggests that the element of context is rather substantial when extracting meaningful information from open government data. Based on the notion that “without the meaningful use of data, data engineering is just a bunch of cool tricks” (Bizer, Boncz, Brodie, & Erling, 2011, p. 56), firms should make concrete efforts to integrate their knowledge of the specific business context with the ability to statistically handle the data. This can be achieved, for example, by 1) fostering collaboration between data scientists and senior managers that hold a good understanding of the application domain; or 2) conducting a thorough and individual assessment of costs and benefits of endeavoring in open government data use.

7.1.3 *Suggestions for public officials*

Besides the importance of increasing the ability of firms to effectively use open government data, it is evident that fine tuning business models to make a more effective use of the data will only partly improve the overall situation. In fact, if “government data is to generate its potential value for society, then government information policies and practices need to be better aligned with the needs of secondary users” (Dawes & Helbig, 2010, p. 52).

Our findings regarding the perception of use barriers in different scenarios of use are useful to government officials who are engaged in the implementation of open government data strategies, namely ones that aim to foster innovation. Below, we offer government officials some insights and suggestions that can be used as relevant information when designing or implementing open government data policies:

- Faced with the need to increase the engagement of external stakeholders, governments should actively look for “champions”, both inside and outside the public sector.

Internally, the ability to identify public servants who are invested in the ideas of open government data is critical to the success of open government data policies. In addition to the obvious benefits of having a motivated employee in their structure, the identification an internal champion can be a very important cog to counteract some of the difficulties of government in supplying the data. These challenges can include, for example, the lack of awareness of the public officials about the benefits and risks of opening up data, high employee turnover in administrative positions, and deficient operationalization of internal policies at the middle-management level.

Simultaneously, by finding champions outside the public sector, more specifically in the private sector, governments can gain from, for example, having these individuals spread the word of the rising opportunity of open government data, or, through their personal networks, connect to other potential users in the private sector.

- There is a pressing desire by commercial users to interact with government agencies beyond the technical level. We anticipate that in the next few years, governments of local, state/regional, or federal/national levels will be expected to improve and expand the interaction with their users to a level of dialogue. Fundamentally, this new interaction implies a two-way conversation in which the role of the participants in the ecosystem is strengthened (e.g. increased communication between users and providers), and broadened (e.g. by including other stakeholders in the conversation).

We propose that government officials consider the development of modern data portals that feature a modernized social layer, where users can maintain a rich conversation with public servants, and among themselves. Besides the technical feedback regarding the data itself, this interaction should include other features that can be of value to users, such as the opportunity to request a new dataset, collaboratively developing open data standards, and showcasing data uses.

- According to the firms managers interviewed in our study, the issue of licensing can be confusing to the user, potentially leading to a decrease in use of open government data. Examples of this lack of clarity can be found when the license of the dataset is not explicit, or when the ownership of the data is ambiguous.

Thus, we posit that progress should be attempted regarding the explicitness of the license agreement, as well as in the standardization of licensees. While the issue of clarity regarding the ownership and use terms of public data is imperative with respect to the principle of universal use of open data, the standardization of licenses can be a decisive factor for the business strategy of firms.

- Finally, regarding the actual data itself, two aspects are worth noting.

First, we find that, despite its central role in the public discourse, the issue of *data quality* is still a fuzzy concept. The ambiguity of the concept of data quality in the discussion of government data can be a problem when setting up strategies that aim to increase the quality of the data available for users. In this context, we suggest that government agencies collaborate with the users of their data to, in each particular case, understand more clearly what is perceived by them as high-quality data.

Furthermore, some attention should also be given to the format of the data. Typically, governments release *aggregated data*, which in contrast to *raw data*, refers to data about which assumptions and grouping operations have already been made. Aggregated data is most suitable for the large majority of users. However, in some specific scenarios (e.g. low-level data exploration, or to learn statistics), *raw data*, such as data resulting from surveys, may be the most appropriate data format. Thus, we encourage government officials to ponder the release of more raw datasets in order to cater for the needs of users that pursue more detailed data for their specific purposes.

7.2 Future work

Considering the scarce understanding of this emergent field, various opportunities for further research arise from our research. For example, our approach could be reproduced in specific contexts within the realm of the commercial use of open government data (e.g. sectors, locations, etc.) in order to understand how findings may vary. Future research could also examine the role of socio-cultural factors in the commercial use of the data by studying and comparing findings from different regions or countries. Furthermore, it would be useful to extend this study by investigating whether the barriers that were identified in this study are being counterbalanced with actionable policies and guideline. Finally, it would be rather valuable to comparatively explore which barriers are predominant, and what is their impact, in other groups of open government data users (e.g. citizens, non-profits, civic software developers). Studying other groups of open government data users can lend essential information for the enactment of policies and legislation that account more equitably for differences between stakeholders.

7.3 Concluding notes

There is a significant opportunity for businesses to tap into the value of open government data in a wide range of application domains. This opportunity is likely to grow further as more data becomes available every day.

We anticipate that the knowledge output of our research, both in theory and in practice, contributes to a more dynamic and inclusive open government data ecosystem. In particular, we find reassurance in the view that “to be successful, open government data programs should not be approached as ‘technical exercises’ but require active engagement and collaborations between different stakeholders”

(p. 3). Accordingly, we celebrate the involvement not only of the private sector, but also of every other stakeholder in the open government data ecosystem.

In an era where the notion of *data* is paramount to most areas of our lives, and value is co-created through networked of interdependent relationships, we posit that the development of an ecosystem that is simultaneously just and competitive, can only be attained through the commitment of all stakeholders. More importantly, considering that the issue of open government data is not an apolitical one, we urge that the ideas or actions that find support in the findings of this dissertation are in the direction of the common good.

8 BIBLIOGRAPHY

- Access Info Europe, & Open Knowledge Foundation. (2010). *Beyond Access: Open Government Data and the Right to Reuse*. Retrieved from https://www.access-info.org/wp-content/uploads/Beyond_Access_10_Aug_2010_consultation.pdf
- Agrawal, D., Zhang, C., & Kettinger, W. J. (2014). The Openness Challenge: Why Some Cities Take It On and Others Don't. In *20th Americas Conference on Information Systems*.
- Ahmadi-Zeleti, F. (2015). Government must not take quality of open data for granted. *Egov News*. Retrieved from <http://egov.news/2015/11/26/government-must-not-take-quality-of-open-data-for-granted/>
- Ahmadi-Zeleti, F., Ojo, A., & Curry, E. (2014). Business Models for the Open Data Industry: Characterization and Analysis of Emerging Models. In *15th Annual International Conference on Digital Government Research* (pp. 215–226).
- Aitamurto, T., Sirkkunen, E., & Lehtonen, P. (2011). *Trends in data journalism*. next Media programme of TIVIT publication. Retrieved from http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2011/D3.2.1.2.B_Hyperlocal_Trends_In Data_Journalism.pdf
- Albano, C. S., & Reinhard, N. (2014). Open Government Data: Facilitating and Motivating Factors for Coping with Potential Barriers in the Brazilian Context. In M. Janssen, H. J. Scholl, M. A. Wimmer, & F. Bannister (Eds.), *13th IFIP WG 8.5 International Conference, EGOV 2014, Dublin, Ireland, September 1-3, 2014. Proceedings* (pp. 181–193).
- Alexopoulos, C., Loukis, E., & Charalabidis, Y. (2014). A Platform for Closing the Open Data Feedback Loop based on Web 2.0 functionality. *Journal of eDemocracy and Open Government*, 6(1), 62–68.
- Alhojailan, M. I., & Ibrahim, M. (2012). Thematic Analysis: A Critical Review of Its Process and Evaluation. In *WEI International European Academic Conference Proceedings* (pp. 8–21).
- Aporta. (2012). *Characterization Study of the Infomediary Sector (english version)*. Gobierno de España. Retrieved from http://datos.gob.es/sites/default/files/121001_RED_007_Final_Report_2012_Edition_vF_en.pdf

- Appelgren, E., & Nygren, G. (2014). Data Journalism in Sweden. *Digital Journalism*, 2(3), 394–405.
- Artigas, F., & Ae Chun, S. (2013). Visual analytics for open government data. In *Proceedings of the 14th Annual International Conference on Digital Government Research* (pp. 298–299). New York, New York, USA.
- Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., ... Wouters, P. (2004). Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal*, 3(November), 135–152.
- Baiocchi, G., & Ganuza, E. (2014). Participatory Budgeting as if Emancipation Mattered. *Politics & Society*, 42(1), 29–50.
- Bates, J. (2012). “This is what modern deregulation looks like”: co-optation and contestation in the shaping of the UK’s Open Government Data Initiative. *The Journal of Community Informatics*, 8(2). Retrieved from <http://www.ci-journal.net/index.php/ciej/article/view/845/916>
- Bates, J. (2014). The strategic importance of information policy for the contemporary neoliberal state: The case of Open Government Data in the United Kingdom. *Government Information Quarterly*, 31(3), 388–395.
- Bauer, F., & Kaltenböck, M. (2011). *Linked Open Data: The Essentials*. Semantic Web Company: edition mono/monochrom. Retrieved from <https://www.semantic-web.at/LOD-TheEssentials.pdf>
- Beghin, N. (2014). *Avaliando os websites de transparência orçamentária nacionais e sub- nacionais e medindo impactos de dados abertos sobre direitos humanos no Brasil*. Inesc — Instituto de Estudos Socioeconômicos. Retrieved from <http://www.inesc.org.br/biblioteca/publicacoes/textos/pesquisa-transparencia-orcamentaria-nos-websites-nacionais-e-sub-nacionais>
- Behkamal, B., Kahani, M., Bagheri, E., & Jeremic, Z. (2014). A metrics-driven approach for quality assessment of linked open data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 64–79.
- Belianska, V. (2014). Societal Benefits and Economic value of Open Data. *COMSODE*. Retrieved from <https://www.comsode.eu/index.php/2014/10/societal-benefits-and-economic-value-of-open-data/>

- Benjamin, S., Bhuvanewari, R., & Rajan, P. (2007). *Bhoomi: E-Governance, Or , An Anti-Politics Machine Necessary to Globalize Bangalore? A CASUM-m*. Retrieved from <https://casumm.files.wordpress.com/2008/09/bhoomi-e-governance.pdf>
- Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3), 264–271.
- Bertot, J., Lee, S. M., Hwang, T., & Choi, D. (2012). Open innovation in the public sector of leading countries. *Management Decision*, 50(1), 147–162.
- Bertot, J., McDermott, P., & Smith, T. (2012). Measurement of open government: Metrics and process. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2491–2499.
- Bitner, S. (2014). How to Use Free External Data Sources for Your BI Advantage. *Sand Hill*. Retrieved from <http://sandhill.com/article/how-to-use-free-external-data-sources-for-your-bi-advantage/>
- Bizer, C., Boncz, P., Brodie, M. L., & Erling, O. (2011). The meaningful use of big data: four perspectives - four challenges. *ACM SIGMOD Record*, 40(4), 56–60.
- Borglund, E., & Engvall, T. (2014). Open data?: Data, information, document or record? *Records Management Journal*, 24(2), 163–180.
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. *Transforming Qualitative Information Thematic Analysis and Code Development*. Sage Publications.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679.
- Brantley, W. (2015). The Data Briefing: The Open Government Data Revolution Is Just Getting Started. *Digitalgov.gov*. Retrieved from <http://www.digitalgov.gov/2015/12/02/the-data-briefing-the-open-government-data-revolution-is-just-getting-started/>
- Braun, V., & Clarke, V. (2008). Using thematic analysis in psychology. *Qualitative Research in Psychology*, (October 2012), 37–41.
- Braunschweig, K., Eberius, J., Thiele, M., & Lehner, W. (2012). *The State of Open Data: Limits of Current*

Open Data Platforms Categories and Subject Descriptors. Technische Universität Dresden Faculty of Computer Science, Database Technology Group. Retrieved from https://www.db.inf.tu-dresden.de/opendatasurvey/www2012_short.pdf

Brito, J. (2010). All Your Data Are Belong to Us: Liberating Government Data. In D. Lathrop & L. Ruma (Eds.), *Open Government: Collaboration, Transparency, and Participation in Practice*. O'Reilly Media.

Broad, E., Smith, F., Duhaney, D., & Carolan, L. (2015). *Open data in government; how to bring about change*. Open Data Institute. Retrieved from <https://theodi.org/open-data-in-government-how-to-bring-about-change>

Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of "big data"? *McKinsey Quarterly*. Retrieved from http://www.mckinsey.com/insights/strategy/are_you_ready_for_the_era_of_big_data

Bryman, A. (2004). *Social Research Methods*. *Social Research* (2nd ed.). SAGE Publications, Ltd.

Burns, R. (1997). *Introduction to Research Methods* (3rd ed.). Melbourne: Addison Wesley Longman.

Buteau, S., Larquemin, A., & Mukhopadhyay, J. P. (2015). Open Data and Applied Socio-economic Research in India: An Overview. In *Open Data Research Symposium*. Ottawa, Canada.

Carolan, L. (2015). Open all hours: How open data is helping some of the world's most impoverished countries. *Civil Service World*. Retrieved from <http://www.civilserviceworld.com/articles/analysis/open-all-hours-how-open-data-helping-some-worlds-most-impooverished-countries>

Carvalho, A. I., & Lafuente, R. (2015). *Interface Design Strategies for Data Portals*. European Public Sector Information Platform. Retrieved from http://www.epsplatform.eu/sites/default/files/Interface_Design_Strategies_for_Data_Portals_v1.pdf

Cashman, T. (2014). The Economic Impact of Open Data. *Socrata Blog*. Retrieved from <http://www.socrata.com/economic-impact-open-data/>

Cerrillo-i-Martínez, A. (2012). The Reuse of Public Sector Information in Europe and Its Impact on Transparency. *European Law Journal*, 18(6), 770–792.

- Chan, C. M. L. (2013). From Open Data to Open Innovation Strategies: Creating E-Services Using Open Government Data. In *2013 46th Hawaii International Conference on System Sciences* (pp. 1890–1899).
- Charalabidis, Y., Alexopoulos, C., & Loukis, E. (2016). A Taxonomy of Open Government Data Research Areas and Topics. *Journal of Organizational Computing and Electronic Commerce*.
- Charalabidis, Y., Loukis, E., & Alexopoulos, C. (2014). Evaluating Second Generation Open Government Data Infrastructures Using Value Models. In *47th Hawaii International Conference on System Science (HICSS 2014)* (pp. 2114–2126).
- Chattapadhyay, S. (2014). *Opening Government Data Through Mediation: Exploring the Roles, Practices and Strategies of Data Intermediary Organisations in India*. Open Data Research Network. Retrieved from http://www.opendataresearch.org/sites/default/files/publications/sumandro_oddc_project_report_0.pdf
- Chignard, S. (2012). *L'Open data: Comprendre l'ouverture des données publiques*. FYP éditions.
- Chun, S. A., Shulman, S., Sandoval, R., & Hovy, E. (2010). Government 2.0: Making connections between citizens, data and government. *Information Polity*, 15(1–2), 1–9.
- CISCO. (2016). *The Zettabyte Era: Trends and Analysis*. Retrieved from <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.pdf>
- Clarke, A., & Margetts, H. (2014). Governments and Citizens Getting to Know Each Other? Open, Closed, and Big Data in Public Management Reform. *Policy & Internet*, 6(4), 393–417.
- Conradie, P., & Choenni, S. (2012). Exploring process barriers to release public sector information in local government. In *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance - ICEGOV '12* (pp. 5–13).
- Cornford, J., Wilson, R., Baines, S., & Richardson, R. (2013). Local governance in the new information ecology: the challenge of building interpretative communities. *Public Money & Management*, 33(3), 201–208.
- Courmont, A. (2012). How to Govern Open Data? The politics of open data portals. In *IPSA Conference*.

Madrid, Spain.

- Crane, H. (2013). Public sector staff know open data matters but fail to get government plan. *The Guardian*. Retrieved from <http://www.theguardian.com/public-leaders-network/2013/feb/18/public-sector-staff-open-data>
- Cranefield, J., Robertson, O., & Oliver, G. (2014). Value in the mash: Exploring the benefits, barriers, and enablers of open data apps. In *ECIS 2014 Proceedings*.
- Craveiro, G. S., Machado, J. A. S., & Machado, J. S. (2015). Open Government Data Initiatives and its Impacts on Citizen Empowerment: the case of "Caring For My Neighborhood." In *Open Data Research Symposium*. Ottawa, Canada.
- Cucciniello, M., Nasi, G., & Valotti, G. (2012). Assessing Transparency in Government: Rhetoric, Reality and Desire. In *2012 45th Hawaii International Conference on System Sciences* (pp. 2451–2461).
- Cuillier, D., & Piotrowski, S. J. (2009). Internet information-seeking and its relation to support for access to government records. *Government Information Quarterly*, 26(3), 441–449.
- D'Aveni, R. A. (1994). *Hypercompetition: Managing the Dynamics of Strategic Maneuvering*. (R. Gunther, Ed.) New York (Vol. 21). Free Press.
- Dander, V. (2013). How to gain knowledge when data are shared? Open Government Data from a media pedagogical perspective. In *MIT8: media in transition international conference*. Boston. Retrieved from http://web.mit.edu/comm-forum/mit8/papers/Dander_MIT8_draft.pdf
- David, P. A., & Foray, D. (2003). Economic Fundamentals of the Knowledge Society. *Policy Futures in Education*, 1(1), 20.
- Davies, T. (2010). *Open data , democracy and public sector (Master's thesis)*. University of Oxford.
- Davies, T. (2011). *Open Data: Infrastructures and ecosystems* (Open Data Research).
- Davies, T. (2013). Transparency and open data. In S. Burall & T. Hughes (Eds.), *Mapping the new terrain: public dialogue on science and technology*.
- Davies, T. (2014a). Open Data in Developing Countries: Emerging Insights from Phase I. *The World*

Wide Web Foundation, 1–38. Retrieved from

<http://www.opendataresearch.org/sites/default/files/publications/Phase 1 - Synthesis - Full Report-print.pdf>

Davies, T. (2014b). Open Data Policies and Practice: An International Comparison. In *European Consortium for Political Research Panel P356*. Retrieved from

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2492520

Davies, T. (2014c). Unpacking Open Data: Power, Politics and the Influence of Infrastructures. Berkman Center, Harvard. Retrieved from <https://www.youtube.com/watch?v=VgEIt7MnxfA>

Davies, T., & Bawa, Z. (2012). The Promises and Perils of Open Government Data (OGD). *The Journal of Community Informatics*, 8(2). Retrieved from <http://communityjournal.net/index.php/ciej/article/view/929/926>

Davies, T., & Frank, M. (2013). “There’s no such thing as raw data’. Exploring the socio-technical life of a government dataset. In *5th Annual ACM Web Science Conference* (pp. 75–78). New York, New York, USA.

Dawes, S. S. (2010). Information Policy Meta-Principles: Stewardship and Usefulness. In R. H. Sprague Jr. (Ed.), *43rd Hawaii International Conference on System Sciences HICSS43* (pp. 1–10). IEEE.

Dawes, S. S. (2012). *A Realistic Look at Open Data*. W3C. Retrieved from https://www.w3.org/2012/06/pmod/pmod2012_submission_38.pdf

Dawes, S. S., & Helbig, N. (2010). Information strategies for open government: Challenges and prospects for deriving public value from government transparency. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6228 LNCS, pp. 50–60).

Dawes, S. S., Pardo, T. A., & Cresswell, A. M. (2004). Designing electronic government information access programs: a holistic approach. *Government Information Quarterly*, 21(1), 3–23.

Denis, J., & Goota, S. (2014). Exploration, Extraction and Rawification. The Shaping of Transparency in the Back Rooms of Open Data. In *Open Questions on Closed Systems - Neil Postman Graduate Conference*. New York, New York, USA.

- DeSantis, L., & Ugarriza, D. N. (2000). The concept of theme as used in qualitative nursing research. *Western Journal of Nursing Research*, 22(3), 351–372.
- Detlor, B., Hupfer, M. E., Ruhi, U., & Zhao, L. (2013). Information quality and community municipal portal use. *Government Information Quarterly*, 30(1), 23–32.
- DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, 40(4), 314–321.
- Dietrich, D. (2015). *The Role of Civic Tech Communities in PSI Reuse and Open Data Policies*. ePSI Platform. Retrieved from https://www.europeandataportal.eu/sites/default/files/library/201505_the_role_of_civic_tech_communities_in_psi_reuse_and_open_data_policies.pdf
- DiFranzo, D., Graves, A., Erickson, J. S., Ding, L., Michaelis, J., Lebo, T., ... others. (2011). The web is my back-end: Creating mashups with linked open government data. *Linking Government Data*, 205–219.
- Downey, E. (2012). *Public Service, Governance and Web 2.0 Technologies: Future Trends in Social Media*. IGI Global.
- Dubois, A., & Gadde, L. (2002). Systematic combining : an abductive approach to case research. *Journal of Business Research*, 55, 553–560.
- Dumpawar, S. (2015). *Open Government Data Intermediaries: Mediating Data to Drive Changes in the Built Environment (Master's thesis)*. Massachusetts Institute of Technology.
- Dwivedi, Y. K., Weerakkody, V., Janssen, M., Millard, J., Hidders, J., Sniijders, D., ... Slade, E. L. (2015). Driving innovation using big open linked data (BOLD) panel. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. http://doi.org/10.1007/978-3-319-25013-7_1
- Eaves, D. (2010). Case Study: How Open data saved Canada \$3.2 Billion. Retrieved from <https://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>
- Eberius, J., Braunschweig, K., Thiele, M., & Lehner, W. (2012). Identifying And Weighting Integration

Hypotheses On Open Data Platforms Categories and Subject Descriptors. *Wod*, 22–29.

- Edwards, S. (2014). Participatory Budgeting: Increasing the Power of the Taxpayer over Public Expenditures. *Chicago Policy Review*. Retrieved from <http://chicagopolicyreview.org/2014/11/14/participatory-budgeting-increasing-the-power-of-the-taxpayer-over-public-expenditures/>
- Elo, S., Kaariainen, M., Kanste, O., Polkki, T., Utriainen, K., & Kyngas, H. (2014). Qualitative Content Analysis: A Focus on Trustworthiness. *SAGE Open*, 4(1), 1–10.
- European Commission. (2013). *Galileo, Europe's GPS, opens up business opportunities and makes life easier for citizens*. European Commission. Retrieved from http://europa.eu/rapid/press-release_MEMO-13-718_en.htm
- Evans, A. M., & Campos, A. (2008). Open Government Initiatives: Challenges of Citizen Participation. *Journal of Policy Analysis and Management*, 32(1), 173–203.
- Fagan, C. (2016). Why open data can stop corruption. *Transparency International*. Retrieved from <http://blog.transparency.org/2016/03/24/why-open-data-can-stop-corruption/>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27–34.
- Ferro, E., & Osella, M. (2012). *Business Models for PSI Re-Use: A Multidimensional Framework*. “Using Open Data: Policy Modeling, Citizen Empowerment, Data Journalism” Workshop. Retrieved from https://www.w3.org/2012/06/pmod/pmod2012_submission_16.pdf
- Ferro, E., & Osella, M. (2013). *Eight Business Model Archetypes for PSI Re-Use*. *Open Data on the Web*. “Open Data on the Web” Workshop. Retrieved from https://www.w3.org/2013/04/odw/odw13_submission_27.pdf
- Fitzgerald, A., Hooper, N., & Cook, J. S. (2013). Implementing open licensing in government open data initiatives: a review of Australian government practice. In *Proceedings of the 9th International Symposium on Open Collaboration - WikiSym '13* (p. Article No. 39).
- Floridi, L. (2005). Is Semantic Information Meaningful Data? *Philosophy and Phenomenological Research*,

70(2), 351–370.

- Foulonneau, M., Martin, S., & Turki, S. (2014). How Open Data Are Turned into Services? In M. Snene & M. Leonard (Eds.), *Exploring services science: 5th International Conference, IESS 2014. Proceedings* (pp. 31–39). Geneva, Switzerland.
- Fung, A., Russon Gilman, H., & Shkabatur, J. (2013). Six Models for the Internet + Politics. *International Studies Review*, 15(1), 30–47.
- Gibbons, S. (2011). Crime nudge. *LSE Spatial Economics Research Centre*. Retrieved from <http://spatial-economics.blogspot.com/2011/07/crime-nudge.html>
- Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journal*, 204(6), 291–295.
- Gitelman, L., & Jackson, V. (2013). *Raw Data Is an Oxymoron*. (L. Gitelman, Ed.). MIT Press.
- Goldstein, B., & Dyson, L. (2013). *Beyond Transparency: Open Data and the Future of Civic Innovation*. Code for America Press.
- Gonzalez-Zapata, F., & Heeks, R. (2014). The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly*, pp. 441–452.
- González-Zapata, F., & Heeks, R. (2015). Understanding Multiple Roles of Intermediaries in Open Government Data. In *13th International Conference on Social Implications of Computers in Developing Countries*. Negombo, Sri Lanka.
- Gorden, R. (1992). *Basic Interviewing Skills*. *Basic Interviewing Skills* (1st ed.). Waveland Pr.
- Gorham, U., Jaeger, P. T., Sarin, L. C., & Choi, H. (2014). Big data, open government and e-government: Issues, policies and recommendations. *Information Polity*, 19(1–2), 5–16.
- Gould, E., & Gomez, R. (2010). Community Engagement & Infomediaries : challenges facing libraries , telecentres and cybercafés in developing countries. In *iConference*. Urbana Champaign, Illinois, USA.

- Gould, E., & Gomez, R. (2010). New challenges for libraries in the information age: a comparative study of ICT in public libraries in 25 countries. *Information Development*, 26(2), 166–176.
- Granickas, K. (2013). *Understanding the impact of releasing and re-using open government data*. European Public Sector Information Platform. Retrieved from http://www.epsiplatform.eu/sites/default/files/2013-08-Open_Data_Impact.pdf
- Granickas, K. (2015). *Building Community around Open Government Data*. European Public Sector Information Platform. Retrieved from https://www.europeandataportal.eu/sites/default/files/library/201409_building_community_around_open_government_data.pdf
- Graves, A., & Hendler, J. (2013). Visualization tools for open government data. In *Proceedings of the 14th Annual International Conference on Digital Government Research* (pp. 136–145).
- Gray, J. (2014). Towards a Genealogy of Open Data. In *General Conference of the European Consortium for Political Research*. Glasgow.
- Gruen, N., Houghton, J., & Tooth, R. (2014). *Open for Business: How Open Data Can Help Achieve the G20 Growth Target*. Lateral Economics. Retrieved from https://www.omidyar.com/sites/default/files/file_archive/insights/ON_Report_061114_FNL.pdf
- Gurin, J. (2014). *Open Data Now: The Secret to Hot Startups, Smart Investing, Savvy Marketing, and Fast Innovation* (1st ed.). McGraw-Hill.
- Gurstein, M. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*, 16(2), 1–10.
- Hagel, J., & Rayport, J. F. (1997). The coming battle for customer information. *Harvard Business Review*, 75(1), 53–55, 58, 60–61 passim.
- Hall, W., Shadbolt, N., Tiropanis, T., O'Hara, K., & Davies, T. (2012). *Open data and charities*. Nominet Trust. Retrieved from [https://www.nominettrust.org.uk/sites/default/files/Open Data and Charities.pdf](https://www.nominettrust.org.uk/sites/default/files/Open%20Data%20and%20Charities.pdf)
- Halonen, A. (2012). *Being Open About Data: Analysis of the UK open data policies and applicability of open*

data. Finnish Institute in London. Retrieved from

[http://finnishinstitute.cdn.coucouapp.com/media/W1siZiIsIjIwMTMvMDEvMzEvMTNfMDNfMTNfNjM1X2JlaW5nXzIwb3Blbl8yMGFib3V0XzIwZGF0YS5wZGYiXV0/being open about data.pdf?sha=32c50d21c5cabb01](http://finnishinstitute.cdn.coucouapp.com/media/W1siZiIsIjIwMTMvMDEvMzEvMTNfMDNfMTNfNjM1X2JlaW5nXzIwb3Blbl8yMGFib3V0XzIwZGF0YS5wZGYiXV0/being%20open%20about%20data.pdf?sha=32c50d21c5cabb01)

Hargrave, S. (2016). Open data feeds the poor 20,000 tonnes of food per year in Washington DC. *The Guardian*. Retrieved from <http://www.theguardian.com/media-network/2016/may/10/washington-dc-food-bank-open-data-deliver-20000-tons-food-per-year>

Harrison, T., Pardo, T., & Cook, M. (2012). Creating Open Government Ecosystems: A Research and Development Agenda. *Future Internet*, 4, 900–928.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). *Big Data for Big Business? A Taxonomy of Data-Driven Business Models Used by Start-Up Firms*. Cambridge Service Alliance. Retrieved from <http://cambridgeservicealliance.eng.cam.ac.uk/news/March2014Paper>

Heimstädt, M., Saunderson, F., & Heath, T. (2014a). Conceptualizing Open Data ecosystems: A timeline analysis of Open Data development in the UK. In *In Proceedings of the International Conference for E-Democracy and Open Government*.

Heimstädt, M., Saunderson, F., & Heath, T. (2014b). From Toddler to Teen: Growth of an Open Data Ecosystem. *eJournal of eDemocracy & Open Government*, 6(2), 123–135.

Helbig, N., Cresswell, A. M., Burke, G. B., & Luna-Reyes, L. (2012). *The Dynamics of Opening Government Data*. Center for Technology in Government - University at Albany. Retrieved from <https://www.ctg.albany.edu/publications/reports/opendata/opendata.pdf>

Hernon, P. (1991). Government information policy principles. *Government Information Quarterly*, 8(4), 393–399.

Hogge, B. (2010). *Open Data Study – New Technologies*. Open Society Foundation. Retrieved from http://www.transparency-initiative.org/wp-content/uploads/2011/05/open_data_study_final.pdf

Holsti, O. R. (1968). Content analysis. In Gardner Lindzey and Elliott Aronson (Ed.), *The Handbook of Social Psychology* (pp. 596–692). Addison-Wesley.

- Horrigan, J. B., & Rainie, L. (2015). *Americans' Views on Open Government Data*. Pew Research Center. Retrieved from <http://www.pewinternet.org/2015/04/21/open-government-data/>
- Howard, A. (2012). Data for the Public Good. *O'Reilly Radar*. Retrieved from <http://radar.oreilly.com/2012/02/data-public-good.html>
- Howard, A. (2013). Open data economy: Eight business models for open data and insight from Deloitte UK. *O'Reilly Radar*. Retrieved from <http://radar.oreilly.com/2013/01/open-data-business-models-deloitte-insight.html>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Hudson-Smith, A., Batty, M., Crooks, A., & Milton, R. (2008). Mapping for the Masses: Accessing Web 2.0 through Crowdsourcing. *Social Science Computer Review*, 27(4), 524–538. Retrieved from <http://discovery.ucl.ac.uk/15198/>
- Huijboom, N., & Van den Broek, T. (2011). Open data: an international comparison of strategies. *European Journal of ePractice*, 12. Retrieved from http://www.epractice.eu/files/European Journal epractice Volume 12_1.pdf
- Huizingh, E. K. R. E. (2011). Open innovation: State of the art and future perspectives. *Technovation*, 31(1), 2–9.
- Iemma, R. (2012). *Open Government Data: A Focus on Key Economic and Organizational Drivers*. Nexa Center for Internet & Society. Retrieved from http://porto.polito.it/2520895/1/SSRN_id2262943.pdf
- Iglesias, C., & Boyera, S. (2015). *Case study: How Aidsplan use open data to track health spending*. Open Data Institute. Retrieved from <http://theodi.org/case-studies/aidsplan-case-study-a-watchdog-organisation-in-the-health-sector>
- Immonen, A., Palviainen, M., & Ovaska, E. (2014). Towards open data based business: Survey on usage of open data in digital services. *International Journal of Research in Business and Technology*, 4(1), 286–295.
- Janssen. (2011a). The influence of the PSI directive on open government data: An overview of recent

- developments. *Government Information Quarterly*, 28(4), 446–456.
- Janssen. (2011b). The role of public sector information in the European market for online content: a never-ending story or a new beginning? *Info*, 13(6), 20–29.
- Janssen. (2012). Open Government Data and the Right to Information: Opportunities and Obstacles. *The Journal of Community Informatics*, 8(2). Retrieved from <http://ci-journal.net/index.php/ciej/article/view/952>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258–268.
- Janssen, M., & van den Hoven, J. (2015). Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy? *Government Information Quarterly*. <http://doi.org/10.1016/j.giq.2015.11.007>
- Janssen, M., & Zuiderwijk, A. (2014). Infomediary Business Models for Connecting Open Data Providers and Users. *Social Science Computer Review*, 32(5), 694–711.
- Jeffery, K. G., Asserson, A., Houssos, N., Brasse, V., & Jörg, B. (2014). From Open Data to Data-intensive Science through CERIF. *Procedia Computer Science*, 33, 191–198.
- Jellema, A., Meijninger, W., & Addison, C. (2015). *Open Data and Smallholder Food and Nutritional Security*. Technical Centre for Agricultural and Rural Cooperation. Retrieved from http://www.cta.int/images/Opendataforsmallholders-report_.pdf
- Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2012). The Value of Open Government Data: A Strategic Analysis Framework. In *SIG eGovernment pre-ICIS Workshop*. Orlando.
- Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2013). The generative mechanisms of open government data. In *21st European Conference on Information Systems*.
- Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2014). Data-driven innovation through open government data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 100–120.
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013). Generating value from open government data. *International Conference on Information Systems (ICIS 2013): Reshaping Society Through Information*

Systems Design, 2(July), 1737–1756.

Johnson, J. A. (2014). From Open Data to Information Justice. *Ethics and Information Technology*, 16(4), 263–274.

Johnson, P., & Robinson, P. (2014). Civic Hackathons: Innovation, procurement, or civic engagement? *Review of Policy Research*. <http://doi.org/10.1111/ropr.12074>

Kaasenbrood, M., Zuiderwijk, A., Janssen, M., de Jong, M., & Bharosa, N. (2015). Exploring the Factors Influencing the Adoption of Open Government Data by Private Organisations. *International Journal of Public Administration in the Digital Age*, 2(2), 75–92.

Kalampokis, E., Hausenblas, M., & Tarabanis, K. (2011). Combining social and government open data for participatory decision-making. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6847 LNCS, 36–47.

Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly*, 30(4), 508–513.

Kassen, M. (2013). Globalization of E-government: Open Government as a Global Agenda, Benefits, Limitations and Ways Forward. *Information Development*, 1–8.

Kerschberg, B. (2011). Metadata, The Freedom of Information Act, and Government Hypocrisy. *Forbes*. Retrieved from <http://www.forbes.com/sites/benkerschberg/2011/04/11/metadata-the-freedom-of-information-act-and-government-hypocrisy/#72e00f9a5ac3>

Keserú, J. (2015). A new approach to measuring the impact of open data. *The Sunlight Foundation Blog*. Retrieved from <https://sunlightfoundation.com/blog/2015/05/05/a-new-approach-to-measuring-the-impact-of-open-data/>

Keserú, J., & Chan, J. K. (2015). *The social impact of open data*. Sunlight Foundation. Retrieved from <http://www.opendataresearch.org/dl/symposium2015/odrs2015-paper20.pdf>

King, J. (2013). Canada Post goes postal, asserts trademark rights for the term “postal code.” *Canadaone.com*. Retrieved from http://www.canadaone.com/ezine/briefs.html?StoryID=13Apr24_1

Kitchin, R. (2013). Four critiques of open data initiatives. *London School of Economics Blog*. Retrieved

from <http://blogs.lse.ac.uk/impactofsocialsciences/2013/11/27/four-critiques-of-open-data-initiatives/>

- Klievink, B., Zuiderwijk, A., & Janssen, M. (2014). Interconnecting governments, businesses and citizens - A comparison of two digital infrastructures. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8653 LNCS, pp. 84–95). Springer Verlag.
- Kondracki, N. L., Wellman, N. S., & Amundson, D. R. (2002). Content Analysis: Review of Methods and Their Applications in Nutrition Education. *Journal of Nutrition Education and Behavior*, 34(4), 224–230.
- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology* (2nd ed., Vol. 79). Thousand Oaks.
- Kroes, N. (2011). Unlocking the goldmine: New legal proposals to open up Europe's public sector. *European Commission - Neelie's Blog*. Retrieved from <http://blogs.ec.europa.eu/neelie-kroes/opendata/>
- Kube, M., Hilgers, D., Fueller, J., & Koch, G. (2013). Re-Thinking Citizen Participation in the Public Sector: Adopting Open Innovation for Collaborative Idea Generation. *SSRN Electronic Journal*, 1–43.
- Kuk, G., & Davies, T. (2011). The Roles of Agency and Artifacts in Assembling Open Data Complementarities. *ICIS 2011 Proceedings*, 1–16.
- Kulk, S., & Loenen, B. Van. (2012). Brave New Open Data World? *International Journal of Spatial Data Infrastructures Research*, 7, 196–206.
- Lamb, C. M. (2001). Creating a collaborative environment: the human element. *Information Outlook*, 5(5), 22–25.
- Lambert, S. (2006). A Business Model Research Schema. In *19th Bled eConference eValues* (pp. 1–13). Flinders University, School of Commerce, Australia.
- Lassinantti, J., Bergvall-Kareborn, B., & Stahlbrost, A. (2014). Shaping local open data initiatives:

- Politics and implications. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 17–33.
- Latham, M. (2003). Democracy and infomediaries. *Corporate Governance International Review*, 11(2), 91–101.
- Lathrop, D., & Ruma, L. (2010). *Open Government: Collaboration, Transparency, and Participation in Practice* (1st ed.). O'Reilly Media.
- Latif, A., Saeed, A. U., Hoefler, P., Stocker, A., & Wagner, C. (2009). The linked data value chain: A lightweight model for business engineers. *Proceedings of ISEMANTICS09 International Conference on Semantic Systems*, (September), 568–575.
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management*, 40(2), 133–146.
- Lemieux, V. (2014). Why we're failing to get the most out of open data. *World Economic Forum*. Retrieved from <https://www.weforum.org/agenda/2014/09/open-data-information-governance-quality/>
- Linders, D. (2012). How can open development improve the effectiveness of aid? Leveraging open data, open standards, and web 2.0 interactivity for better development outcomes. In *ACM International Conference Proceeding Series* (pp. 155–164).
- Linders, D. (2013). Towards open development: Leveraging open data to improve the planning and coordination of international aid. *Government Information Quarterly*, 30(4), 426–434.
- Linders, D., & Wilson, S. C. (2011). What is open government?: one year after the directive. In *12th Annual International Conference on Digital Government Research* (pp. 262–271). ACM.
- Lindman, J., & Nyman, L. (2014). The Businesses of Open Data and Open Source: Some Key Similarities and Differences. *Technology Innovation Management Review*, (January), 12–17. Retrieved from <http://timreview.ca/article/757>
- Lindman, J., Rossi, M., & Tuunainen, V. (2013). Open Data Services: Research Agenda. In *Hawaii International Conference on System Sciences (HICSS-46)* (pp. 1239–1246). Grand Wailea, Maui, Hawaii.

- Living Cities. (2012). *Field Scan of Civic Technology*. Living Cities. Retrieved from <https://www.livingcities.org/resources/131-field-scan-of-civic-technology>
- Lohr, S. (2012). The age of Big Data. *The New York Times*. Retrieved from <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>
- Longo, J. (2011). OpenData: Digital-Era Governance Thoroughbred or New Public Management Trojan Horse? *Public Policy Governance Review*, 2(2), 38–51.
- Luna-Reyes, L. F., & Chun, S. A. (2012). Open government and public participation: Issues and challenges in creating public value (introduction). *Information Polity: The International Journal of Government & Democracy in the Information Age*, 17(2), 77–81.
- Lundvall, B.-Å., & Johnson, B. (1994). The learning economy. *Journal Of Industry Studies*, 1(2), 23–42.
- Maali, F., Cyganiak, R., & Peristeras, V. (2010). Enabling interoperability of government data catalogues. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6228 LNCS, pp. 339–350).
- Magalhaes, G., Roseira, C., & Strover, S. (2013). Open government data intermediaries. In *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance - ICEGOV '13* (pp. 330–333). New York, New York, USA: ACM Press.
- Manyika, J., Chui, M., Farrell, D., Kuiken, S. Van, Groves, P., & Doshi, E. A. (2013). *Open data: Unlocking innovation and performance with liquid information*. McKinsey Global Institute.
- Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousand Oaks.
- Martin, C. J. (2014). Barriers to the Open Government Data Agenda: Taking a Multi-Level Perspective. *Policy and Internet*, 6(3), 217–240.
- Martin, Foulonneau, M., Turki, S., & Ihadjadene, M. (2013). Risk Analysis to Overcome Barriers to Open Data. *Electronic Journal of E-Government*, 11(2), 348–359.
- Maxwell, J. A. (2013). *Qualitative Research Design: An Interactive Approach* (3rd ed.). Sage Publications.
- Mayer-Schönberger, V. (2012). Mapping the Ecology of Open Data Development. *Policy by the Numbers*.

Retrieved from <http://policybythenumbers.blogspot.com/2012/10/mapping-ecology-of-open-data-development.html>

- McDermott, P. (2010). Building open government. *Government Information Quarterly*, 27(4), 401–413.
- McLeod, J. (2012). Thoughts on the opportunities for records professionals of the open access, open data agenda. *Records Management Journal*, 22(2), 92–97.
- Meijer, A. (2009). Understanding modern transparency. *International Review Of Administrative Sciences*, 75(2), 255–269.
- Meijer, A., Hoog, J. de, Twist, M. van, Steen, M. van der, & Scherpenisse, J. (2014). Understanding the Dynamics of Open Data From Sweeping Statements to Complex Contextual Interactions. In *Open Government: Opportunities and Challenges for Public Governance* (pp. 101–114). New York: Springer.
- Meijer, A. J. (2007). Publishing public performance results on the Internet. Do stakeholders use the Internet to hold Dutch public service organizations to account? *Government Information Quarterly*, 24(1), 165–185.
- Meijer, Curtin, D., & Hillebrandt, M. (2012). Open government: connecting vision and voice. *International Review of Administrative Sciences*, 78(1), 10–29.
- Meng, A. (2014). Investigating the Roots of Open Data's Social Impact. *JeDEM - eJournal of eDemocracy and Open Government*, 6(1), 1–13.
- Mercado-Lara, E., & Gil-Garcia, J. R. (2014). Open government and data intermediaries: the case of AidData. In *Proceedings of the 15th Annual International Conference on Digital Government Research - dg.o '14* (pp. 335–336). New York, New York, USA: ACM Press.
- Miles, M., & Huberman, A. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). SAGE Publications.
- Misuraca, G., Mureddu, F., & Osimo, D. (2014). Policy-Making 2.0: Unleashing the Power of Big Data for Public Governance. In *Open Government* (pp. 171–188). Springer.
- Misuraca, G., & Viscusi, G. (2014). Is Open Data Enough?: E-Governance Challenges for Open Government. *International Journal of Electronic Government Research*, 10(1), 18–34.

- Molloy, J. C. (2011). The Open Knowledge Foundation: Open Data Means Better Science. *PLoS Biology*, 9(12), e1001195.
- Noveck, B., & Gorof, D. L. (2013). *Information for Impact: Liberating Nonprofit Sector Data*. The Aspen Institute.
- O'Hara, K. (2012a). Data quality, government data and the open data infosphere. In *AISB/IACAP World Congress 2012: Information Quality Symposium*. Birmingham.
- O'Hara, K. (2012b). Transparency, open data and trust in government. In *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12* (pp. 223–232). New York, New York, USA: ACM Press.
- O'Reilly, T. (2010). Government as a platform. In D. Lathrop & L. Ruma (Eds.), *Open Government: Collaboration, Transparency, and Participation in Practice* (pp. 11–39). O'Reilly Media.
- Obama, B. (2009). *Transparency and open government. Memorandum for the heads of executive departments and agencies*. The White House. Retrieved from https://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment
- Obama, B. (2013). *Executive Order -- Making Open and Machine Readable the New Default for Government Information*. The White House. Retrieved from <http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government->
- OECD. (2012). *OECD Science, Technology and Industry Outlook 2012*. OECD. Retrieved from http://www.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-industry-outlook-2012_sti_outlook-2012-en
- Ohm, P. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*, 57, 1701–1777.
- Onetti, A., Zucchella, A., Jones, M., & McDougall-Covin, P. (2012). Internationalization, innovation and entrepreneurship: business models for new technology-based firms. *Journal of Management and Governance*, 16(3), 337–368.
- Orszag, P. R. (2009). *Memorandum for the heads of executive departments and agencies*. The White House. Retrieved from

https://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf

- Osterwalder, A. (2004). *The Business Model Ontology: A proposition in a design science approach* (Doctoral thesis). Business. Université de Lausanne. Retrieved from http://www.hec.unil.ch/aosterwa/PhD/Osterwalder_PhD_BM_Ontology.pdf
- Otto, B., & Aier, S. (2013). Business Models in the Data Economy: A Case Study from the Business Partner Data Domain. In *11th International Conference on Wirtschaftsinformatik* (pp. 475–489). Leipzig, Germany.
- Palmer, C. (2015). Wisconsin votes to hide all government actions from the public. *Inquisitr*. Retrieved from <http://www.inquisitr.com/2223882/wisconsin-votes-to-hide-all-government-actions-from-the-public/>
- Park, T., & Vanroekel, S. (2013). The American Dream, Aided By Open Government Data. *The White House Blog*. Retrieved from <https://www.whitehouse.gov/blog/2013/08/07/american-dream-aided-open-government-data>
- Patel, M., Sotsky, J., Gourley, S., & Houghton, D. (2013). *The Emergence of Civic Tech: Investments in a Growing Field*. Knight Foundation. Retrieved from http://www.knightfoundation.org/media/uploads/publication_pdfs/knight-civic-tech.pdf
- Patil, D. (2015). Two Years of Transformative Open Data for Public Good. *The White House Blog*. Retrieved from <https://www.whitehouse.gov/blog/2015/05/11/two-years-transformative-open-data-public-good>
- Peixoto, T. (2013). The Uncertain Relationship between Open Data and Accountability: A Response to Yu and Robinson's "The New Ambiguity of Open Government." *60 UCLA L. Rev. Disc.* 200, 200–213.
- Peled, A. (2011). When transparency and collaboration collide: The USA Open Data program. *Journal of the American Society for Information Science and Technology*, 62(11), 2085–2094.
- Peled, A. (2013). Re-Designing Open Data 2.0. *eJournal of eDemocracy & Open Government*, 5(2), 187–199.
- Perricos, C., Branch, D., & Lewis, H. (2012). *Open growth: Stimulating demand for open data in the UK*.

Deloitte LLP. Retrieved from <http://www2.deloitte.com/uk/en/pages/deloitte-analytics/articles/stimulating-demand-for-open-data-in-the-uk.html>

- Petychakis, M., Vasileiou, O., Georgis, C., Mouzakitis, S., & Psarras, J. (2014). A state-of-the-art analysis of the current public data landscape from a functional, semantic and technical perspective. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 34–47.
- Pollock, R. (2009). *The Economics of Public Sector Information* (Cambridge Working Papers in Economics No. 920).
- Pollock, R. (2011). Building the (Open) Data Ecosystem. *Open Knowledge Blog*. Retrieved from <http://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/>
- Porter, M. E., & Millar, V. E. (1985). How information gives you competitive advantage. *Harvard Business Review*, 63(4 (July–August 1985)), 149–160.
- Rath, S. (2012). Data Template For District Economic Planning. *Journal of Community Informatics*, 8(2). Retrieved from <http://ci-journal.net/index.php/ciej/article/view/826/924>
- Ren, G.-J., & Glissmann, S. (2012). Identifying Information Assets for Open Data: The Role of Business Architecture and Information Quality. In *2012 IEEE 14th International Conference on Commerce and Enterprise Computing* (pp. 94–100).
- Robinson, D., Yu, H., Zeller, W., & Felten, E. W. (2009). Government Data and the Invisible Hand. *Yale Journal of Law Technology*, 11(11), 160–175.
- Rogers, S. (2012). Open data journalism. *The Guardian*. Retrieved from <http://www.theguardian.com/news/datablog/2012/sep/20/open-data-journalism>
- Rose, F. (1999). *The Economics, Concept, and Design of Information Intermediaries* (1st ed.). Physica-Verlag Heidelberg.
- Ruppert, E. (2013). Doing the Transparent State: open government data as performance indicators. In J. Mugler & S. J. Park (Eds.), *A World of Indicators: The production of knowledge and justice in an interconnected world* (pp. 51–78). Cambridge: Cambridge University Press.
- Sadiq, S., Yeganeh, N. K., & Indulska, M. (2011). 20 Years of Data Quality Research : Themes , Trends

- and Synergies. *Proceedings Australian Database Conference, Perth, Australia*, 115, 1–10.
- Sandoval-Almazan, R. (2012). Open government 2.0: citizen empowerment through open data, web and mobile apps. *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance - ICEGOV '12*, 30–33.
- Sandoval-Almazan, R. (2015). Gobierno abierto y transparencia: Construyendo un marco conceptual. *Convergencia*, 22(68), 203–227.
- Sattler, K.-U. (2009). Data Quality Dimensions. In L. Liu & M. T. Özsu (Eds.), *Encyclopedia of Database Systems* (pp. 612–615).
- Sayogo, D. S., & Pardo, T. A. (2012). Exploring the Motive for Data Publication in Open Data Initiative: Linking Intention to Action. In *2012 45th Hawaii International Conference on System Sciences* (pp. 2623–2632). IEEE.
- Schalkwyk, F. van, Cañares, M., Chattapadhyay, S., & Andrason, A. (2015). *Open data intermediaries in Developing countries*. World Wide Web Foundation. Retrieved from http://webfoundation.org/wp-content/uploads/2015/08/ODDC_2_Open_Data_Intermediaries_15_June_2015_FINAL.pdf
- Schrier, B. (2010). Toads on the Road to Open Government Data. In D. Lathrop & L. Ruma (Eds.), *Open Government: Collaboration, Transparency, and Participation in Practice* (pp. 311–320). O'Reilly Media.
- Shkabatur, J. (2012). Transparency With(out) Accountability: Open Government in the United States. *Yale Law & Policy Review*, 31(1), 79–140.
- Simpson, A. C. (2011). On Privacy and Public Data: A study of data.gov.uk. *Journal of Privacy and Confidentiality*, 3(1), 51–65.
- Sofaer, S. (1999). Qualitative methods: what are they and why use them? *Health Services Research*, 34(5 Pt 2), 1101–1118.
- Solaimani, S., & Bouwman, H. (2012). A framework for the alignment of business model and business processes: A generic model for trans-sector innovation. *Business Process Management Journal*, 18(4), 655–679.
- Song, J., & Zahedi, F. (2002). A Theoretical Framework for the use of Web Infomediaries. In *AMCIS*

2002 Proceedings (p. Paper 308).

- Steininger, D. M., Huntgeburth, J. C., & Veit, D. J. (2011). Conceptualizing Business Models for Competitive Advantage Research by Integrating the Resource and Market-Based Views. In *AMCIS 2011 Proceedings* (pp. 1–12).
- Steininger, D. M., Trenz, M., & Veit, D. J. (2013). Building Taxonomies in IS and Management – A Systematic Approach Based on Content Analysis. In *Wirtschaftsinformatik Proceedings* (p. Paper 90).
- Stott, A. (2014). *Open data for economic growth*. World Bank (Transport & ICT Global Practice Series). Retrieved from <http://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf>
- Sun, L. (2016). Internet of Things in 2016: 6 Stats Everyone Should Know. *The Motley Fool*. Retrieved from <https://www.fool.com/investing/general/2016/01/18/internet-of-things-in-2016-6-stats-everyone-should.aspx>
- Sweeney, L. (1997). Weaving technology and policy together to maintain confidentiality. *The Journal of Law, Medicine & Ethics: A Journal of the American Society of Law, Medicine & Ethics*, 25(2–3), 98–110.
- Tammisto, Y., & Lindman, J. (2012). Definition of Open Data Services in Software Business. In *Third International Conference, ICSOB 2012* (pp. 297–303). Cambridge, MA.
- Tennison, J. (2015). How can government support open data without doing it all itself? *Government Computing*. Retrieved from <http://business-applications.governmentcomputing.com/features/how-can-government-support-open-data-without-doing-it-all-itself-4566315>
- The Economist. (2015). Open government data: Out of the box. *The Economist Newspaper*. Retrieved from <http://www.economist.com/news/international/21678833-open-data-revolution-has-not-lived-up-expectations-it-only-getting>
- The White House. (2013). Obama Administration Releases Historic Open Data Rules to Enhance Government Efficiency and Fuel Economic Growth. *Secretary Office of the Press*. Retrieved from <https://www.whitehouse.gov/the-press-office/2013/05/09/obama-administration-releases-historic-open-data-rules-enhance-governmen>

- Tinholt, D. (2012). *The Open Data Economy Unlocking Economic Value by Opening Government and Public Data*. Capgemini Consulting.
- Tong, A., Irshad, H., & Ward Revell, D. (2013). *Market assessment of public sector information*. UK Government - Department for Business Innovation and Skills.
- Tran, A. (2012). Summit Week: Anatomy of A Civic Startup. *Code for America*. Retrieved from <http://codeforamerica.org/2012/10/02/summit-week-anatomy-of-a-civic-startup/>
- Ubaldi, B. (2013). Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. *OECD Working Papers on Public Governance*, 22.
- UK Cabinet Office. (2012). *Open Data White Paper - Unleashing the potential*. Retrieved from https://data.gov.uk/sites/default/files/Open_data_White_Paper_10.pdf
- United Nations. (2003). *UNDP Practice Note: Access to Information*. Retrieved from http://www.undp.org/content/dam/aplaws/publication/en/publications/democratic-governance/dg-publications-for-website/access-to-information-practice-note/A2I_PN_English.pdf
- United Nations. (2010). Stimulus funds, transparency and public trust. In *United Nations E-Government Survey* (pp. 9–23).
- Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences*, 15(3), 398–405.
- van Brakel, R. (2014). Stimulating stakeholders to open up education data. *European Public Sector Information Platform*. Retrieved from <http://www.epsiplatform.eu/content/stimulating-stakeholders-open-education-data>
- Van Veenstra, A. F., & Van Den Broek, T. A. (2013). Opening moves-drivers, enablers and barriers of open data in a semi-public organization. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8074 LNCS, pp. 50–61).
- Vasa, M., & Tamilselvam, S. (2014). Building apps with open data in india: an experience. *Proceedings of the 1st International Workshop on Inclusive Web Programming - Programming on the Web with Open Data for Societal Applications - IWP 2014*, 1–7.

- Verhulst, S., & Caplan, R. (2015). *Open Data: A 21st Century Asset for Small and Medium Sized Enterprises*. GovLab. Retrieved from <http://images.thegovlab.org/wordpress/wp-content/uploads/2015/08/OpenData-and-SME-Final-Aug2015.pdf>
- Vickery, G. (2011). *Review of recent studies on PSI re-use and relevant market developments*. European Commission. Retrieved from http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
- Viscusi, G., Castelli, M., & Batini, C. (2014). Assessing Social Value in Open Data Initiatives: A Framework. *Future Internet*, 6(3), 498–517.
- von Hippel, E. (2010). Comment on “Is open innovation a field of study or a communication barrier to theory development?” *Technovation*, 30(11–12), 555.
- Voss, K. (2015). Open Data Emerges as Driving Force in Civic Technology. *Onvia*. Retrieved from <http://www.onvia.com/blog/open-data-emerges-as-driving-force-in-civic-technology>
- Walker, H. (2014). Harvard Poll Shows Millennials Have “Historically Low” Levels Of Trust In Government. *Business Insider*. Retrieved from <http://www.businessinsider.com/poll-millennials-have-historically-low-levels-of-trust-in-government-2014-4>
- Webber, S. (2004). *The Success of Open Source*. Harvard University Press.
- Weber, M. (1978). *Economy and Society. An Outline of Interpretative Sociology*. *Economy and Society. An Outline of Interpretative Sociology*. University of California Press.
- Whitmore, A. (2014). Using open government data to predict war: A case study of data and systems challenges. *Government Information Quarterly*, 31(4), 622–630.
- Wihbey, J. (2013). Open government and conflicts with public trust and privacy: Recent research ideas. *Journalist's Resource*. Retrieved from <http://journalistsresource.org/studies/politics/digital-democracy/government-transparency-conflicts-public-trust-privacy-recent-research-ideas>
- Williams, O. (2015). 2015 worldwide Internet, mobile and social media trends: get into 376 pages of data. *TWNNews*. Retrieved from <http://thenextweb.com/socialmedia/2015/01/21/2015-worldwide-internet-mobile-social-media-trends-get-376-pages-data/>
- Wirtz, B. W., & Birkmeyer, S. (2015). *Open Government: Origin, Development, and Conceptual*

Perspectives. *International Journal of Public Administration*, 38(5), 381–396.

Woods, J. (2001). The Commercial Exploitation of Private Sector Information in the European Union. *Computer Law & Security Report*, 17(4), 244–246.

Woods, M. (2011). Interviewing for research and analysing qualitative data: An overview. *Massey University*.

Worthy, B. (2013). David Cameron's Transparency Revolution? The Impact of Open Data in the UK. *SSRN Electronic Journal*, 1–24.

Yang, Z., & Kankanhalli, A. (2013). Innovation in Government Services: The Case of Open Data. In Y. K. Dwivedi, H. Z. Henriksen, D. Wastell, & R. De' (Eds.), *Grand Successes and Failures in IT. Public and Private Sectors* (Vol. 402, pp. 644–651). Berlin, Heidelberg: Springer Berlin Heidelberg.

Yankova, M. (2016). Journalism in the Age of Open Data. *Ontotext*. Retrieved from <http://ontotext.com/journalism-in-the-age-of-open-data/>

Yannoukakou, A., & Araka, I. (2014). Access to Government Information: Right to Information and Open Government Data Synergy. *Procedia - Social and Behavioral Sciences*, 147, 332–340.

Yates, D., & Paquette, S. (2011). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *International Journal of Information Management*, 31(1), 6–13.

Yu, H., & Robinson, D. G. (2012). The New Ambiguity of "Open Government." *59 UCLA L. Rev. Disc.* 178, 178–208.

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. In *Applications of social research methods to questions in information and library science* (pp. 308–319). Westport, CT: Libraries Unlimited.

Zott, C., Amit, R., & Massa, L. (2010). *The business model: Theoretical roots, recent developments, and future research* (IESE Working Paper). *IESE Research Papers* (Vol. 3). IESE Business School.

Zuiderwijk, A., Helbig, N., Gil-Garcia, J. R., & Janssen, M. (2014). Special Issue on Innovation through Open Data - A Review of the State-of-the-Art and an Emerging Research Agenda: Guest Editors'

Introduction. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 1–2.

Zuiderwijk, A., & Janssen, M. (2014a). Barriers and Development Directions for the Publication and Usage of Open Data: A Socio-Technical View. In M. Gascó-Hernández (Ed.), *Open Government* (pp. 115–135). Springer New York.

Zuiderwijk, A., & Janssen, M. (2014b). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17–29.

Zuiderwijk, A., & Janssen, M. (2014c). The Negative Effects of Open Government Data - Investigating the Dark Side of Open Data. *Proceedings of the 15th Annual International Conference on Digital Government Research*, 147–152.

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of E-Government*, 10(2), 156–172.

Zuiderwijk, A., Janssen, M., & Dwivedi, Y. K. (2015). Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Government Information Quarterly*, pp. 429–440. Elsevier Ltd.

Zuiderwijk, A., Janssen, M., Poulis, K., & van de Kaa, G. (2015). Open data for competitive advantage. In *Proceedings of the 16th Annual International Conference on Digital Government Research - dg.o '15* (pp. 79–88). New York, New York, USA: ACM Press.

Zurada, J., & Karwowski, W. (2011). Knowledge Discovery Through Experiential Learning From Business and Other Contemporary Data Sources: A Review and Reappraisal. *Information Systems Management*, 28(3), 258–274.

Appendix I (Presentation of firms)

Amazon Web Services

<http://aws.amazon.com>

Industry: Information Technology and Services

Size: 10,001+ employees

Founded: 2006

Amazon Web Services provides an elastic, open and flexible, secure infrastructure platform in the cloud that powers hundreds of thousands of businesses in 190 countries around the world. With data center locations in the U.S., Europe, Singapore, and Japan, they have customers across all industries.

Aunt Bertha

<http://www.auntbertha.com>

Industry: Computer Software

Size: 11-50 employees

Founded: 2010

Aunt Bertha is the leading social services search engine in the United States. It is the developer of an enterprise software platform addressing the social determinants of health, used by most of the top medicaid managed care providers in the US as well as a growing number of hospitals and health care providers.

Billguard

Billguard is a personal finance analytics company that uses both crowdsourcing as well as big data analytics in order to help consumers better protect and do more with their money, saved money. In that context Billguard aggregates financial data and consumer knowledge about financial data from multiple sources including a community of over 1 million users, as well as from banks, the Internet, and from the consumer financial protection Bureau.

<http://www.billguard.com>

Industry: Internet

Company Size: 11-50 employees

Founded: 2010

Calcbench

<http://www.calcbench.com>

Industry: Financial Services

Size: 1-10 employees

Founded: 2011

Calcbench is the a market financial data platform powered by XBRL. We provide financial and accounting data from over 9,000 US listed public companies faster, in more detail, and at a better value than traditional providers. Financial analysts, accountants, academics and investors instantly access, work with and share SEC filing data, both online and with our Excel add-in. Calcbench also provides data wholesaling, XBRL analysis, and white labeling services to enterprise customers.

Civic Insight

<http://civicinsight.com>

Industry: Computer Software

Company Size: 1-10 employees

Civic Insight makes real-time government data about a community's underutilized spaces – think abandoned homes, empty storefronts, or vacant lots – both publicly available and extremely easy to use and understand. By getting everyone on the same page, Civic Insight makes it easier for government staff, motivated citizens, and local organizations to collaborate more effectively to improve the quality and value of their neighborhoods.

Fastcase

<http://www.fastcase.com>

Industry: Legal Services

Company Size: 51-200 employees

Founded: 1999

Fastcase is an American provider of online legal research. The company's research software provides online access to millions of cases, statutes, and regulations, and at a fraction of the cost of existing alternatives. Fastcase was built by lawyers and for lawyers, with the input of specialists in legal research, library science, and law. Fastcase's mission is to make better research available to more people through innovation, value, and choice.

Relationship Science

<http://www.relsoci.com>

Industry: Information Services

Size: 501-1000 employees

Founded: 2010

Rel Sci empowers organizations to make smarter business decisions through a deeper understanding of an underutilized asset: their relationship capital with influential decision makers. By turning individual knowledge into institutional knowledge across the enterprise, Rel Sci helps clients be more efficient, collaborative, agile and opportunistic by helping them build better businesses.

TrialX

<http://www.trialx.com>

Industry: Internet

Size: 1-10 employees

TrialX is a web-based platform that combines aspects of decision engines (algorithms) and social networks (the power of the community) to help patients find clinical trials that match their health conditions and connect with investigators conducting trials quickly and effortlessly. The company's mission is to help save lives by enabling patients to find new treatments in an easy and timely manner.

TuvaLabs

<http://www.tuvalabs.com>

Industry: E-Learning

Size: 11-50 employees

Founded: 2012

Tuva (tuvalabs.com) is a data literacy company that empowers data novices to learn essential data and statistics literacy skills. Tuva offers data literacy solutions for schools, businesses, and organizations working in international development. Through our growing library of interactive, authentic, and relevant modules and lessons, participants are able to learn and master the most-critical, foundational concepts in data, statistics, and analytics.

Zebu Compliance

<http://www.YEIHealthcare.com>

Industry: Computer Software

Size: 11-50 employees

Founded: 1999

Zebu Compliance Solutions (formerly YEI Healthcare) improves medical coding efficiency, reimbursement assurance, auditing toolsets, claim accuracy, and staff licensing/exclusion monitoring. Clients reduce labor costs, increase revenue, and maintain audit-readiness.

Zillow

<http://www.zillow.com>

Industry: Internet

Size: 1001-5000 employees

Founded: 2006

Zillow serves the full lifecycle of owning and living in a home: buying, selling, renting, financing, remodeling and more. It starts with Zillow's living database of more than 110 million U.S. homes – including homes for sale, homes for rent and homes not currently on the market, as well as Zestimate home values, Rent Zestimates and other home-related information. Zillow operates the most popular suite of mobile real estate apps, with more than two dozen apps across all major platforms.

Beyond Value

<http://www.beyondvalue.com>

Industry: Internet

Size: 1-10 employees

Founded: 2003

Beyond Value was created initially as a real estate valuation firm but its primary product is now software. The company's main product is Zonability, a web application that provides an efficient way for professionals and everyday people to access local zoning. It is geared for economic development corporations and title companies. The power of having structured data on a single platform allows for a consistent user experience and regional analysis.

Appendix II (Interview questions)

Introductory questions

Can you please introduce yourself?

Can you briefly describe what does [company name] do?

Could you explain how does [company name] use open government data?

How important is open government data to your business and why? How will this importance evolve in the future?

Perceptions on the open government data network and value

Who are the main actors in OGD ecosystem? What are their roles?

Where would you say your company is located in that ecosystem?

Accessibility

How do you find the datasets you need? Do you experience any difficulties in the process?

How do you access the dataset?? Do you encounter any barriers in accessing the data you need?

Which ones?

Do you have difficulty finding open data in the format that you like to work with? How do you deal with that?

Data usability

A number of open data users have reported issues of data quality (namely accuracy, completeness, consistency, and timeliness). Has the issue of data quality been a problem for you?

Similarly, other open data users have complained about the lack of metadata. Has this been an issue in your use of open data?

RESOURCES

Has it been a challenge for [company name] to have the necessary resources, namely hardware and software, to use open data?

Some users have noted difficulty in find human resources with the necessary skills or competences to work with open data efficiently. Has this been an issue for [company name]?

INTERACTION

What types of contacts (social, technical, information) is it?

Is/Would this be important to your business?

Do have contact with other users that use the same data that you use? What types of contact is it (social, technical, information)?

Is/Would this be important to your business?

LEGAL

Other open data users have remarked that the inexistence or inconsistency of licenses in open datasets can be an issue. How do you deal with this at [company name]?

Similarly, is the aspect of the privacy of the data an issue for you? For example, do you ever feel the threat of being legally liable if data is misused? If so, how do you deal with that?

CONCLUDING QUESTIONS

Are there other issues, besides the ones covered in this interview, that you would like to flag as currently being barriers in the use of open data by the private sector?