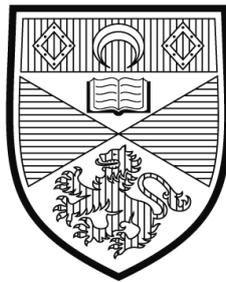


UNIVERSITY OF ST ANDREWS
SCHOOL OF COMPUTER SCIENCE

MAXIMILIAN HEIMSTÄDT

MATRICULATION NUMBER: 120016757
M.SC. MANAGEMENT & INFORMATION TECHNOLOGY

The British Open Data Ecosystem



University of
St Andrews

Author:
Maximilian HEIMSTÄDT

Supervisor:
Dr. Tom KELSEY

February 28, 2014

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Declaration

"I hereby certify that this dissertation, which is approximately 13.500 words in length, has been composed by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree. This project was conducted by me at the University of St Andrews from June 2013 to August 2013 towards fulfilment of the requirements of the University of St Andrews for the degree of MSc under the supervision of Tom Kelsey."

"In submitting this project report to the University of St Andrews, I give permission for it to be made available for use in accordance with the regulations of the University Library. I also give permission for the title and abstract to be published and for copies of the report to be made and supplied at cost to any bona fide library or research worker, and to be made available on the World Wide Web. I retain the copyright in this work."

Date: February 28, 2014

Signature: Maximilian Heimstädt

Acknowledgement

I am very grateful to a number of people and organisations that made this piece of academic work possible. I would like to thank my family and friends for their kind support during the last year. Furthermore, I would like to thank the Foundation of German Business, which provided me with a scholarship to pursue my studies. Gratitude goes towards my supervisor at the University of St Andrews, Tom Kelsey, who provided me with guidance and feedback during the course of writing. This MSc dissertation was made possible due to a research grant of the Open Data Institute in London. My special gratitude goes to Tom Heath, who supervised and greatly supported my work at the ODI in July 2013. Finally, I would like to thank my interview partners, proof-readers and reviewers that helped me to shape this thesis.

Abstract

This dissertation investigates the British Open Data ecosystem as a network of public, private and formalised non-governmental actors, using an inductive case study approach. This research is grounded in the theory of business ecosystems, complexity and evolutionary economics and is backed by empirical data from interviews and observations. The macro analysis of the ecosystem revealed a self-reinforcing, reciprocal duality of actors interested in economic growth and civic empowerment. The micro analysis took a closer look at the interests and actions of individual stakeholders. Boundary organisations like the Cabinet Office or the Open Data Institute help to accelerate the circular reinforcement process between data providers and private sector companies. Stakeholders mainly interested in civic empowerment, respectively the Open Knowledge Foundation, reach their goal by working very collaboratively with stakeholders from the growth perspective. Future research in this field should focus on the relation between global NGOs and national governments and their strategies to unlock the potential of Open Data.

Contents

1	Introduction	8
2	Literature Review	10
2.1	Open Government Data	10
2.1.1	Conceptual Understanding	10
2.1.2	Research Approaches	12
2.2	Business Ecosystems	15
2.2.1	Ecosystem Analogies	15
2.2.2	Complexity and Business Ecosystems	16
2.2.3	Evolutionary Economics and Business Ecosystems	17
2.2.4	Research Framework	18
3	Methodology and Research Design	20
3.1	Research Objective	20
3.2	Case Study Approach	21
3.3	Structural Taxonomy of Ecosystem Participants	22
3.4	Case Study Database	24
4	Analysis	26
4.1	Macro Analysis: Timeline of the Ecosystem	26
4.2	Micro Analysis: Dynamics of the Ecosystem	31
4.2.1	The Interest in Economic Growth	31
4.2.2	The Interest in Civic Empowerment	34
4.2.3	Hybrid Perspectives	37
5	Conclusion	39

List of Tables

2.1	Different Types of Data Formats and Structural Properties	11
3.1	Case Database	24

List of Figures

2.1	Open Data Value Chain	12
2.2	Open Data Research Fronts	13
2.3	Coevolution, Self-organisation and Emergence	17
2.4	Integrated Business Ecosystem Framework	19
3.1	Open Data Ecosystem Taxonomy	23

List of Abbreviations

- CAQDAS** Computer-Assisted Qualitative Data Analysis Software
- CC** Creative Commons
- CEO** Chief Executive Officer
- CRD** Core Reference Data
- FOIA** Freedom of Information Act
- HMRC** Her Majesty's Revenue and Customs
- LOD** Linked Open Data
- ODI** Open Data Institute
- OECD** Organisation for Economic Co-operation and Development
- OGD** Open Government Data
- OGL** Open Government License
- OKF** Open Knowledge Foundation
- PSI** Public Sector Information
- RDF** Resource Description Framework
- SME** Small and Medium Enterprises
- SPARQL** SPARQL Protocol and RDF Query Language
- TIMMS** Trends in International Mathematics and Science Study

CHAPTER 1

Introduction

In 1957 the US lawyer and politician Wallace Parks argued in the *Washington Law Review* that "the accessibility and availability of information about executive and administrative agencies and their operations affect the distribution of power within our system of government and the functioning of our political institutions and processes". Subsequently he criticised that "a large proportion of the significant information about governmental affairs, which is known to government employees or in government files and archives, is not available outside the agencies directly concerned" (Parks, 1957). With his article, Parks demanded transparency of this Public Sector Information (PSI) and framed his idea as "Open Government". Since he died in the year his article was published, Parks was not able to rejoice when in 1966 the US government signed the Freedom of Information Act (FOIA) into law, which grants the disclosure of information and documents controlled by the United States government to the public. With his request for government information Parks referred to textual documents like working papers, reports or contracts. However, with the gradual digitalisation of governmental bodies and public administration in the last decades the doctrine of Open Government was extended with a new dimension: Open Data. Citizens, which have access to administrative data sets, can make informed and thereby more democratic decisions. Governments that publish their data sets can be held accountable for what they do more effectively. Developers and entrepreneurs can create services that are based on data sets, which have been assembled using taxpayer's money already. Open Data, Public Sector Information made available for unrestricted use and reuse, is an almost untapped resource that has the potential to unlock far-reaching societal and economic benefits. Berners-Lee (2013), one of the spokespersons of the movement, recently proclaimed that "Open Data [...] has the potential to create a better world". A precondition for this pleasant forecast to come true is that Open Data as a resource is produced, processed and consumed within a sound system to unlock its projected benefits. We might call this system an Open Data ecosystem.

1 Introduction

In 2009, the UK and the US almost simultaneously launched their national Open Data portals after the movement had slowly developed for almost a decade. This dissertation investigates the formation, development and structure of the British Open Data ecosystem. Using a longitudinal case study approach with empirical evidence from interviews and observations, this research investigates the different stakeholders that form the British Open Data ecosystem in order to understand the underlying dynamics. Based on a theoretical framework assembled from the fields of business ecosystems, complex systems and evolutionary economics, this dissertation provides a historical analysis of the whole system as well as an in-depth analysis of selected stakeholders. Primary data in the form of interviews and observations has been collected in cooperation with the Open Data Institute in London. The rationale behind this research is to provide an overview of the dynamics behind Open Data as a starting point for future research on selected areas. Findings from this and continuative research can help other ecosystems to improve and thereby create a potentially better, but certainly a more open world.

CHAPTER 2

Literature Review

The literature review is divided in two main parts to guide this research. In the first part I cover Open Data as the subject of investigation. I explain the conceptual understanding and provide an overview on the different research fronts. In the second part I derive the theoretical framework that will be used to investigate the British Open Data ecosystem. The framework is based within the theory of business ecosystems and is influenced by the theory of complex systems in social sciences and evolutionary economics.

2.1 Open Government Data

2.1.1 Conceptual Understanding

The Open Knowledge Foundation (OKF) defines Open Data as data that is free to use, reuse, and redistribute – “[...] subject only, at most, to the requirement to attribute and/or share-alike” (Open Knowledge Foundation, 2013). Openness is hereby defined from two perspectives: technical and philosophical-legal openness. Data is technically open when it comes as non-textual raw data and in an open format for platform-independent processing. These formats are not restricted to any proprietary products and can be viewed and processed with Open Source software. Table 2.1 shows a selection of data formats and indicates their structural properties in terms of openness. Philosophical openness describes the democratic access to data, which is determined by licenses and the terms of use. The application of open licenses is of particular importance for the commercial use of Open Data. One of the most renowned content license models are the Creative Commons (CC) copyright licenses. These licenses offer different terms of use and are already widespread and highly comprehensible (Kloiber, 2012; Yu & Robinson, 2012). A central idea of Open Data is the aspect of interoperability. Interoperability means the possibility to work with and combine data sets from different sources

2 Literature Review

Data Formats	Machine Readable	Specifications Available	Open Format
Text (.txt)	yes	yes	yes
Comma Separated Value (.csv)	yes	yes	yes
Hypertext Markup Language (.html)	yes	yes	yes
Extensible Markup Language (.xml)	yes	yes	yes
Resource Description Framework (.rdf)	yes	yes	yes
Open Document Formats (.odt, .ods)	yes	yes	yes
Newsfeed/Webfeed Syndication (.rss)	yes	yes	yes
Portable Document Format(.pdf)	no	yes	yes
Microsoft Word (.doc, .docx)	no	yes	no
Microsoft Excel (.xls, .xlsx)	yes	yes	no
Microsoft Rich Text Format (.rtf)	yes	yes	no
Graphics Interchange Format (.gif)	no	yes	yes
JPEG format (.jpg, .jpeg)	no	yes	yes
Portable Network Graphics (.png)	no	yes	yes
Tagged Image File format (.tiff, .tif)	no	yes	no
Geography Markup Language (.gml)	yes	yes	yes
GPS Exchange Format (.gpx)	yes	yes	yes
Autodesk Drawing Format (.dwg)	yes	no	no

Table 2.1: Different Types of Data Formats and Structural Properties
(cf. Geiger & von Lucke, 2011)

in different systems. Technical and philosophical-legal openness is necessary to secure this interoperability (Open Knowledge Foundation, 2013). In 2007, an Open Government working group defined eight principles of Open Data that sum up the given explanation: Open Data has to be complete, primary, timely, accessible, machine-readable, available in non-proprietary formats under open license and the access must be non-discriminatory, which means it must be possible to access the data as an anonymous user (Tauberer, 2007).

As Yu & Robinson (2012) explain, the term Open Data remains neutral in regards to the content of the data sets and only describes their technical and legal shape. Kloiber (2012), however, mentions that in the majority of articles, report and strategy papers the term is used synonymously for Open Government Data (OGD). This dissertation refers to Open Data as data sets from public bodies that are made accessible to the public under the principles defined above. Due to the complex nature of public administration, this includes a wide range of different data categories. The Open Knowledge Foundation identifies eight categories: geo-data, culture, science, finance, statistics, weather, environment and transport (Open Knowledge Foundation, 2013). Confidential data sets and data sets that contain any kind of private information are excluded from the ambition to be published as Open Data. This constraint is crucial and widely misunderstood by the general public.

2 Literature Review

Governmental efforts to promote Open Data are framed as Open Data initiatives. They generally work towards three major objectives: Increasing the accountability and transparency of the government and the administration, strengthening civic participation, and unlocking economic potential through data-driven innovation and efficiency (Kloiber, 2012). However, not all categories of Open Data are useful for the completion of each of these goals. They can be distinguished by their societal or economic potential. Yu & Robinson (2012) classify data categories on an axis between the two poles "service delivery" and "public accountability". Although most of the data sets may serve both purposes, this categorisation helps to evaluate the underlying intention of data disclosures. Data sets that are categorised as "service delivery", like real-time data of public transport or geodata, can be used to develop public service applications or products with a marketable business model.

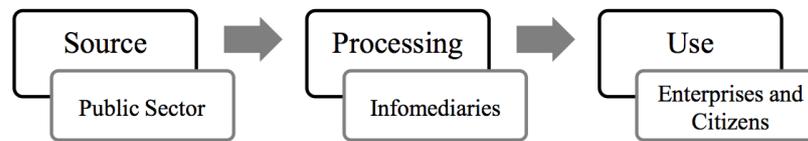


Figure 2.1: Open Data Value Chain
(cf. Rojas et al., 2013)

Data about the federal budget or the financing of political parties can serve the purpose of "public accountability". Edited and visualised by infomediaries, like journalists or developers, this data increases the accountability of the government and the public administration and therefore strengthens the democracy. The belief that a more direct access to data and information leads to an improvement of democratic processes is an essential part within the societal discourse around Open Data and Open Government. Bertot et al. (2010) additionally argues that the disclosure of this kind of data is likely to curtail nepotism and the abuse of power. Fig. 2.1 shows a simplified Open Data value chain. The data sets are transferred from its source, the public sector, through infomediaries that process the data and thereby increase its value, to the end-users. These end-users can be individual citizens as well as enterprises or other organisations. In chapter 3.3 I will show that this model is not sufficient and provide an improved version.

2.1.2 Research Approaches

Open Data initiatives are taking place for a relatively short amount of time. However, several different research directions have evolved. Davies et al. (2013) suggest to classify research on Open Data into three broad groups, which can be tackled under different research paradigms and from different subjects. This tri-partite classification, however, focuses solely on the social sciences and has to be enhanced by more technical areas of research, like computer science or application design. In this section I will outline the different approaches, give examples and classify this dissertation within the research landscape.

2 Literature Review

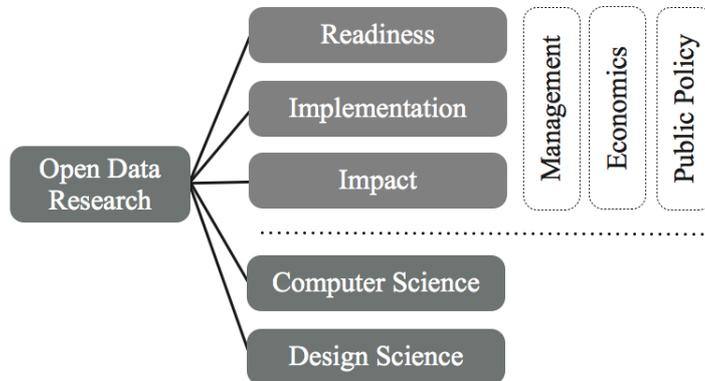


Figure 2.2: Open Data Research Fronts
(cf. Davies et al., 2013)

Davies et al. (2013) subdivides social science research on Open Data into readiness assessment, implementation studies and impact studies. Readiness assessments investigate whether a geographical region, be it a town, state, or nation, meets the preconditions to implement an Open Data initiative. These studies have a distinct set of assumptions of what constitutes an effective Open Data initiative and make recommendations on what kind of improvements in the infrastructure are necessary. Recent implementation studies have all benchmarked on Open Data initiatives in developed countries like the UK or US (cf. Grewal et al., 2011; Stott & Kaplan, 2013; CIPESA & APC, 2012). Implementation studies follow a more objective and less normative approach than readiness studies. They investigate whether and how Open Data policies are implemented and cover a wide array of indicators. Some measure the availability of data sets, indicate the proportion of organisations publishing their data openly or look at the quality of the published data. A comprehensive list of research projects can be found with Davies et al. (2013). Existing case studies on Open Data in developing countries historically describe the implementation of Open Data initiatives and highlight, which obstacles occurred. Two of these initiatives have been implemented in Kenya and Moldova. Both were sponsored by the World Bank, which investigates Open Data as a tool for economic and political development according to the ICT4D (Information and Communication Technology for Development) approach. (cf. Rahemtulla et al., 2011, 2012; Majeed, 2012; Davies & Edwards, 2012).

Open Data promises benefits from two directions. On the one side, the release of data sets under an open license enables entrepreneurs to develop new commercial services and business models. On the other side, advocates of Open Data expect democratic empowerment and civic participation. Impact studies directly assess these promises. According to Kloiber (2012) countries with progressive data policies like France and the UK already claim to observe positive economic effects. An EU commissioned research on the use of Open Data estimated the increase of business activity as EUR 40 billion per year. The indirect benefits

2 Literature Review

(benefits for people who use the services) are estimated on EUR 140 billion per year (Vickery, 2011). A Finnish geodata study claims that SMEs grow by 15%, if required geodata is free of charge (Koski, 2011). In Canada, Open Data helped to expose a case of misuse of charity status, which cost the taxpayer CAN\$ 3.2 billion (Eaves, 2012). Nonetheless, as Davies et al. (2013) points out, so far there has only been a selection of isolated case studies (Davies, 2010; Hammell et al., 2011; Tong et al., 2013) or vague economic estimations that extrapolate on single use cases, but no comprehensive large-scale studies. They describe two paths how future studies could address this deficit: Studies from a macro perspective could assess correlations between measures of general openness and economic key indicators that may capture the positive effects of Open Data. Studies from a micro perspective could investigate the dynamics of Open Data use on a granular level, for example the use of a unique data set through different access points or by different organisations.

Figure 2.2 shows that readiness, implementation and impact can serve as research paradigms for different fields within the social sciences. Management studies might assess how Open Data affects organisational structures, innovation processes and accountability within an enterprise. Economists are able to derive quantitative models, which capture the impact of Open Data and research on public policy can investigate how digital openness affects the requirements for public administration. Besides these approaches from the social sciences, technical research plays a major role in the development of Open Data ecosystems. In the UK the idea of Open Data partly emerged from efforts around semantic web technologies. Technological researchers argue that putting data on the web openly is necessary but not sufficient to unlock its full potential. To transform “data on the web” into “the web of data”, the concept of Linked Open Data (LOD) is promoted (Alani et al., 2007; Berners-Lee, 2009; Böhm et al., 2010; Hoxha & Brahaj, 2011). The term LOD describes a set of best practices, which have evolved in recent years and have served to publish and connect information online (Bizer et al., 2009). To promote a unique standard throughout the web, Berners-Lee (2011) developed a five-star rating system for the quality of LOD including, for example, RDF triples and SPARQL queries as standards. This five-star rating system now globally functions as a fundamental principle of Open Data and has been promoted by all kinds of stakeholders. In comparison to social science research, technological work rather follows the paradigm of action research. This means that research proceeds through pilot projects in the field, which intend to make immediate practical use of Open Data. To enhance the citizens’ use of Open Data services, technology action research is paired with research on human-computer interaction and application design. In her work Kloiber (2012) describes several real-life applications of Open Data from the UK, US and Germany that would fall in the category of technological action research. However the line between academic, knowledge driven action research and non-academic civil activism is blurred.

This dissertation research can be regarded as a cross-cutting approach to study the implementation of the British Open Data initiative and combines viewpoints from management science, public policy and technological history. With this work I pursue a descriptive approach and generalise on the emergence on Open Data ecosystems. The conducted research serves as an entry point for future research on Open Data ecosystems.

2.2 Business Ecosystems

From a manifold array of options to describe a network of organisations, I identified business ecosystem theory as the most appropriate one to describe the dynamic network of Open Data stakeholders. In this section I give an overview on business ecosystem theory and its intersections to complexity theory and evolutionary economics.

2.2.1 Ecosystem Analogies

The analogy between biological and economic ecosystems has been stressed in academic literature various times. Hannon (1997) states that ecology and economics share many common features, as both study dynamic systems that incorporate methods of production, exchange, capital stocks and storage. According to Lewin (1999), "biological ecosystems and economic systems are complex adaptive systems and thus follow the same deep laws". Furthermore, Lewin elaborates on the importance to understand economy as an ecosystem to "grasp fundamental truths about what makes the economy work". The analogy between economy and ecology seems plausible but also shows some obvious limitations. In business ecosystems the actors are intelligent, thus able to plan and predict the future to some extent (Iansiti & Levien, 2004). Lewin (1999) calls this distinction the ability of people to make conscious decisions. Another weak point of this analogy is that business ecosystems act upon the maxim to constantly deliver innovation, whereby natural ecosystems only exist to survive (Iansiti & Levien, 2004).

Peltoniemi & Vuori (2004) review existing ecosystem analogies, for example the industrial ecosystem, the economy as an ecosystem approach, the digital business ecosystem and the social ecosystem. At the end of their conceptual analysis they derive a definition of a business ecosystem: "As a conclusive definition we consider a business ecosystem to be a dynamic structure which consists of an interconnected population of organisations. These organisations can be small firms, large corporations, universities, research centres, public sector organisations, and other parties which influence the system." This definition seems to be the most accurate one at hand and will be used within this dissertation. In subsequent work Peltoniemi (2006) identifies some characteristics that are components of every business ecosystem. Business ecosystems consist of a large number of loosely connected participants that interact in both constellations, in cooperation as well as in competition. In other words members of a business ecosystem "depend on each other for their mutual effectiveness" (Iansiti & Levien, 2004). They can form alliances to be more efficient but could die despite their best efforts when changes in another part of the network are propagated throughout the system: They have a shared fate (Lewin, 1999).

2.2.2 Complexity and Business Ecosystems

In this section I cover three areas of complexity theory and explain how they interfere with business ecosystem theory. The three areas are coevolution, self-organisation and emergence.

In a general understanding evolution, and thereby coevolution, is a term with a positive connotation. When things evolve they become better. Evolutionary theory however does not include such beneficial implications by default. Lenway & Murtha (2004) rather describes evolution as a “cumulative and transmissible change”. Coevolution as a property of complex systems is present when “the change in fitness of one system changes the fitness of another system, and vice versa. Coevolution is the evolutionary mutual changes of species (or organisations) that interact with each other” (Merry, 1999). Lenway & Murtha (2004) state that coevolution takes place when the participants “have a significant causal impact on each other’s ability to persist”. Pagie (1999) discusses three types of coevolution from a biological perspective: competitive, mutualistic and exploitative coevolution. These types can be transferred to the coevolutionary process happening within business ecosystems. “Competitive coevolution occurs between species, which are limited by the same resources” (Pagie, 1999). In a business ecosystem a price war is an example of competitive coevolution, as it effects all competing organisations, while fighting for the customers spending as a scarce resource. Competitors are forced to use or acquire the resource more efficiently. Biologists frame this the red queen effect (Peltoniemi, 2006). In mutualistic coevolution all participants benefit from structural changes. The development of the hardware and software market can be seen as a mutualistic coevolution. The products complement each other and together trigger an increasing market volume. Finally, in exploitative coevolution none of the participants benefits from the interaction (Peltoniemi, 2006). This can be observed between large corporations and their suppliers.

Anderson (1999) describes self-organisation as “a process where pattern and regularity emerge without the intervention of a central controller”. This is quite close to Goldstein’s idea of a process “whereby new emergent structures, patterns, and properties arise without being externally imposed on the system” (in Choi et al., 2001). Business ecosystems adjust without an external controller and therefore satisfy the definition of self-organisation in a broad sense. In comparison to a centrally planned economy, the market economy system enables decentralised decision-making. However, in real life business ecosystems and especially in ecosystems developing around Open Data there are obvious interventions by the public sector. The participation of publicly funded bodies put the attribute of self-organisation in question. These interventions can either be regarded as inhibiting self-organisation or as supporting the macro-structures to make self-organisation possible (Peltoniemi, 2006).

Phan (2004) describes emergence as “a property of a complex adaptive system that is not contained in the property of its parts”. More specific, Smith & Stacey (1997) argue that emergence “means that the links between individual agent actions and the long-term systemic outcome are unpredictable”. According to them “emergence confuses the links between cause and effect, which makes it impossible for one actor to control the whole system.” For business ecosystems it thereby becomes clear that emergence is the link between micro (the individual agents performance) and macro behaviour (the ecosystems economic performance), which is

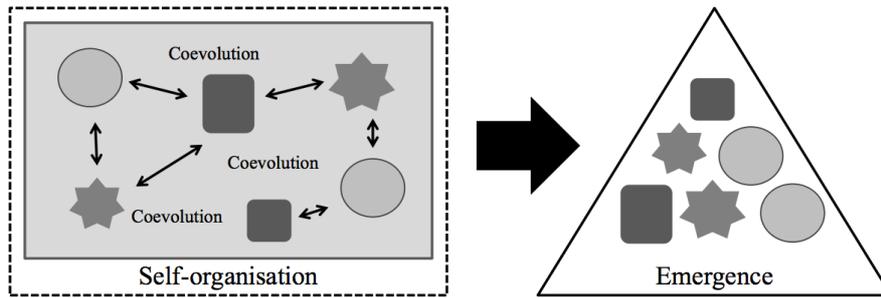


Figure 2.3: Coevolution, Self-organisation and Emergence
(cf. Peltoniemi, 2006)

not strictly summative. Peltoniemi (2006) concludes that in business ecosystems "emergence is a phenomenon that arises from organisation-level motives and actions that lead to unpredictable and even surprising population-level behavior". The connection between coevolution, self-organisation and emergence within business ecosystems is visualised in figure 2.3.

2.2.3 Evolutionary Economics and Business Ecosystems

Since the early 1980s economists make use of Darwin's theory of biological evolution to describe industry structures. Although this branch has found numerous advocates, there are some doubts as well. Hodgson (1998) argues that "if economic development is determined by some process of natural selection, with something analogous to genetic replication and to random variation or mutation, then what role remains for the notions of intentionality, purposefulness or choice, which economists of many schools of thought have held so dear?". However, as explained in section 2.2.1, evolution in socio-economic systems can be studied without overstretching the ecological analogy. This dissertation points at evolutionary economics as a way to explain the evolution of technology-driven business ecosystems.

As pioneers of evolutionary economics Nelson & Winter (1982) explain that this approach emphasises "the tendency of the most profitable firms to drive the less profitable ones out of business." According to their theory, firms "are modelled as simply having, at any given time, certain capabilities and decision rules." These capabilities and rules change over time due to deliberate efforts and random events. The market determines which firms are profitable and which are not. The latter are excluded from the market. This mechanism can be seen as the economic analogue to the natural selection process (Nelson & Winter, 1982). In ecological ecosystems genes determine a creature's physical and mental condition and therefore its capability to adapt to the environment. In the evolutionary theory of Nelson & Winter (1982) organisational routines play the role of genes. Therefore, routines as well as the environment has an effect on an organisation's performance. When the organisation does not perform sufficiently it is likely to be excluded from the population. However, Hodgson (1998) remarks that

2 Literature Review

“it is widely, but wrongly, assumed that evolutionary processes lead generally in the direction of optimality and efficiency.” According to him, natural selection, whether in ecology or economy, does not lead to the superlative fittest but to the tolerably fit. Again this can be found with Nelson & Winter (1982), who argue that firms in the market are engaged in “profit seeking” or “profit-motivated striving” but not necessarily in “profit maximisation”. Foster & Metcalfe (2001) state that the members of a population must show variation in their characteristics to enable meaningful selection. They argue that in an economic context this variation is derived from innovations in products, means of production or routines. Again it is important to bear in mind that ‘meaningful’ in this context does not necessarily describes an optimal process in terms of fitness or economic outcome. At this point the intersection to complexity theory as explained in section 2.2.2 becomes clear. In this section I pointed at the importance of routines and decision rules when investigating business ecosystems. Furthermore, I explained that besides routines as a factor, market participants get selected depending on their fit of innovation and environmental change.

2.2.4 Research Framework

The diverse theoretical underpinnings have to be subsumed in a framework to guide meaningful research. A comprehensive framework is able to embrace the underlying theory and to break it down in key variables. With this dissertation I lean on the integrated business ecosystem framework described by Peltoniemi (2006).

Business ecosystems consist of micro and macro structures that develop in reciprocal cycles. Changes within the individual organisations influence the performance of the ecosystem on a macro level. These changes on the macro-level however force individual organisations to adapt. As figure 2.4 shows, the two main themes for the investigation of business ecosystems are therefore feedback and conscious choice. Selection, coevolution and emergence relate to the feedback mechanisms within an ecosystem. For example the selective pressure an organisation encounters is likely to be the echo of its own actions, which have shaped the other organisations behaviour. Conscious choice is the other key variable within the framework. Due to the conscious choice of organisations, the evolutionary process is not just mere random but can be meaningful. Without the ability to think ahead and remember the past, organisations would purely react to immediate triggers from their environment. "Conscious choice enables the organisation to think with a longer perspective, to integrate new knowledge with the old knowledge, to try to guess a competitor's next move, to create tactics and even to try to cheat" (Peltoniemi, 2006). Conscious choice therefore differentiates the business ecosystem from the ecological ecosystems and leads to dynamics of non-linear cooperation and competition. Within organisations it is embodied in leading individuals like founders or directors. Therefore, the analysis will not only cover organisations but some key individuals within the ecosystem as well. Summing up the combination of conscious decisions under limited knowledge and the interconnected dynamics of feedback mechanisms results in a nondeterministic, nonlinear and unpredictable future constructed by organisations, which can be investigated under the research framework at hand (Peltoniemi, 2006).

2 Literature Review

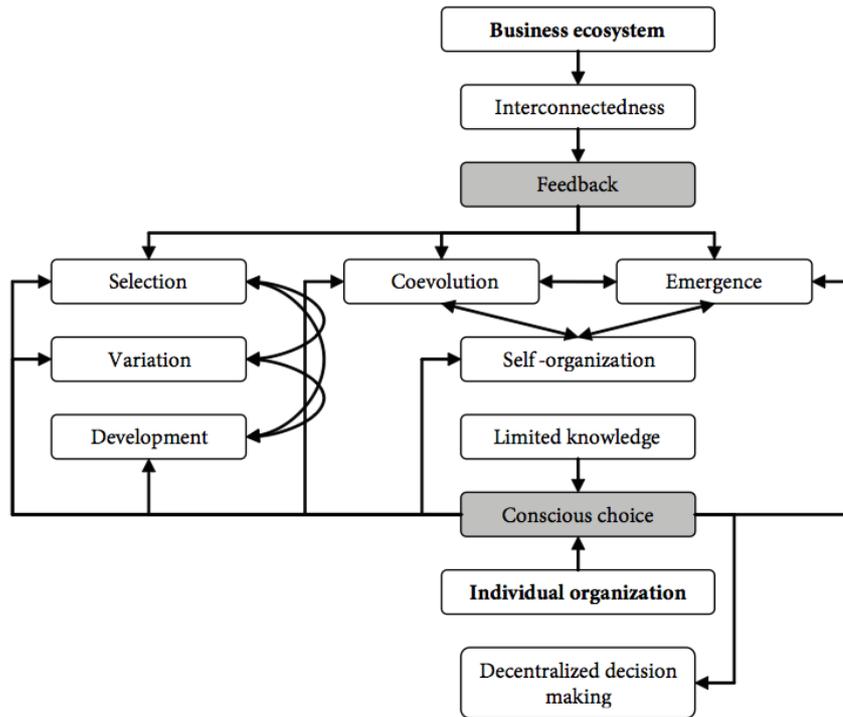


Figure 2.4: Integrated Business Ecosystem Framework (Peltoniemi, 2006)

CHAPTER 3

Methodology and Research Design

In this chapter I define the research questions and describe the methodical approach to answer them. In the first section I provide a working definition of Open Data ecosystems, a primary research question and a set of subordinate questions. In the second section I justify the chosen case study approach and explain why I investigate the British Open Data ecosystem as a case. In the third section I formulate a structural taxonomy of ecosystem participants to provide a framework for data collection. In the final section I describe the process of data collection and the range of data I gathered.

3.1 Research Objective

In the literature review I give insights to the foundations of Open Data and the theory of business ecosystems. In this dissertation I study business ecosystems that evolve through the use of Open Data and are therefore named Open Data ecosystems. Based on the literature, I derived the following working definition of Open Data ecosystems to formulate a clear research question:

An Open Data ecosystem is a dynamic structure, which consists of an interconnected population of conscious agents, who either directly produce and process data sets that are technically and legally open, or indirectly enable these processes. These agents are exposed to various feedback processes and can be all forms of governmental bodies, small firms, large corporations, universities, research centres, public sector organisations, individuals or other parties, which influence the system. Open Data ecosystems are limited by the geographical and constitutional area, which the involved data sets refer to. Therefore Open Data ecosystems can exist on a municipal, state, federal, supranational and global level.

3 Methodology and Research Design

The definition acknowledges that agents within Open Data ecosystems are interconnected, but it does not include their mutual dependency. As Open Data ecosystems include public bodies, which are not necessarily subject to market economy mechanisms, their dependence on other private sector agents is in doubt. The same applies to the assumption that Open Data ecosystems are self-organising. According to existing literature it remains unclear whether the participation of governmental bodies within a business ecosystem inhibits self-organisation or supports the macro-structures to make self-organisation possible (Peltoniemi, 2006). With this working definition at hand the primary research question is:

- How did the network of public, private and formalised non-governmental actors shape the development of the British Open Data ecosystem?

Three subordinate research questions tackle subdomains of the primary question and guide the research:

- Which are the focal stakeholders that act within the British Open Data ecosystem?
- What are their structures, interests and actions?
- How does the appearance of the ODI influence micro- and macro structures within the ecosystem?

3.2 Case Study Approach

Saunders et al. (2011) describe the way to develop a research methodology as an analogy to the researcher peeling an onion, where different layers have to be detached consecutively – from the research philosophy as the outer layer towards the techniques and procedures as the core of the onion. I play on this analogy to describe my research approach. This dissertation follows a pragmatic research approach and thereby regards the research question as the most important determinant of the epistemology, ontology and axiology (Saunders et al., 2011). Tashakkori & Teddlie (1998) argue that a pragmatic approach allows the researcher to focus on his topic of interest without engaging in debates about concepts like truth and reality. As the literature review showed, the existing literature about Open Data ecosystems is limited and therefore this research follows a rather inductive strategy. Data is collected to understand the nature of the problem at hand and to derive some theoretical findings. Easterby-Smith et al. (2008) argue that researchers following an inductive approach "are more likely to work with qualitative data and to use a variety of methods in order to establish different views of a phenomenon".

To answer the research question I use a case study approach. Robson (2002) defines the case study approach as "a strategy for doing research which involves an empirical investigation of a particular contemporary phenomenon within its real life context using multiple sources of evidence". Flyvbjerg (2006) explains that case studies complement rule-based context-independent knowledge with the creation of context-dependent knowledge, which is necessary to develop expert knowledge within qualitative science. This is aligned with Eysenck

3 Methodology and Research Design

(1976) stating that "sometimes we simply have to keep our eyes open and look carefully at individual cases – not in the hope of proving anything, but rather in the hope of learning something".

This project studies the British Open Data ecosystem as a single and embedded base, which means that the whole ecosystem is treated as one case but consists of several sub-units – the different stakeholders – which are analysed separately. The case study uses the British Open Data ecosystem as an "extreme case" (Flyvbjerg, 2006), as it is one of the most developed ecosystems of this kind worldwide. Interviews with stakeholders of the ecosystem form an integral part of the case study. The major stakeholder categories have been identified through the background literature and informal discussions with experts at the beginning of the field research. The selection of interview partners was done by purposive sampling, which enabled me to select interview partners that provided the best input to answer the research question. This form of sampling is common when the population of potential interview partners is fairly small and when the researcher wishes to select partners that are particularly informative (Neuman, 2005). Despite demonstrated advantages, Flyvbjerg (2006) also highlights some weaknesses of the approach. As the researcher is an integral part of the research there might be a selection bias when collecting the data that may overstate or understate relationships. Furthermore, the statistical significance of phenomena within a case study might often be unknown or unclear. However, the above listed advantages clearly outweigh the limitations.

3.3 Structural Taxonomy of Ecosystem Participants

Similar to the biological flora and fauna, participants in the business ecosystem can be divided into different species. This taxonomy of participants not only helps to plan a sound data collection, but moreover allows to draw generalised conclusions from the data afterwards. According to the definition from section 3.1, participants of an Open Data ecosystem produce and process data sets directly and indirectly. Figure 3.1 shows the schematic composition of the ecosystem.

The central value chain, as discussed in section 2.1.1, leads from data suppliers over data intermediaries towards the final data consumers. In this case study data suppliers are all public bodies that produce valuable data sets. This includes institutions, which are entirely tax-funded, like ministries, as well as institutions that generate most of their revenue from selling data to the government, like Trading Funds in the UK. Intermediaries are all kind of organisations that process raw data to enhance its value. This can be for example data journalists, who visualise data sets and therefore make their information content available to the broad public. However, journalists are only one handy example for this group. Generally, all individuals and especially businesses that add value to data by processing it and making it available for reuse, whether commercially or for free, belong to the category of data intermediaries. A standard value proposition for an intermediary might be to match geodata with data sets from other categories and offer the derived map as a service. Data consumers therefore are individuals or organisations, which consume data to extract information. This can be the single user consuming a geodata tool, a business which commissions an intermediary to combine its

3 Methodology and Research Design

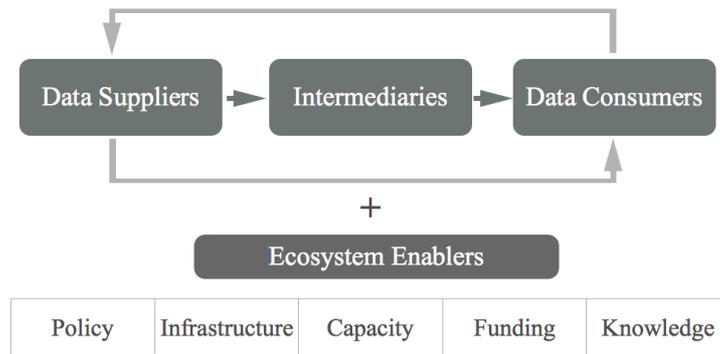


Figure 3.1: Open Data Ecosystem Taxonomy

closed data with Open Data, or the government itself. This last case is indicated as a feedback loop in figure 3.1. When governmental bodies open up their data sets, which will then get processed and enhanced, the body is likely to get hold of the processed data to learn more about its own organisation. Figure 3.1 also indicates a shortcut from data suppliers towards data consumers. Through FOI requests or portals like data.gov.uk end users can access raw data directly. However, from a practical point of view only a fractional amount of citizens in the UK are able to derive insight from the raw data by themselves, due to a lack of data literacy.

Besides the central value chain the ecosystem consists of focal stakeholders, which do not necessarily interact with the actual data sets in any way. For my investigation I subdivided this group of ecosystem enablers into five categories, whereby single stakeholders can take up more than one of the five depicted characteristics. Stakeholders like the Open Data Institute or the Open Knowledge Foundation are engaged in policy making around information and transparency legislature and thereby crucially influences other market participants. At the same time the ODI creates capacity within the ecosystem, as it offers seminars and even intends to offer a postgraduate diploma. Philanthropic investors, like Omidyar Network or the Open Society Foundations (OSF), offer grant funding for Open Data projects within the UK and even the first for-profit investors have funded Open Data startups. The supply of funds is crucial for the survival and growth of the ecosystem. Working with Open Data in most cases means working with large data sets. Therefore initiatives like the Open Source Software project Apache Hadoop are important elements of the ecosystem, because they produce the technical infrastructure. Lastly, the mere creation of knowledge around the ecosystem enables its development as this makes it more accessible for potential participants. Universities are the main bodies to generate well-grounded knowledge, however institutions like the Finnish Institute in London also create knowledge by sponsoring publications on the digital commons. These examples illustrate the class of ecosystem enablers, which I introduced to enhance the classical value chain. For this research I met interview partners from all parts of the ecosystem as schematised in figure 3.1. A list of interview partners can be found in appendix A2.

3.4 Case Study Database

In the final section of this chapter I describe the data that I collected in London for three weeks in July 2013. I conducted semi-structured interviews and broadly followed an interview manual that was designed to answer the research questions for each stakeholder. The manual can be found in appendix A6. However, during most interviews the questions have been rephrased and adapted to the individual interviewee to improve the quality of the responses. I started the interviews with broad questions to inspire narration. This urge for narration is helpful to decrease ex-post-rationalisation of the interview partners (Schütze, 1983). Afterwards I followed up with some more specific questions to test for my key variables. For example, I asked the interviewees for pivotal events that influenced the development of the ecosystem to identify cases of extreme selection pressure in the ecosystem. As another example I asked for the interviewees co-operation partners and competitors in the market to find out more about the effect of conscious choice in this relationships. Within a data collection period of three weeks I conducted 15 interviews. The interviews lasted between 15 and 45 minutes each with an average duration of 25 minutes.

Table 3.1: Case Database

Case	ODI	External	Total
Interviews	7	8	15
Events	4	3	7
Total	11	11	22

Most of the interviews have been transcribed verbatim. An exemplary transcribed interview can be found in appendix A4. Some of the interviews within the ODI had the character of a briefing and were therefore written up in a third-person summary style. They rather serve exploratory purposes. The interviews were recorded and transcribed using the tool ExpressScribe. One interview was conducted via phone and recorded using Skype Call Recorder. In total, I gathered around 16,000 words of transcribed interview data. All audio recordings and raw interview transcriptions are stored on the researchers computer, as well as on a cloud-based service. All interviews have a unique identification code to simplify the referencing. A table explaining the identification code logic is attached in appendix A1. One of the interviews is on request anonymised down to the organisation's name.

My second source of data are events. Events within the Open Data ecosystem serve as meeting point of stakeholders and have been a crucial driver for its development. I took part in six events in London and one in Edinburgh. The events ranged from quite informal gatherings like the OKF Maker Night, an event to sit together and develop apps, to very formal events like the roundtable discussion of the Foundation for Science and Technology, mostly attended by senior academics and officials from the public sector. According to ethnographical

3 Methodology and Research Design

standards I always disclosed my identity as an academic researcher. Based on observation notes from these events I could draw conclusions about the relationship between participating stakeholders and the mechanics behind these events as meeting points within the ecosystem. The distribution of interviews and events is shown in table 3.1. A list of events and interview partners can be found in the appendices A2 and A3. During the discussions, a lot of interview partners referred to documents as milestones for the development of the ecosystem. Most of these recommendations have been included in the historical overview of the ecosystem in chapter 4.1.

CHAPTER 4

Analysis

In this chapter I analyse the empirical data according to the research framework and the research questions. In the first section I describe the development of the British Open Data ecosystem from a macro perspective and over time. The chapter relies on different existing timelines, background research and confirmation by several interview partners, who played key roles in the development themselves. Media reports or other official documents provide empirical validity. In the second section I analyse the dynamics of the ecosystem from a micro perspective and is partitioned by the stakeholders' main interest. Therefore the interviews have been transferred into tables for qualitative coding. I coded the interviews according to the research framework and provided interpretation based on my experiences and observations in the field. An exemplary coding table can be found in appendix A5. The amount of data gathered from different techniques allowed a certain degree of data triangulation and was sufficient to derive some generalised findings.

4.1 Macro Analysis: Timeline of the Ecosystem

The emergence of the British Open Data ecosystem has started long before the term itself was coined. Since the late 1990s initiatives to unlock the potential of Public Sector Information have been driven by a community of activists and civil servants. Over the last years Open Data in the UK transformed from a rather small movement to an issue of national interest, where different parties try to shape it between the rationales of economic growth and civic empowerment. This section describes the development of the ecosystem from a macro perspective and analyses the different streams that sometimes complement and sometimes oppose one another.

4 Analysis

A licensing policy that allows the commercial reuse of Open Data lies at the heart of every initiatives economic potential. In the UK the 1998 green paper “Crown Copyright in the Information Age” initiated a liberalisation process crucial for todays ecosystem. As one future scenario the paper suggested the abolishment of crown copyright, but when in 2000 the Office of Public Sector Information took over the regulation of crown copyright licensing from HMSO, they did not abolish copyright licensing completely, but introduced the PSI Click-Use licensing scheme¹. This transactional but free class license allowed the commercial and non-commercial use of crown copyright material under the precondition of attribution. Three years later, in April 2003, Douglas Alexander, Minister for the Cabinet Office, established the Advisory Panel on Public Sector Information (APPSI) as a non-departmental public body to advise the government on the reuse of Public Sector Information². Later that year, in November, the European Union adopted the “Directive on the reuse of Public Sector Information” to create a common legislative framework how public bodies across the EU should make their PSI data available³. In January 2005 the Freedom of Information Act, which got the royal assent in 2000, came into force giving citizens, businesses and especially journalists the legal basis to directly request a broad range of information held by public authorities. When in November of the same year the PSI Directive was implemented on a national level, two strong preconditions for today’s Open Data ecosystem interlocked.

Back in 2004 Rufus Pollock, along with others, incorporated the Open Knowledge Foundation, a non-profit organisation to develop tools and communities around different aspects of digital information. In October 2005 the OKF organised the World Summit on Free Information Infrastructure, which in following years turned into the annual OK Conference. Shortly after, they released the first draft of the Open Knowledge Definition to define “openness” in relation to data and content⁴⁵. Besides the efforts of rethinking the use of PSI from the governmental side, this indicated the increasing formalisation of civil activism. In March 2006 the Guardian newspaper’s technology section launched the long-lasting “Free our Data” campaign, lead by the journalists Michael Cross and Charles Arthur. They argued that government-funded agencies like the Ordnance Survey or the UK Hydrographic Office should provide the citizens easy access to their data, as the taxpayer funds the data collection anyways⁶. The campaign succeeded when in April 2010 Ordnance Survey, under the label of OS OpenData, released important geodata for reuse⁷. In line with the beginning of this campaign, the Office of Fair Trading published the report “The commercial use of public information” in December 2006 to investigate whether the Trading Fund model was skewing the market⁸. Geospatial data as produced by the Ordnance Survey stayed in the focus of attention when in May 2007 the European Union’s INSPIRE directive came into force, which imposes national agencies on developing processes and routines of sharing certain geospatial data sets⁹.

In June 2007 Tom Steinberg and Ed Mayo together with the Cabinet Office published the Power of Information Review taking a “practical look at the use and development of citizen and state-generated information in the UK”¹⁰. In reaction to a list of recommendations given in the report, the UK government established the Power of Information (POI) Task Force in March 2008. The POI Task Force consisted of representatives from business, civil activism and the Cabinet Office with the goal to investigate Steinberg’s and Mayo’s recommendations. At the same time Cambridge Professor David Newbery, together with Rufus Pollock, pub-

4 Analysis

lished the study “Models of Public Sector Information via Trading Funds” criticising the bodies’ commoditisation of tax-funded data. In parallel, the government made the commitment that data sets held by trading funds will be made available as widely as possible for the use in downstream markets¹¹. This event can be seen as a cumulated result of the reports published by Newbery and the Office of Fair Trading as well as the Guardian’s ongoing campaign. Parts of the valuable Trading Fund data sets finally came into reach. Three months after the POI Task Force has been established they announced the competition “Show Us A Better Way” together with the Cabinet Office. The competition made large data sets of public information available to developers, which have been unavailable beforehand¹². One of the competitions winners was the budget visualisation application “Where Does My Money Go?”¹³ that launched in December 2009 and later became part of the OKF-run service OpenSpending¹⁴. Until today OpenSpending is one of the lighthouse projects when arguing about ways to foster transparency and government accountability using Open Data.

The discussion about the reuse of PSI noticeably intensified from early 2009 on. As Rufus Pollock stated “even the phrasing changed. We started talking about Open Government Data[...]” (OKF-130719-RP). In January 2009 Barack Obama assumed the office as President of the United States and on his first day in charge issued a memorandum on the Freedom of Information Act clearly committing to openness as a government doctrine¹⁵. Besides some beguiling political rhetoric, this memorandum clearly indicated the United State’s role as a future driver of digital innovation in government. When the POI Task Force published their final report in February 2009 the Cabinet Office immediately started with its implementation. One of the recommendations from the report was to create a single point of access for government data. Richard Stirling reports how this recommendation was approached:

“I was part of the secretariat for that task force. I sort of helped to draft some of the data sections of the report and then when the report was published, the task force was disbanded. I established a new unit within the Cabinet Office, staffed it up, got the resources, hired the staff and build the initial data.gov.uk.”(CO-130718-RS)

The British Open Data portal launched in a beta version in September 2009 with a call for developers and data sets in RDF format. Earlier that year, in February 2009, the OKF released the first version of their Open Database License (ODbL)¹⁶ and thereby set important groundwork for an international application of Open Data. Although copyright has been widely harmonised around the world, the legal situation for databases is not as clear. Databases in the US for example are not necessarily protectable by law, where in the EU they are. The ODbL, combined with an appropriate content license, allows the reuse of data sets under the paradigm of Open Data¹⁷. Already in May 2009 the US government launched the data portal, data.gov, containing 47 data sets¹⁸. In 2009 the British government appointed Tim Berners-Lee and Nigel Shadbolt as advisors to the government on how to open up government data¹⁹. Their ongoing lobbying in the following years will clearly shape the nations Open Data policy. Shortly after his appointment, Tim Berners-Lee officially launched data.gov.uk to the general public in January 2010²⁰.

4 Analysis

In the first half of 2010 Open Data entered the campaigns for the upcoming general election. In March the Labour Party as well as the Conservative Party framed their ideas of how information technology should shape Britain's future. David Cameron started off when he released the "Technology Manifesto" outlining his ideas on how to get the nation, as he put it, "out of Gordon Brown's recession". One of his key points was a legislative change towards a "Right for Government Data"²¹. Tom Steinberg, who previously worked on the POI Task Force, was one of the central figures in the creation of this strategy document²². In my interview Steinberg states: "I stopped advising [Labour] government on these issues [the POI Task Force] – and I gave lots of advise – when I was unable to get them to give citizens precise powers around data"(MS-130717-TS). Under Cameron he was able to frame more precise plans for these powers. Shortly after, on March 22nd, the incumbent Prime Minister Gordon Brown revealed the "National Digital Strategy" for his next mandate, including the creation of a Web Science Institute under the direction of Tim Berners-Lee and Nigel Shadbolt. In his press release on "Building Britain's digital future" Brown announced a funding of GBP 30 million for the institute to create a "modern Domesday Book for the 21st century"²³. However, the results of the general election crossed this plan and right after David Cameron assumed the office as new Prime Minister in May 2010, the department for Business, Innovation and Skills (BIS) cancelled the formerly announced Web Science Institute as part of Chancellor George Osborne's savings plans²⁴. Instead Cameron followed the ideas from the Technology Manifesto setting up a Transparency Board including the established data and transparency advocates Berners-Lee, Shadbolt, Steinberg and Pollock, as well as Minister of Cabinet Office Francis Maude²⁵. In June David Cameron send a first letter to the Cabinet Ministers demanding for practical implementation on the transparency agenda including central government and local government spending data, crime data and some data (e.g. positions and salaries) on key civil servants²⁶. As the idea of Open Data at that time had neither reached nor convinced all relevant data holders in public administration, one of the interview partners symbolically described the action as the "Prime Minister running around with a stick" to get the data out of the departments (ODI-130711-JT).

In September 2010 the British government released the Open Government License (OGL) and therefore took a big leap towards a prospering Open Data ecosystem²⁷. The OGL replaces the Click-Use License for central government works. On the data.gov.uk blog Nigel Shadbolt describes the new license:

*"Based on the world-leading Creative Commons family of licences, the new licence works in parallel with them and mirrors their Attribution Licence and the Open Data Commons Attribution Licence, whilst covering a broad range of information, including Crown Copyright, databases and source codes, and applying to the whole of the UK. It is also available in a machine-readable form so that people's tools can automatically know that it is open to use."*²⁸

4 Analysis

One of the interview partners described this change from a transactional to a non-transactional license as one of the focal points in the development of the whole British Open Data ecosystem (NA-130709-XY). In July 2011 David Cameron addressed the Cabinet Ministers in a second letter concerning his transparency agenda. This time he asked for the provision of key data on the NHS, education, criminal courts and transport²⁹. In September 2011 the governments of the US and Brazil launched the Open Government Partnership, an international initiative to promote transparency, civil participation and digital administration. As one of the eight founding members the UK released its first OGP National Action Plan, which from then on had to be continuously revised with the civil society³⁰. In April 2012 the UK took over the 18-month co-chairmanship from the US³¹ and in June 2013 launched a draft of the second OGP National Action Plan, titled “From Open Data to Open Government”³².

Although plans for the Web Science Institute were cancelled in 2010, interested parties kept on lobbying until in May 2012 the government released an implementation plan for the Open Data Institute³³ to be set up in London. The idea for a government funded institution to support the economic exploitation of Open Data was first explicitly mentioned by George Osborne in his autumn statement in November 2011³⁴. Similar to the initial plan around the Web Science Institute, Tim Berners-Lee and Nigel Shadbolt have been appointed as president and chairman of the ODI. The institute officially opened in November 2012.

In May 2012 Francis Maude, Minister for the Cabinet Office, appointed Heather Savory as the first chair of the Open Data User Group (ODUG), a committee to include the users’ perspective into the process of Open Data policy making. According to Jeni Tennison, member of the ODUG, it works like a “proxy” for businesses and the civil society (ODI-130711-JT)³⁵. The British government, respectively the Cabinet Office, continued consultations with data experts and policy makers and in June 2012 published the report “Unleashing the Potential – The Open Data White Paper” along with departmental Open Data strategies and an updated version of data.gov.uk³⁶. With this white paper the government demonstrated a growing internal subject-related capacity by addressing a number of quite specific topics that clearly exceeded generic political rhetoric. Two of the most striking points within the white paper were the passage concerning a developer engagement strategy and the release of the Postcode Address File (PAF). Although technically every citizen can access data sets from data portals, developers are the important link to turn this data into valuable applications. In the white paper the government expressed the need to focus more on this demand side of the data. The PAF produced by the Royal Mail is one of the most valuable data sets hold by the UK government and could be a resource for a large number of applications. Within the white paper the government acknowledged this value, but also points towards the unsettled economic trade-off for releasing this valuable asset free of charge. As Jeni Tennison, member of the ODUG, states, “a lot of energy has been put in this one [demanding the free access of the PAF] but it has not been successful. This data set is routinely requested, but the [HM] Treasury is not convinced of the economic benefit” (ODI-130711-JT).

The study “Open Growth”, released by the consulting company Deloitte in December 2012 tried to quantify the economic value of Open Data for the UK economy by using broad anecdotal and some statistical evidence³⁷. The report was conducted in collaboration with the

4 Analysis

Open Data Institute and formed an integral part of the widely received "Shakespeare Review of Public Sector Information" published in May 2013. The review was commissioned by the BIS jointly with Cabinet Office and conducted by Stephan Shakespeare, who at that time chaired the government's Data Strategy Board³⁸. The comprehensive report was received by the UK government as a foundation for future policy decisions on data and transparency, as reflected in the "Government Response to Shakespeare Review" published in June 2013³⁹. Although the report has been warmly welcomed in most of the public coverage, one of my interview partners criticised that it shows essential similarities to the 2004 APPSI annual report and therefore exhibits the reluctant development of tangible progress in the field (NA-130709-XY). In June 2013 the G8 leaders signed the Open Data charter and therefore agreed on five strategic principles to act on, concerning the policies around PSI⁴⁰.

Covering a timespan of 15 years, the historical analysis shows two major streams that influenced the development. On the one hand, civil activists pushed the idea of a better democracy through data access from a niche movement to a topic on the national agenda, by formalising and collaborating with governmental actors. On the other hand, Open Data entered the political agenda as a mechanisms for economic growth, but government was only able to create meaningful policies by engaging businesses and the civil society. Economic growth through Open Data directly constitutes governmental transparency, as more civic services and applications are released. However, a higher transparency and democratic empowerment leads to economic growth, as an intense civic audit detects economic potential that can be skimmed. This reciprocal effects show that both sides of the movement, the side interested in growth as well as the one interested in civic empowerment, work towards the same state of the ecosystem but with a different order of priorities.

4.2 Micro Analysis: Dynamics of the Ecosystem

In this section I take a closer look at the dynamics between individual stakeholders in the British Open Data ecosystem. I analyse their relevance, development and interaction with other actors. The analysis is based on the transcribed interviews and observations from the field. The transcriptions have been coded for motives and patterns in order to answer the research questions. As section 4.1 indicates, the Open Data movement takes place between two main motives: Economic growth and civic empowerment. This section investigates the stakeholders according to their predominant motive.

4.2.1 The Interest in Economic Growth

Open Data as a mechanism for economic growth is implemented as a top-down approach by the British government. Mandated by the Prime Minister himself, all public bodies are prompted to open up their data sets. Since the topic entered the political agenda, the Cabinet Office is in charge to facilitate this process and is therefore one of the focal stakeholders to look at. With funds provided by the Technology Strategy Board, the Cabinet Office facilitated the setup of the Open Data Institute. The purpose of this non-profit is to unlock the economic potential of Open Data in the UK and is another key element of this analysis. Finally, the

4 Analysis

transition from Open Data to economic growth is expected to be achieved by startups from the private sector. Therefore, this diverse community is the third main stakeholder within the growth perspective.

Under former administrations the Cabinet Office functioned as an accelerator for emerging topics on the political agenda. It developed the topics for a certain amount of time and then passed them along to an appropriate department. However, under the Cameron administration this behaviour slightly changed and even after half a decade of work, Open Data is still affiliated with the Cabinet Office. Under the Conservative's government, the Cabinet Office is "in the business of doing things [...] which it didn't used to be"(CO-130718-RS). Richard Stirling, who was contributing to the POI Task Force report and subsequently was responsible for the development of data.gov.uk, explains that the Cabinet Office's process of policy making around Open Data involved "bringing in the best ideas from the outside"(CO-130718-RS). These ideas not only derived from the circle around the senior academics Berners-Lee and Shadbolt, but particularly from civic activists like Rufus Pollock or Tom Steinberg. Data.gov.uk serves as a tangible example of this collaboration, as it has always been based on CKAN, a data management system developed by the Open Knowledge Foundation. The Cabinet Office consulted Pollock throughout the whole process of development, from the initial POI Task Force report to the implementation of CKAN as an underlying structure for data.gov.uk. During the interviews both, Stirling and Pollock, referred to a specific informal meeting where they "co-designed" the data portal (CO-130718-RS, OKF-130719-RP). Besides this specific collaboration, the Cabinet Office constantly communicated with the national community of developers to work towards their requirements on the one hand and to spread the word on the other hand. In absence of a formalised association of developers as a single point of contact, members of the Cabinet Office approached the developers at diverse meet-ups and universities:

"There was a period where twice a week, maybe sometimes five days a week, I went to developer meet-ups. And I would speak to an audience of between 50 and 250 developers about the things the government is doing to open up their data [...]. Somebody that worked for me [...] gave talks at undergraduate courses on computer science at about eight different campuses. We did all of these things to try to get the message out." (CO-130718-RS)

The development of startups lies at the heart of economic growth through Open Data. Interviews and observation have shown some common features and needs within the community. Startups that develop products and services around Open Data have a very lean setup with predominantly technical founders and minimal sunk cost. Open Data startups tend to work with Open Source software and partially contribute in its development (MC-130712-FB). It is likely that the startup community will fragment itself as the ecosystem grows. As one of the interviewees mentioned "there already is kind of a geospatial community" (NA-130709-XY). Besides a lean setup, Open Data startups need more infrastructure to support the economy in the desired scale. One of the reasons the British government set up the Open Data Institute was to provide this infrastructure efficiently.

4 Analysis

As described in chapter 4.1 the Open Data Institute was established in late 2012 as a non-profit organisation limited by guarantee. It received GBP 10 million from the Technology Strategy Board, an "innovation vehicle" (ODI-130726-GS) of the government, with the commitment to secure match funding over a period of five years. Furthermore, the ODI received GBP 500,000 as a grant from the philanthropic investor Omidyar Network. The general mission of the ODI is to unlock the economic potential of Open Data. To achieve this mission the ODI set up a startup programme, works on technical infrastructure around Open Data, offers diverse training opportunities, supports research projects and establishes a national and international Open Data community. The ODI reports to the TSB and Omidyar Network using financial and impact metrics.

The ODI provides a number of stepping stones for the Open Data startup community. As Francine Bennett, founder of Mastodon C, explains, the ODI "is a central point for people to come to, to talk to. It's a focus point even in a physical way. People meet here" (MC-130712-FB). Furthermore it helps startups with access to data sets and to gather media attention for the resulting products and services. Promoting successful Open Data projects is an essential task to turn the promises around Open Data into tangible evidence. Bennett states:

"The project ['Prescribing Analytics', a joint venture analysing NHS data] was quite helpful, it helped people understand why analysing is useful. It makes us credible. It was a big story in the news. The work with NESTA [National Endowment for Science, Technology and the Arts] resulted from that. Also the work with the Connected Digital Economy catapult project." (MC-130712-FB)

To identify governance structures and dependencies within the ecosystem, it is important to investigate the relationship between the ODI and the Cabinet Office. Several interviewees confirmed that the Cabinet Office has no operational influence on the ODI (ODI-130726-GS, CO-130718-RS, ODI-130715-SC). However, the relationship is multi-layered, as Gavin Starks, CEO of the ODI, explains:

"We sit outside of government but part of our function is to hold government to account. We are in a very interesting position of sitting outside of government but having ties inside government, including Cabinet Office support [...]. That's incredibly powerful." (ODI-130726-GS)

The operational independence of the ODI is granted, as there is no board member with direct affiliation to the government. TSB and Omidyar Network hold observer seats on the board. According to Starks, this setup creates a governance gap between investors and directors and therefore keeps the organisation "mission focused" (ODI-130726-GS). At the time of research the ODI was drafting a working paper to modify the framework for government procurement:

"Having open by default in the contracts means that service level agreements have to be open, performance against those service level agreements has to be open, so that any derivative data that is created by the contractor on behalf of the public is open by default." (ODI-130726-GS)

4 Analysis

Contradicting my assumption that the government would interfere in the ODI's operational work, these modifications would result in the ODI affecting the public bodies' day-to-day work. Another assumption that has been contradicted is that the government might fund certain startups using the ODI and therefore interferes in the market. However, as Stuart Coleman, Commercial Director of the ODI confirms, there has not been any investment from the ODI. All provided services, like free desk space, are funded using the grant from Omidyar Network to "avoid [...] coming into contact with issues relating state aid" (ODI-130715-SC). This separation between an enabling environment provided by the state and the ownership of companies is crucial within any business ecosystem. When speaking to Olivia Burman from the Cabinet Office, she clarifies the government's position on this:

“Government is not here to innovate, we are here to provide the environment, or to foster the environment for innovation to take place. But we are not responsible for doing it ourselves. [...] The government does not create a dynamic marketplace when it innovates on its own.” (CO-130725-OB).

Besides all the necessary infrastructure, unlocking the economic potential is a reciprocal effect between public bodies and the private sector. Public bodies only develop routines in publishing high-quality data on a regular basis when they experience a direct benefit (CO-130725-OB). This benefit, in the form of products or services can only be created by private sector companies, if they get valuable data in a good quality. The reciprocal loop is evident. The ODI and the Cabinet Office bridge these needs and function as boundary organisations to enable mutual benefit. The ODI pools the demand side of the market by building a startup community. These specific demands can then be processed to the data holders directly or via the Cabinet Office, which again serves as single point of contact for the supply side.

4.2.2 The Interest in Civic Empowerment

A large number of NGO's in the UK work towards a better democracy and more transparency in political processes. The research for this dissertation has shown that over the last ten years two organisations have been significantly influential in the development of the Open Data ecosystem. The Open Knowledge Foundation and mySociety both started their work in 2004 and from there on, although with different approaches, shaped the way public information was made available to citizens.

The Open Knowledge Foundation was legally incorporated in 2004 to provide a framework for the growing social movement around digital commons. From an operational perspective the OKF builds tools and communities around freely available information goods. In recent years their work focused on Open Data as a subdomain of open knowledge. However, long before the term Open Data was introduced people started to improve the way digital information was made available and shared. An early example is the Online Computer Library Center (OCLC), which was set up in 1967 as a non-profit cooperative for libraries to share bibliographic metadata. Two civic initiatives, which more directly influenced the development of the OKF were launched in 2001: With his organisation Creative Commons Lawrence Lessing started to provide several copyright licenses to enable users to communicate which rights they

4 Analysis

reserve or waive for the use of their work. In the same year, Jimmy Wales and Larry Sanger opened up Wikipedia, which should become the worlds most popular online encyclopedia. When talking to Rufus Pollock, founder of OKF, he referred to projects like Everything²⁴¹ or H2G2⁴² as predecessors of the Wikipedia and influences for his understanding of collaboratively assembled open knowledge. He describes that the projects were hosting "thousand of articles and thousands of people and nobody knew what they were allowed to do with them" (OKF-130719-RP). Besides this aspect of freely available information, Pollock also engaged in the semantic web community working with the metadata data model RDF. "Before I did the OKF I was using MusicBrainz⁴³. I was really into RDF by that time, in 2003 or 2002 and MusicBrainz had a lot of data in RDF" (OKF-130719-RP). This semantic web community, headed by Tim Berners-Lee, later on became one of the key drivers of the Open Data movement. As this historic detour shows, the foundations of the Open Knowledge Foundation were set by early projects around the collective creation of information online and the technology to enhance the semantics of metadata. However, the organisation of the OKF was influenced by even another famous strand of digital openness.

The OKF is a global organisation and organised in national chapters, which are coordinated by a central unit located in London. Like other transnational NGOs the chapters develop their own national projects and can attract own funding. The funding streams of the OKF in general consist of grants, donations and occasional consulting work. As an incorporated non-profit the OKF has a board of directors, however, Pollock, benignly smiling, describes the governance structure as him being "a bit like the benevolent dictator" (OKF-130719-RP), a term originating in the Open Source Software (OSS) community. This reveals another major influence on the OKF. Pollock states that the formalisation process of the OKF was strongly inspired by the Apache Software Foundation (ASF), one of the well known boundary organisations in the OSS community. According to Von Hippel & Von Krogh (2009), these boundary organisations are established to negotiate between commercial organisations or governments and social movements. In the case of the ASF the social movement consists of programmers with the belief in openly available source code. For the OKF the movement believes in openly available knowledge and – more specific – openly available government data sets. The OKF has a "very loose concept of membership", because, although interested individuals can officially register on okf.org, "there are no rights attached" (OKF-130719-RP). With growing maturity the OKF slowly develops into an internationally distributed network. The national chapters grow in number and, as an example, the OKF community in Berlin is already bigger than the one in London. Pollock expresses his hope that "in two years time the chapters have more staff then OKF central has by, let's say, an order of magnitude" (OKF-130719-RP). Therefore, membership might move from the central organisation towards the network.

As shown in section 4.2.1, the OKF closely collaborated with the Cabinet Office on projects like the development of data.gov.uk. The observations at several OKF events during the period of data collection showed that this collaborations continue. Members of the Cabinet Office were at least present or were even presenting their efforts to the audience on all of the observed events. The OKF internally communicates via mailing lists and uses these semi-formal meet-ups to communicate with external stakeholders. On the meet-ups the OKF informs about ongoing projects and engages businesses, academics and civil servants to present their efforts

4 Analysis

as well. Besides these obvious win-win relationships it is essential to analyse the relationship between the OKF and the ODI. During my interview Gavin Starks states that the ODI and the OKF have different visions and missions: “We [ODI] are focused on the commercial sector, we are focused on driving the [...] demand side and not just unlocking supply” (ODI-130726-GS). Pollock acknowledges the collaborative functioning of the two organisations as well and highlights the ODI’s ability to “bug” government for data and to “shout about” the positive results of Open Data projects (OKF-130719-RP). The ODI’s abilities to build evidence and talk about it are similar to functions the OKF has delivered over the last decade. What is different is that the evidence provided by the ODI is backed by tangible business models:

“The UK government is interested in growth, that is sort of why they are funding the ODI, like startups, but that also drives transparency. I think combining these logics is quite powerful, similar to Open Source.” (OKF-130719-RP)

Despite this gratitude toward the ODI, the analysis also revealed some concerns towards it. The ODI’s development of Open Data Certificates, as an international standard for the quality of data sets, might become a threat for the movement when a fee is applied for getting accredited. As Pollock puts it, in this case “the incentives become quite complicated” (OKF-130719-RP), as commoditised certificates might contradict the maxim of Open Data to provide it at marginal cost. When, for example, these audited certificates become an industry standard, the cost for the data publisher increases and might be passed on to the consumer. To visualise the threat of introducing even another definition of Open Data, he refers to the Open Access community:

“It’s the case that I worry that you see with Open Access where people get into internal fights about what is Open Access. This is damaging.” (OKF-130719-RP)

The analysis pins down to the question, whether the emerging ODI with its commercial focus is likely to substitute the OKF in some parts of their agenda. This finding can serve as a starting point for future research.

MySociety, founded in 2003 and becoming operational in 2004, is an e-democracy project for civic empowerment. It develops and maintains services like FixMyStreet⁴⁴, WhatDoTheyKnow⁴⁵ or TheyWorkForYou⁴⁶ and was during this research often referred to as one of the focal actors in the Open Data movement. However, it has to be distinguished between the broader movement of democratic empowerment through information technology and the more distinct Open Data movement. MySociety mainly contributes to the civic empowerment by providing tools to communicate with the government in a more efficient way or to give citizens a better access to textual documents held by the government. This kind of data fosters democratic empowerment but does not contribute to the aims of the Open Data movement. In my interview, Steinberg elaborates on the role of Open Data and describes the relationship to the OKF:

4 Analysis

“Data is only one thing amongst a lot that mySociety need to achieve its goals to empower citizens. Its about democracy and civil society. [...] Arguably the existence of free programming languages, web frameworks and operating systems is much more important to mySociety’s story, as the availability to free data. If we don’t have free data it is a small problem, but if we don’t have a free operating system it would be a major problem. That’s the big difference between OKF and mySociety, they are about data and for us it just happens to be one of ten things we need to function.” (MS-130717-TS)

This makes clear, and thereby contradicts my previous assumption, that mySociety plays an important role in the national implementation of e-democracy projects in the UK, but only a minor role in the British Open Data ecosystem. However, Tom Steinberg’s long lasting influence on the process of shaping the digital agenda — including Open Data policies — is evident.

4.2.3 Hybrid Perspectives

The structural taxonomy from section 3.3 has outlined the different types of ecosystem participants. With the analyses in section 4.2.1 and 4.2.2 I have investigated participants following the two complementary schools of thought within the Open Data movement: Growth and civic empowerment. However, in a business ecosystem there are multitudinous hybrid perspectives or even perspectives not covered by the two dimensions at all. The empirical evidence from the field allows a brief look at two stakeholders with inconclusive perspectives: Financial intermediaries and the academic community.

Non-profit initiatives as well as startups (at least in most cases) need financial support to get established and to flourish subsequently. In the UK financial intermediaries supporting Open Data initiatives are scarce and diverse. On the one side there is the government investing in the ODI itself and in adjacent initiatives like the Immersion Programme⁴⁷. On the other side there are philanthropic investors like the Open Society Foundations and Omidyar Network, set up by highly affluent individuals that pursue societal and in the case of Omidyar Network financial goals as well. The latter appears to be the most active non-governmental financier of Open Data activities in the UK. In my interview, Martin Tisne, director of policy at Omidyar Network, explains that they "support a lot of groups in the Open Data movement. But the most prominent ones in the UK are probably the OKF and mySociety" (ON-130722-MT). But Omidyar Network not only supports non-profit organisations with a societal agenda, but has also given a grant to the ODI, which seems likely to be extended after the funding period of two years. Martin Tisné describes the organisation’s current interest in this kind of grants as "encouraging transparency in governments, encouraging citizen engagement in governments and ultimately [increasing] accountability of governments to citizens" (ON-130722-MT). However, Omidyar Network as an "impact investor" also pursues investments in for-profit businesses worldwide, which attempt to solve societal problems with mechanisms of the free market. In the US Omidyar Network has invested in the for-profit startup SeeClick-Fix⁴⁸. Therefore, Omidyar Network, as one of the few financial intermediaries in the Open Data ecosystem, takes up a central role with an hybrid motivation.

4 Analysis

Finally, the last stakeholder I cover with my empirical research is the academic community. As section 2.1.2 shows, the number of researchers working on and with Open Data in the UK is still small. At the time of research the only Universities officially affiliated with the ODI were the University of Southampton and the Manchester Business School. Berners-Lee and Shadbolt both are professors at the University of Southampton, which is currently developing a postgraduate diploma in cooperation with the ODI. The Manchester Business School is the only academic subscriber to the ODI membership scheme. In my interview with the person responsible for this subscription, Tom Kirchmaier, he states that his motivation as an academic to affiliate with the ODI is the expectation to improve the quality and accuracy of data sets provided by public bodies. According to his experiences as a researcher in the field of economics, "Open Government Data is mainly unexplored within the academic landscape" and "not used in any form of teaching yet" (MBS-130705-TK). Informal talks with members of the OKF from an academic environment gave the impression that the Open Data movement is mainly anchored within the computer science community and that researchers from other empirical sciences rather affiliate with the Open Access movement. Generally, Open Data has originated out of academic research within the computer science community but nowadays develops outside of academia. Academic interest from fields like economics will only emerge, when the necessary amount of preprocessing per data set will be reduced significantly and the data becomes more accurate.

CHAPTER 5

Conclusion

In an ongoing process new data sets, formerly locked on state-owned servers, are openly uploaded, downloaded, used, reused and commercialised. The British Open Data ecosystem is constantly evolving – it is a dynamic system. In this dissertation I captured the ecosystem at a certain point in time to describe how the network of public, private and formalised non-governmental actors shapes its development.

From a macro perspective the ecosystem is fuelled by the self-reinforcing reciprocal duality of actors interested in economic growth and civic empowerment. Emerged from the social movement around collaborative aggregation of information, civil activists formalised around ten years ago and promoted Open Data, what was back then still labeled as reusable Public Sector Information. Later on individual civil servants carried the idea inside government, where it ended up as a topic in both opposing electoral campaigns. Due to the Open Government Partnership and the G8 commitment to the Open Data charter, the reciprocal duality is likely to develop on a supranational level as well. However, following the theory of business ecosystems all higher-level effects emerge from self-organising and co-evolving processes on a more granular level. The micro analysis of this dissertation took a closer look at the interests and actions of individual stakeholders. Stakeholders mainly interested in growth, respectively the government and companies from the private sector, perform a circular reinforcement process. Public bodies only develop routines for publishing high-quality data regularly when they experience a direct benefit through products or services developed by innovators. However, these innovators can only perform when they get valuable data in a good quality. To launch and accelerate this circle the ODI and the Cabinet Office serve as boundary organisations. Stakeholders mainly interested in civic empowerment reach their goal by working very collaboratively with stakeholders from the growth perspective. Although the major stakeholder in this perspective, the Open Knowledge Foundation, is organised as a distributed network, it

5 Conclusion

is unofficially led by an individual, who has long-lasting experiences in the field. This element refers to the phenomenon of conscious choice, a crucial element within a business ecosystem.

There are certain limitations to this research, which are determined by the accessibility of data. The data collection was organised with kind support of the Open Data Institute. This helped me to conduct significantly more interviews than I would have been able to organise without support. The credibility as an associated researcher of the ODI also helped me to collect data that was very rich. However, the convenient access to persons affiliated with the ODI clearly created a selection bias of my results. My insights into the activities of the ODI have been much deeper than into all the other stakeholders. This helped me to understand the ODI's role in depth but might have skewed my interpretation. Furthermore, I was not able to conduct interviews with representatives from major data holding bodies like the Ordnance Survey or the Met Office due to a lack of time. Therefore, their position in the ecosystem could not be clarified.

Future research on the British Open Data ecosystem might focus more on the relation between the OKF and the government, as this relation on the one hand revealed many different facings and on the other hand provides a lot of accessible data in form of public documents and public mailing lists.

NOTES

Notes

- ¹<http://webarchive.nationalarchives.gov.uk/20071204130045/opsi.gov.uk/click-use/index.htm>
- ²<http://www.nationalarchives.gov.uk/appsi%5Cdefault.htm>
- ³http://ec.europa.eu/information_society/policy/psi/rules/eu/index_en.htm
- ⁴<http://blog.okfn.org/2005/10/19/open-knowledge-definition-released/>
- ⁵<http://opendefinition.org/>
- ⁶<http://www.freeourdata.org/>
- ⁷<http://blog.ordnancesurvey.co.uk/2010/04/os-opensdata-goes-live/>
- ⁸http://www.oft.gov.uk/713560/publications/reports/consumer-protection/oft861.Ue_ds1N5Ey5
- ⁹<http://inspire.jrc.ec.europa.eu/>
- ¹⁰<http://www.opsi.gov.uk/advice/poi/power-of-information-review.pdf>
- ¹¹<http://epsiplatform.eu/content/cambridge-study>
- ¹²<http://webarchive.nationalarchives.gov.uk/20100402134053/showusabetterway.com/>
- ¹³<http://wheredoesmymoneygo.org/>
- ¹⁴<http://openspending.org/>
- ¹⁵<https://www.eff.org/files/filenode/foia/2009foia.mem.rel.pdf>
- ¹⁶<http://opendatacommons.org/2009/02/27/open-database-license-draft-available-for-comments/>
- ¹⁷<http://opendatacommons.org/faq/licenses/db-versus-contents>
- ¹⁸<http://www.data.gov/>
- ¹⁹http://webarchive.nationalarchives.gov.uk/20090707073427/http://cabinetoffice.gov.uk/newsroom/news_releases/2009/090610_web.aspx
- ²⁰<http://www.guardian.co.uk/technology/blog/2010/jan/21/timbernerslee-government-data>
- ²¹http://www.conservatives.com/News/News_stories/2010/03/Conservative_Technology_Manifesto_launched.aspx
- ²²<http://steiny.typepad.com/premise/2009/10/yes-im-going-to-be-advising-the-opposition-on-it.html>
- ²³http://www.publicservice.co.uk/feature_story.asp?
- ²⁴<http://www.bbc.co.uk/news/10152929>
- ²⁵<http://data.gov.uk/blog/new-public-sector-transparency-board-and-public-data-transparency-principles>
- ²⁶<http://webarchive.nationalarchives.gov.uk/+http://www.cabinetoffice.gov.uk/newsroom/statements/transparency/pm-letter.aspx>
- ²⁷<http://www.nationalarchives.gov.uk/doc/open-government-licence/version/1/open-government-licence.htm>
- ²⁸<http://data.gov.uk/blog/new-open-government-licence>
- ²⁹<https://www.gov.uk/government/news/letter-to-cabinet-ministers-on-transparency-and-open-data>
- ³⁰<http://www.opengovpartnership.org/countries/united-kingdom>
- ³¹<https://www.gov.uk/government/news/uk-sets-out-objectives-as-joint-chair-of-global-partnership-on-global-government>
- ³²<https://www.gov.uk/government/consultations/open-government-partnership-uk-draft-national-action-plan-2013>
- ³³http://www.innovateuk.org/_assets/0511/ODI_Implementation_Plan_May_Release.pdf
- ³⁴<https://www.gov.uk/government/news/autumn-statement-2011-3>
- ³⁵<https://www.gov.uk/government/news/chair-of-the-open-data-user-group-appointed>
- ³⁶<https://www.gov.uk/government/publications/open-data-white-paper-unleashing-the-potential>
- ³⁷http://www.deloitte.com/assets/Dcom-UnitedKingdom/Local_Assets/Documents/Market_insights/Deloitte_Analytics/uk-insights-deloitte-analytics-open-data-june-2012.pdf
- ³⁸https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198752/13-744-shakespeare-review-of-public-sector-information.pdf
- ³⁹<https://www.gov.uk/government/publications/government-response-to-shakespeare-review>
- ⁴⁰<https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>
- ⁴¹<http://www.everything2.com/>
- ⁴²<http://www.h2g2.com/>
- ⁴³<http://musicbrainz.org/>
- ⁴⁴<http://www.fixmystreet.com/>
- ⁴⁵<https://www.whatdotheyknow.com/>
- ⁴⁶<http://www.theyworkforyou.com/>
- ⁴⁷https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/84197/FAQs_Open_Data_Immersion_Programme.pdf
- ⁴⁸<http://seeclifix.com/>

Appendix A

A1 Logic of Interview Identification Codes

The identification codes that appear in the body of this dissertation refer to interviews, which are pooled in the case study database (figure 3.1). The following briefly describes the denotation scheme. Anonymous interview partners are indicated by the initials 'XY'.

Scheme: Stakeholder-YearMonthDay-Initials/Comment

Example: ODI-130711-XY

ODI Open Data Institute

MC MastodonC

NA National Archives

ON Omidyar Network

CO Cabinet Office

OKF Open Knowledge Foundation

MBS Manchester Business School

MS MySociety

NOTES

A2 List of Interview Partners

Francine Bennett Mastodon C (CEO)

Olivia Burman Cabinet Office (Head of Relationship Management)

Stuart Coleman ODI (Commercial Director)

Leigh Dodds ODI (Associate Technical Team)

Tom Heath ODI (Head of Research)

Tom Kirchmaier Manchester Business School / LSE (Lecturer)

Briony Phillips ODI (Immersion Programme Manager)

Rufus Pollock Open Knowledge Foundation (Founder & Director)

Carl Rodrigues ODI (Associate Business Development)

Gavin Starks ODI (CEO)

Tom Steinberg mySociety (Founder & Director)

Richard Stirling ODI (Membership Programme Manager) / Formerly: Cabinet Office

Jeni Tennison ODI (Technical Director)

Martin Tisné Omidyar Network (Director of Policy UK)

Anonymous The National Archives

A3 List of Events

Event	Date	# of Participants
OKF Open Data London Meet-up	July 2, 2013	40
OKF Open Data Maker Night London	July 16, 2013	10
The Foundation for Science and Technology Roundtable	July 10, 2013	30
Open Data – Better Society? Student Seminar at University of Edinburgh	April 26-27, 2013	25
ODI Lunchtime Lecture	July 5, 2013	15
ODI Lunchtime Lecture	July 12, 2013	25
University of Southampton Focus Group on Neighbourhood Technology	July 9, 2013	5

NOTES

A4 Exemplary Interview

Interview Partner: Richard Stirling

Details: Membership Programme Manager at Open Data Institute, formerly Secretariat of Power of Information Task Force at Cabinet Office

Interviewer: Maximilian Heimstädt

Date: 18/07/13

Location: Open Data Institute

Audio File: Yes

Abbreviations: MH = Maximilian Heimstädt ; RS = Richard Stirling

Identification Code: CO-130718-RS

MH: In Germany Open Data is affiliated with the Home Office, here in the UK it is with the Cabinet Office. Why? Are there historical reasons?

RS: It is with the Cabinet Office because it came out of a research, an internal government research project, the Cabinet Office in the UK tends to be where cross-cutting projects stay. It seems to have application everywhere, but they take projects that don't really fit anywhere, that are not daily business for any department. These projects tend to sit in the Cabinet Office. Then they get a lot of energy input and then they become business as usual. The reason it's still there is that the Cabinet Office slightly changed its function under this administration and takes more business-as-usual functions. It's in the business of "doing things" now which it didn't used to be.

MH: When did the idea to data.gov.uk come up? Did it come out of the research project as well?

RS: You say that you talked to Tom Steinberg. He was part of the Power of Information report team. The following one was the Power of Information Task Force. Established reporting to the Minister of Cabinet Office at the time, Tom Watson. I was part of the secretariat for that task force. I sort of helped to draft some of the data sections of the report and then – when the report was published – the task force was disbanded. I established a new unit within the Cabinet Office, staffed it up, got the resources, hired the staff, build the initial data.gov.uk, established the mechanisms for releasing data internally to government. Changed the rules for enable it to happen. Established a support network for people across departments that were releasing the data – to make it happen.

MH: The idea was developed within the task force? Was it developed within government or was there external pressure?

RS: Any good policy making process involves bringing in the best ideas from the outside. This was no different. There have been lots of people, Gavin [Starks] was one of them, that have been arguing that government is awful in managing its data, and that it should start releasing it as Open Data. There was even a "Free Our Data" campaign out of the Guardian, run by

NOTES

Michael Cross and Charles Arthur. This was the first time when government had responded to the pressure, said "yes", and had done something about it. If you read the Power of Information task force report you see there are twenty quite chunky recommendations. It breaks down roughly to 10 on data and 10 on "there are people on the internet, they are talking about you, don't you want to be part of this conversation". And one of those recommendations was to create a single point of contact for government data sets and to aid discovery.

MH: Data.gov.uk was released under CKAN, which was earlier developed by OKF. Can you elaborate?

RS: That's technically not true.

MH: That's how it is written on the OKF Blog.

RS: CKAN was established before data.gov.uk, the first version was ... I genuinely can't remember... we did a disposable version just to sort of get a working proof of concept, which we opened as a working proof of concept. ...no...I... no, it was based on CKAN. Have you seen Rufus' slides on CKAN? He has a nice little hand drawn sketch of like the CKAN data.gov.uk integration, which he drew in the conversation we had in the coffee shop around the corner of the Cabinet Office. This was the information architecture for data.gov.uk, so we kind of co-designed it.

MH: How was the collaboration with the Open Knowledge Foundation like? When did it start, how did it evolve?

RS: In a good policy making process you are talking to all the people who are interested. You work out what their relative areas of expertise are and bring them into the solution. That's a good government design and implementation process. So it was integral all the way through, I was talking to Rufus when we were writing the independent report, I was talking to Rufus when we were thinking about what we would do, I was talking to Rufus when we were doing the initial design.

MH: Now you are not at the Cabinet Office anymore but can you make an assumption how the relationship to the OKF developed?

RS: It's a different relationship at different stages. When data.gov.uk was first done, both projects were in a much earlier stage. OKF was smaller than it is today and the data team from the cabinet office was smaller.

NOTES

MH: Data.gov.uk launched in September 2009, only four month after data.gov in the US. How did these two projects relate to each other?

RS: This is a real learning point for me. Never, never betray your good ideas. I had a phone call with Beth Noveck before her inauguration. I had a draft of the Task Force report, which I knew is going to be welcomed. So I made the mistake of saying "here is what we are planning on doing" and of course, a new president came in and they could move faster. And they did.

MH: So it was a British idea?

RS: I wouldn't say that. Ideas have no real ownership. It was a collective idea. It was something that people were calling for all over the place and I ended up the person doing.

MH: Did the US and the UK collaborate in the development?

RS: No, it was an independent effort.

MH: And structure wise do they resemble each other?

RS: They didn't and they do more now. CKAN now holds the US one as well and indeed lots of other similar data platforms around the world.

MH: Which were the major stakeholders you pulled together at the Cabinet Office to draft the report?

RS: The best thing to do is to look at the public blog that we maintained when we worked on the report. To look at the members of the task force, to look where they came from. And you will also see that we had a good mix of interest from across the place. We had a recently wide representation of people that were interested in data from the commercial and non-commercial side. And so not just OKF but mySociety and the other activists.

MH: How did you communicate with the "average citizen"? Was there any vehicle?

RS: It was deliberately done as a specialist model. At the beginning of data.gov.uk I was quite explicit that this was a phased approach and phase one was to get the audience of largely developers, because that is the ideal audience to cater for but also the one that gives initial feasibility. The impact of Open Data would then broaden out to businesses and then broaden out to the citizens. That takes a period of years to go from one to the other.

NOTES

MH: How did you communicate with the developers? Who to approach? Where do they meet?

RS: There was a period where twice a week, maybe sometimes five days a week I went to developer meet-ups. And I would speak to an audience of between 50 and 250 developers about the things the government is doing to open up their data to be available for people to use. We have nothing like the Chaos Computer Club here. There is a very healthy market of meet-ups advertised on meetup.com. You also get referrals. People see you speak and you do quite well, then people recommend you to other meet-ups. We co-opted people around the country to speak to people. Somebody that worked for me went around the academic institutions, he walked around and gave talks at undergrad courses on computer science at about eight different campuses. We did all of these things to try to get the message out. At that time it was new and world leading, so we could get some press coverage. The Guardian was quite attentive.

MH: What was the role of the media? The role of the Guardian?

RS: A lot of people at the Guardian wanted the data and they even developed this campaign. The media... eh... they were very grateful for the data when they got it. But the primary motivation was good governance and efficient and effective public services. So there is far too much of me that pretends to be a communist and if you think of the way that markets work then removing data asymmetry is a good thing and if you want to promote well functioning government then you need to have data about how it fits together.

MH: You worked with Ordnance Survey in 2009 as well. Did you try to overcome the Trading Fund model?

RS: We worked with the Ordnance Service around the initial release of data, called "OS Open". But I was still in Cabinet Office, we did all the hard work. Read the excellent paper Rufus Pollock wrote with [David] Newbery.

MH: What is the role of the ODI after the funding period? Is there a chance to become independent from governmental funding? Will it stay an arm of the Cabinet Office?

RS: It's not an arm of the Cabinet Office, it has been set up as an independent organisation. It just happens to have some funding. But we have no board representation from government, they have no influence on our day-to-day business. We have to meet TSB metrics, but they are incidentally met by us achieving our goals anyway. So they are not governmentally focused, these are metrics set up around the health of the organisation. The independence is a question you should ask the management team but I guess that we try to be self sufficient by the end of the funding period.

NOTES

A5 Exemplary Coding Table

Interview Excerpt	Interpretation
<p>“It is with the Cabinet Office because it came out of a research, an internal government research project, the Cabinet Office in the UK tends to be where cross-cutting projects stay. It seems to have application everywhere, but they take projects that don’t really fit anywhere, that are not daily business for any department. These projects tend to sit in the Cabinet Office. Than they get a lot of energy input and then they become business as usual. The reason it’s still there is that Cabinet Office slightly changed its function under this administration and takes more business-as-usual functions. It’s in the business of "doing things" now which it didn’t used to be.” (CO-130718-RS)</p>	<p>The Open Data initiative is associated with the Cabinet Office. Under former governments the Cabinet Office functioned as an accelerator for emerging fields of engagement, which were then passed on to an appropriate department. However, the Cabinet Office’s role changed slightly under the current administration and is responsible for Open Data activities on a day-to-day basis.</p>
<p>“Any good policy making process involves bringing in the best ideas from the outside. This was no different. [...] This was the first time when government had responded to the pressure, said "yes", and has done something about it.” (CO-130718-RS)</p>	<p>The Cabinet Office included experts from the field of e-government into the process of policy making.</p>
<p>“Have you seen Rufus’ slides on CKAN? He has a nice little hand drawn sketch of the CKAN data.gov.uk integration, which he drew in the conversation we had in the coffee shop around the corner of the Cabinet Office. This was the information architecture for data.gov.uk, so we kind of co-designed it.” (CO-130718-RS)</p>	<p>To set up data.gov.uk the Cabinet Office closely collaborated with the Open Knowledge Foundation. In the interviews with Rufus Pollock and Richard Stirling both independently referred to an informal but focal meeting when they co-designed the initial sketch of the CKAN data.gov.uk integration on a napkin.</p>
<p>“In a good policy making process you are talking to all the people who are interested, working out what their relative areas of expertise are and bringing them into the solution. [...] I was talking to Rufus when we were writing the independent report, I was talking to Rufus when we were thinking about what we would do, I was talking to Rufus when we were doing the initial design.” (CO-130718-RS)</p>	<p>Richard Stirling describes the process of collaborative policy making of the Cabinet Office with Rufus Pollock.</p>

NOTES

<p>“This is a real learning point for me. Never, never betray your good ideas. I had a phone call with Beth Noveck before her inauguration. As I had a draft of the Task Force report, which I knew is going to be welcomed. So I made the mistake of saying "here is what were planning of doing" and of course, a new president came in and they could move faster. And they did.” (CO-130718-RS)</p>	<p>It is a task for future research to investigate the co-evolution of the US and UK ecosystem. Plans for the British data portal have been around since the first draft of the POI Task Force report. Richard Stirling, responsible for the development of data.gov.uk reports from a phone call with Beth Noveck, who became CIO of the Obama administration.</p>
<p>“It was deliberately done as a specialist model. At the beginning of data.gov.uk I was quite explicit that this was a phased approach and phase one was to get the audience of largely developers because that is the ideal audience to cater for, but also the one that give initial feasibility. The impact of Open Data would then broaden out to business and then broaden out to the citizens. That takes a period of years to go from one to the other.” (CO-130718-RS)</p>	<p>The UK Open Data initiative was planned by the Cabinet Office as a phased model. Phase one should target the developers, phase two businesses and phase three the citizens.</p>
<p>“There was a period where twice a week, maybe sometimes five days a week I went to developer meet-ups. And I would speak to an audience of between 50 and 250 developers about the things the government is doing to open up their data to be available for people to use. We have nothing like the Chaos Computer Club here.” (CO-130718-RS)</p>	<p>The Cabinet Office used different meet-ups to engage the developers for the first stage. Groundwork. A formalised developers club like the German “Chaos Computer Club” might have facilitated this process.</p>
<p>“Somebody that worked for me went around the academic institutions, he walked around and gave talks at undergrad courses on computer science at about eight different campuses. We did all of these things to try to get the message out.” (CO-130718-RS)</p>	<p>The Cabinet Office targeted computer science students at eight different campuses to engage developers from an academic environment.</p>

NOTES

<p>“A lot of people at the Guardian wanted the data and they even developed this campaign. The media... eh... they were very grateful for the data when they got it. But the primary motivation was good governance and efficient and effective public services.” (CO-130718-RS)</p>	<p>The role of the media is not quite clear. Richard Stirling mentioned the “Free Our Data” campaign set up by the Guardian but names “good governance and efficient and effective public services” as main motivation for the governmental efforts.</p>
<p>“Its not an arm of the Cabinet Office, its been set up as an independent organisation. It just happens to have some funding. But we have no board representation from government they have no influence on our day-to-day business.” (CO-130718-RS)</p>	<p>Stirling confirms that the government has no influence on the operational work.</p>

NOTES

A6 Initial Interview Manual

Administration:

Thank you very much for your time. The interview will take between 20 and 30 minutes. The interview will be audio-recorded and transcribed afterwards. The results will form part of my MSc dissertation at University of St Andrews. I study Management IT. At the end of the interview you can decide whether you may be cited with clear name or anonymised.

The Project:

I study the British Open Data ecosystem. I consider Open Government Data to be a resource that is processed within an ecosystem – quite similar to a biological ecosystem. My theoretical foundation is based on business ecosystem theory, complex systems and evolutionary economics. I like to investigate stakeholders of this ecosystem to learn how it evolves. I identified [your institution] as a key stakeholder of the ecosystem.

0. Contextual Data

- Date and time
- Name of interviewee
- Company/institution
- Position of interviewee
- Location and setting

1. General Questions on the Ecosystem

- Open Data comes in different formats and categories. Which kinds of Open Data do you consider the most valuable ones?
- How do you define the value of Open Data?
- I consider the Open Data ecosystem as a network of data producing, processing and consuming stakeholders within a regulatory and enabling environment. Which are the most important bodies that produce Open Data in the UK?
- Which are the major stakeholders that process Open Data in the UK?
- Citizens are the only consumers of Open Data? Do you agree?
- Which stakeholders create the regulatory framework? Which build capacity?
- Do you consider the UK as the best developed Open Data ecosystem worldwide?
 - If yes, why?
 - If no, which one is better and why?

NOTES

- What do you consider as critical success factors of the UK Open Data ecosystem?
- Have there been pivotal events that influenced the development?
- How does the political climate influence the development?

2. Stakeholder Questions

- Please describe briefly how your organisation engages with Open Data today?
- What limits your growth concerning your Open Data activities?
- With which Open Data stakeholders do you have business relationships in the broader sense?
- With which OGD stakeholders do you compete in a broad sense?
- Can you give positive or negative examples when your development was influenced by other market participants?

3. ODI Questions

- Do you know the function of the ODI?
 - If yes, continue.
 - If no, I will explain.
- How do you relate to the ODI?
- The ODI is partly tax-funded. Do you see any adverse interference in the market?
- Rate the importance of the ODI for your organisation from 1 (lowest) to 10 (highest)?
- Imagine you are profit-maximising Venture Capitalist. Would you invest in the ODI? Why?

Final Remarks:

We reached the end of this interview. Thank you very much for your participation. Are there any more questions, remarks or suggestions for improvement? Can I cite you directly? If you like I can send you a final version of the dissertation when I am done.

Appendix B

B1 Ethical Approval

PRELIMINARY ETHICS SELF-ASSESSMENT CHECKLIST FOR PROJECT SUPERVISORS

Please read this page carefully and then sign it together with the attached ethical self-assessment form before returning them to the ethics secretary.

If you have any unsure whether or not a full ethics application is required, please email ethics-cs@st-andrews.ac.uk.

A full ethics application **is not required** if:

- 1) The answer to question 1 on the self-assessment form is no;
- 2) The answer to question 1 on the self-assessment form is yes AND the answer to questions 2-15 is no.

A full ethics application **is required** if:

- 1) The answer to any of questions 2-15 on the self-assessment form is yes;
- 2) The project involves data from any social networking sites (Facebook etc.);
- 3) The project involves health data (e.g. working with NHS projects);
- 4) The project involves working with children (a Disclosure Scotland form is also required so this should be flagged up well in advance of starting the project).
- 5) The answer to question 1 is yes AND supervisor/ethics convenor thinks a full ethics form is necessary for some reason not covered by the above.

Please complete applicants' name

Maximilian Heimstädt

"I verify that as supervisor, in addition to reading through all of the student's completed self-assessment form, I have also checked that a full ethics form **should/should not** (delete one) be completed."

Signature 

Date 24/4/2013

Name Tom KELSEY.

Bibliography

- Alani, H., Dupplaw, D., Sheridan, J., O'Hara, K., Darlington, J., Shadbolt, N., & Tullo, C. (2007). Unlocking the potential of public sector information with semantic web technology. In *The Semantic Web*, (pp. 708–721). Springer.
- Anderson, P. (1999). Perspective: Complexity theory and organization science. *Organization Science*, 10(3), 216–232.
- Berners-Lee, T. (2009). Putting Government Data Online. Retrieved from <http://www.w3.org/DesignIssues/GovData.html> on August 11, 2013.
- Berners-Lee, T. (2011). Linked Data - Design Issues (2011). Retrieved from <http://www.w3.org/DesignIssues/LinkedData.html> on August 11, 2013.
- Berners-Lee, T. (2013). Details of Web Foundation research. Retrieved from <http://www.webfoundation.org/2013/04/berners-lee-unveils-details-of-web-foundation-research-at-ogp-meeting/> on August 11, 2013.
- 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.
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data - The story so far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5(3), 1–22.
- Böhm, C., Naumann, F., Freitag, M., George, S., Höfler, N., Köppelmann, M., Lehmann, C., Mascher, A., & Schmidt, T. (2010). Linking Open Government Data: What journalists wish they had known. In *Proceedings of the 6th International Conference on Semantic Systems*, (p. 34). ACM.
- Choi, T. Y., Dooley, K. J., & Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: Control versus emergence. *Journal of Operations Management*, 19(3), 351–366.

Bibliography

- CIPESA, & APC (2012). Uganda: Open Government Data Readiness Study. Retrieved from http://www.cipesa.org/?wpfb_dl=39 on August 11, 2013.
- Davies, T. (2010). Open Data, democracy and public sector reform. Retrieved from <http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf> on August 11, 2013.
- Davies, T., & Edwards, D. (2012). Emerging implications of open and linked data for knowledge sharing in development. *IDS Bulletin*, 43(5), 117–127.
- Davies, T., Perini, F., & José, M. A. (2013). *Researching the emerging impacts of Open Data*. ODDC Working Papers #1.
- Easterby-Smith, M., Thorpe, R., Jackson, P., & Lowe, A. (2008). *Management Research*. SAGE Publications Limited.
- Eaves, D. (2012). Case study: How Open Data saved Canada \$3.2 billion. Retrieved from <http://fairwhistleblower.ca/content/case-study-how-open-data-saved-canada-32-billion> on August 11, 2013.
- Eysenck, H. J. (1976). *Case studies in behaviour therapy*. Routledge Kegan & Paul.
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative inquiry*, 12(2), 219–245.
- Foster, J., & Metcalfe, J. S. (2001). Modern evolutionary economic perspectives: An overview. *Frontiers of Evolutionary Economics*.
- Geiger, C. P., & von Lucke, J. (2011). Open Government Data. In *Conference for E-Democracy and Open Government*, (p. 183).
- Grewal, A., Iglesias, C., Alonso, J. M., Boyera, S., & Bratt, S. (2011). Open Government Data – Feasibility Study in Ghana. From: http://public.webfoundation.org/2011/02/OGD_ghana_PR.pdf.
- Hammell, R., Perricos, C., Branch, D., & Lewis, H. (2011). Unlocking growth: How open data creates new opportunities for the uk.
- Hannon, B. (1997). The use of analogy in biology and economics: From biology to economics, and back. *Structural Change and Economic Dynamics*, 8(4), 471–488.
- Hodgson, G. M. (1998). Economics and evolution: Bringing life back into economics. *Polity Press and University of Michigan Press, Cambridge, UK*.
- Hoxha, J., & Brahaj, A. (2011). Open Government Data on the Web: A Semantic Approach. In *Emerging Intelligent Data and Web Technologies (EIDWT), 2011 International Conference on*, (pp. 107–113). IEEE.
- Iansiti, M., & Levien, R. (2004). *The keystone advantage: What the new dynamics of business ecosystems mean for strategy, innovation, and sustainability*. Harvard Business School Press.

Bibliography

- Kloiber, J. (2012). Open Government Data – Zwischen politischer Transparenz und Wirtschaftsförderung. Retrieved from http://igitur-archive.library.uu.nl/student-theses/2012-1004-200536/Masterthesis_3806316_JuliaKloiber.pdf on August 11, 2013.
- Koski, H. (2011). Does marginal cost pricing of public sector information spur firm growth? Tech. rep., The Research Institute of the Finnish Economy.
- Lenway, S. A., & Murtha, T. P. (2004). Knowledge and competitive advantage: The co-evolution of firms, technology and national institutions. *Journal of International Business Studies*, 35(6), 560–563.
- Lewin, R. (1999). *Complexity: Life at the edge of chaos*. University of Chicago Press.
- Majeed, R. (2012). Disseminating the Power of Information: Kenya Open Data Initiative, 2011–2012. Tech. rep.
- Merry, U. (1999). Organizational strategy on different landscapes: A new science approach. *Systemic Practice and Action Research*, 12(3), 257–278.
- Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Belknap Press.
- Neuman, W. L. (2005). *Social research methods: Quantitative and qualitative approaches*. Allyn and Bacon.
- Open Knowledge Foundation (2013). Open Definition. Retrieved from: <http://okfn.org/> on August 11, 2013.
- Pagie, L. (1999). *Information integration in evolutionary processes*. PhD thesis.
- Parks, W. (1957). Open Government Principle: Applying the right to know under the constitution. *George Washington Law Review*, 26.
- Peltoniemi, M. (2006). Preliminary theoretical framework for the study of business ecosystems. *Emergence: Complexity & Organization (E:CO)*.
- Peltoniemi, M., & Vuori, E. (2004). Business ecosystem as the new approach to complex adaptive business environments. In *Proceedings of eBusiness Research Forum*, (pp. 267–281).
- Phan, D. (2004). From agent-based computational economics towards cognitive economics. *Cognitive Economics: An Interdisciplinary Approach*, (p. 371).
- Rahemtulla, H., Custer, S., Tisacova, I., Jhalla, K., Gigler, S., & Brigham, C. (2012). The Journey of Open Government and Open Data Moldova. Tech. rep. Retrieved from <http://de.scribd.com/doc/99595560/The-Journey-of-Open-Government-and-Open-Data-Moldova> on August 11, 2013.

Bibliography

- Rahemtulla, H., Kaplan, J., Gigler, B.-S., Cluster, S., Kiess, J., & Brigham, C. (2011). Open Data Kenya: Case study of the underlying drivers, principal objectives and evolution of one of the first Open Data initiatives in africa. Tech. rep. Retrieved from <http://www.scribd.com/WorldBankPublications/d/75642393-Open-Data-Kenya-Long-Version> on August 11, 2013.
- Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers*, vol. 2. Blackwell Oxford.
- Rojas, L. A., Lovelle, J. M., Tarazona, G. M., & Bermúdez, C. E. (2013). Open Data as a key factor for developing expert systems: A perspective from Spain. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2.
- Saunders, M., Lewis, P., & Thornhill, A. (2011). *Research Methods for Business Students*. Pearson Education India.
- Schütze, F. (1983). Biographieforschung und narratives Interview. *Neue Praxis*, 13(3), 283–293.
- Smith, M. Y., & Stacey, R. (1997). Governance and cooperative networks: An adaptive systems perspective. *Technological Forecasting and Social Change*, 54(1), 79–94.
- Stott, A., & Kaplan, J. (2013). Open Data Readiness Assessment Tool (January 2013 version). Retrieved from <http://data.worldbank.org/about/open-government-data-toolkit/readiness-assessment-tool> on August 11, 2013.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed Methodology: Combining qualitative and quantitative approaches*, vol. 46. SAGE Publications.
- Tauberer, J. (2007). 8 Principles of Open Government Data. Retrieved from <http://www.opengovdata.org/home/8principles> on August 11, 2013.
- Tong, A., Irsha, H., & Revell Ward, D. (2013). Uganda Open Government Data readiness study. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf on August 11, 2013.
- Vickery, G. (2011). Review of recent studies on PSI reuse and related market developments. *Information Economics, Paris*.
- Von Hippel, E., & Von Krogh, G. (2009). Open source software and the 'private-collective' innovation model: Issues for organization science. *MIT Sloan Research Paper*.
- Yu, H., & Robinson, D. (2012). The New Ambiguity of Open Government. *Princeton CITP/Yale ISP Working Paper*.