Master's Thesis
1. June 2016

Exploring the Potential of Businesses as Open Data Publishers

A Case Study on Trustpilot

Bjarke Due Jensen **Reshal Katyal**

Supervisor
Attila Marton

Programme
IT Management and Business Economics

Number of Pages: 116
Number of Characters: 222.390

COPENHAGEN BUSINESS SCHOOL

Abstract

Open data is a significant contributor to the current data revolution we are witnessing. Governments around the world are opening data to foster transparency and innovation. Open data is expected to create more than $3 trillion in annual value (Manyika et al., 2013). The majority of research on open data is concerned with open government data, while little is known about companies as open data providers.

The purpose of this thesis is to explore companies as contributors of open data. In this study, we take the first tentative steps in understanding external data sharing in a business context. To understand why companies would contemplate opening data, it is first necessary to understand how they benefit from sharing data in a loosely coupled manner. Through a case study on Trustpilot, we examine a data sharing initiative in a business context and pose the question, "how is external data sharing justified in business?". The research question is addressed through a qualitative and naturalistic research approach. In the analysis of Trustpilot, we examine the openness of their data sharing initiative, the role of data sharing in the business model and explore Trustpilot as a platform to understand the rationale for sharing data in a business ecosystem.

Based on our empirical findings we propose that external data sharing in business is justified by the opportunity to 1) increase customer utility, 2) extend brand exposure, 3) amplify platform dynamics and 4) leverage complementors in the business ecosystem. Additionally, we find indications that companies as open data providers is inhibited by 1) undermined value capture, 2) loss of control, 3) adverse brand effects and 4) the need for a proactive approach.

The picture that emerged from our findings is that limited data access induces artificial data scarcity. If no data access restrictions applied, existing business models could be threatened. We suggest further investigation of how business models and strategy can accommodate providing open data, while running a profitable business.

Keywords

Open Data, Shared Data, Data Spectrum, Open Data Business Model, External Data Sharing, Business Model, Platforms, Multi-sided platform, Industry Platform, Trustpilot.

PREFACE

This paper is a master thesis at Copenhagen Business School. The authors are graduating students at the programme IT Management and Business Economics.

Acknowledgements

Special appreciation goes to our supervisor Attila Marton from the Department of IT Management, Copenhagen Business School. Thanks for the valuable guidance and input along the process of carrying out an ambitious study into a new direction of an already immature research domain.

Gratitude goes to Trustpilot and their employees who cared to allocate some of their precious time to provide us with insights from their business. The business they are running is truly inspirational. We wish the best for their future endeavours and will follow them closely.

The Open Data phenomenon was initially brought to our attention by Thorhildur Jetzek at a guest lecture in 2013. We thank her for the enlightenment and the exchange of ideas during the design process of this study.

Additionally, we would like to thank family, friends, and colleagues for their understanding and support in the process.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

API	Application Program Interface
ODBM	Open Data Business Model
OD	Open Data
ODI	Open Data Institute
OGD	Open Government Data
HTTP	Hypertext Transfer Protocol
LOD	Linked Open Data
MSP	Multi-sided Platform
PSI	Public Sector Information
SaaS	Software-as-a-Service
SEO	Search Engine Optimization
URI	Uniform Resource Identifier
RDF	Resource Description Framework

INTRODUCTION

The purpose of this thesis is to examine and understand how businesses view exposing data outside the boundaries of the firm. The interest stems from the phenomenon referred to as 'Open Data' and the idea that free and unrestricted data access can generate substantial value and innovation in society. We open the discussion of companies as open data publishers. Through a case study on Trustpilot we demonstrate that external data sharing is justified in business by the opportunity to 1) increase customer utility, 2) extend brand exposure, 3) amplify platform dynamics and 4) leverage complementors in a business ecosystem.

A data revolution

The world's data masses are doubling every other year (IDC Research, 2014). We are all contributors to the data masses through our digital interactions, such as the use of our smartphones, wearables, our interaction on social media and through the e-mails we send at work.

IDC Research (2014) estimates that the world's data masses will expand from 4.4 zettabytes in 2013 to 44 zettabytes by 2020[1]. With cheaper storage, faster processing, and better tools for dealing with big data, we continue to see data-driven value creation across society. We are witnessing a data revolution (Kitchin, 2014): A substantial paradigm shift, which is comparable to the industrial revolution, due to the potential to fundamentally transform society and the way we do business.

A significant contributor to the data revolution is the concept of *Open Data*. Open data is the idea of releasing datasets for everyone to access and use. Manyika et al. (2013) estimate the potential annual value enabled by open data to be $3 trillion. The United States kick-started an open data snowball when Barack Obama in 2009 announced the Open Government Directive (Attard et al., 2015). A transparency strategy that involved opening government data to the public. Today USA continues to be among the top ranked open data countries (World Wide Web Foundation, 2015) and offers more than 180.000 data sets (data.gov, 2016). A figure that

[1] 44 zettabyte = 44 billion TB (terabytes).*
*Equivalent to approximately 6TB per person in the world.

worldwide already reached 1 million back in 2013 (Manyika et al., 2013). However, data first becomes valuable when put to work in organisations and businesses worldwide. Zillow is a popular example of a company that extracts value from open government data (Jetzek, 2015). Zillow is a real estate aggregator that includes data on properties, sales history, pricing patterns and features of the neighborhood to guide consumers and facilitate the transaction. As a result, the property information offered reduces the need for a broker and substantially reduces the buyer's search costs.

Open Data Research

Researchers within the open data domain have addressed issues of finding benefits with open government data (Janssen et al., 2012; Ubaldi, 2013; Davies & Bawa, 2012), challenges in publishing data in a meaningful way, interlinking open data across the World Wide Web (Berners-Lee, 2006; Bizer et al., 2009a), and business models based on open data (Ferro & Osella, 2013; Janssen & Zuiderwijk, 2014; Zeleti et al., 2014).

The Purpose of this Research

So far, open data research has primarily focused on open data offered by governments or grass-root organisations linking data available on the World Wide Web. However, vast amounts of the world's data reside in data silos within companies. We view this as a missing piece in the cumulative pool of open data. Therefore, to reach the full potential of open data, we explore the possibility of businesses as open data publishers.

The underlying premise of a business does not readily align with the idea of providing open data. However, we have seen tentative traits from companies sharing data in a loosely coupled manner with players from their business ecosystem. We regard this as shared data to nuance the debate of open data.

The premise of a business is to generate returns for its shareholders (Brandenburger & Stuart, 1996; Hillier et al., 2011). The act of opening up data and providing it for free to the public, with no restrictions, does not immediately align with the goal of maximising profits and generating returns for shareholders. Successful operation of a business and the act of opening data thus seems contradictory. Through this study, we take tentative steps in understanding businesses as data publishers. We examine why a company would contemplate sharing data and, thereby, provide an understanding of the justification behind such a decision.

We pose the following research question:

How is external data sharing justified in business?

To answer this question, we have conducted a case study on Trustpilot: a rapidly growing digital business, that facilitates communication between online consumers and companies. Trustpilot helps online consumers take informed purchase decisions, and enables companies to manage their online reputation. The study is exploratory in nature and sheds new light to the rationale behind data sharing, through a business model and platform theoretical frame. Our contribution is an in-depth and empirically based insight into the business rationale for loosely coupled data sharing in a business ecosystem. Furthermore, we open the discussion on why companies are not providing "true" open data.

The thesis structure is as follows: First, the thesis provides an overview of prevalent research directions within open data, the research problem, and the research question. Secondly, the thesis provides the methodological and analytical framework, followed by an analysis of the case study. Lastly, we propose and discuss our findings and suggest directions for future research.

Scope and Limitations

A multitude of directions is found within open data literature. The primary direction for this study is what we refer to as open company data. However, findings are compared to those in other directions in open data research. This study is a single case study on Trustpilot and is limited to an internal view of the empirical setting. Open data is a twofold issue, with both data providers and data users. The focus of this study is on businesses as data providers, and empirical evidence was not collected from data users in Trustpilot's environment. The data user's perspective is not irrelevant, yet, the focus of this study is on the data provider's reason for sharing data.

Thesis structure

Introduction	An introduction to the open data phenomenon, the position of this study and a presentation of the research question.
Research Domain	We identify and address prevalent directions within open data literature and position our study.
Analytical framework	A description of our analytical framework covering the theories data spectrum, business models, and platforms.
Research Direction	A presentation of our research direction followed by our research problem and the research question.
Research Design	A description of this study's research design and underlying philosophical worldview, strategy of inquiry and research method.
Analysis	The analysis of the case: Trustpilot, through the application of our analytical framework.
Discussion	A discussion of key findings and elaboration on its meaning and significance in relation to the research domain.
Conclusion	Concluding remarks and proposal of future research directions.

Research Domain

In this section, contemporary issues in the Open Data literature are presented. We find a significant gap in addressing open data in the private sector.

Initially, open data is defined to establish common grounds. Taking stance in a definition of open data, we identify and present prominent contemporary research on the topic. Directions found within literature include linked open data (LOD) and open government data (OGD) and open data business models (ODBM). This paper suggests a new direction in open data research - open company data. The examination of existing directions offers valuable insights into current issues and offers the reader an understanding of the open data research domain. Findings in this study are contrasted and compared to current research in the discussion section (p. 97).

Open data has conceptually existed for many years, but since open government initiatives have been initiated, combined with the emergence of "big data" as a novel phenomenon (Kitchin, 2014), there has been a renewed interest in the topic.

Contemporary directions within open data research, have been identified through library searches on Google Scholar[2], CBS Library[3], CBS Department of IT Management[4] and journals: MIS Quarterly, Journal of Information Management, Information System Management, Information systems and e-business, Information Knowledge and Systems Management.

Our review of directions in literature is not exhaustive, thus not completely representative of all research done within the open data space, but represents predominant topics within research literature.

[2] https://scholar.google.com/
[3] http://www.cbs.dk/en/library
[4] http://www.cbs.dk/en/research/departments-and-centres/department-of-it-management

Research within the open data domain we cover include:

- Open Data Definition
- Linked Open Data (LOD) and the Semantic Web
- Open Government Data (OGD)
    - Opportunities and Challenges of Open Government Data
    - Value Creation from Open Government Data
- Business Models for Open Data (ODBM)

A review of these topics helps identify a gap in the literature and offers an overview of the research body this study contributes to.

Open Data Definition

Before reviewing directions in the open data research domain, a definition of the term open data is required.

A very precise definition of "open works" is offered by OpenDefinition.org: "*Knowledge is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness.*" (OpenDefinition.org, 2015). The Open Knowledge Foundation translated this definition for data in the Open Data Handbook:

> *Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share-alike.*
> (Open Knowledge Foundation, 2013)

To further expand the understanding of "open data", we examine key features of "open". The Open Knowledge Foundation (2014) supports OpenDefinition.org (2015) and highlights three key elements of openness 1) availability and access, 2) reuse and redistribution and 3) universal participation:

Availability and access: Data must be available as a whole and at reasonable reproduction cost. Access must be free of charge, including fee arrangements and royalty. Data should be machine-readable and downloadable on the internet in an open format (Open Knowledge Foundation, 2014; OpenDefinition.org, 2015).

Reuse and redistribution: Data must be provided under terms which allow for free use, including reuse and redistribution. Modifying and intermixing data with other datasets must be allowed. (Open Knowledge Foundation, 2014).

Universal participation: There must be no discrimination against fields of endeavour, persons or groups. Everyone must be able to use the data. Restrictions in use for commercial, non-commercial or certain purposes must not apply for the data (Open Knowledge Foundation, 2014).

To summarise, open data is characterised by not being limited in access and use and offered free of charge to anyone. From our view, definitions and perspectives given of "open data" describes an extreme state of sharing data, which is rarely present in real world scenarios. We address this proposition in our Analytical Framework (p. 30).

Linked Open Data

Linked open data is one of the more thoroughly covered research direction found in the open data literature. It spawns from worldwide initiatives, that aim to turn the internet into more than web pages: a web of data (Berners-Lee, 2006; Bizer et al., 2009b). This direction in research is particularly prevalent in computer science. Most contributors to the linked open data research direction are also active members of the linked open data community and their contributions to the literature are likely biased in promoting linked open data as a practice.

Linked data is the practice of structuring and linking data on the web. A web which has previously been a web of documents is turning into a web of data (Bizer et al., 2009). The inventor of the World Wide Web, Tim Berners-Lee, has in 2006 outlined four rules for publishing data on the web with the goal to assure inter-linkage (Berners-Lee, 2006):

1. Use Uniform Resource Identificators (URIs) as names for things
2. Use HTTP[5] URIs so that people can look up those names
3. When someone looks up a URI. Provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs so that they can discover more things

[5] HTTP: Hyper Text Transport Protocol. The protocol that runs the World Wide Web.

These are technical specifications which are relevant for computer science, but failing to comply with these rules will result in a missed opportunity to interconnect data and thereby hinder the reuse of information. Berners-Lee (2006) explicitly emphasises the internet's ability to create value through unexpected reuse of information.

Linking Open Data is attempted in the Linked Open Data (LOD) project, which is a grassroots driven community. (LOD Community, 2015) It was founded in 2007 and is supported by the W3C Semantic Web Education and Outreach Group (Bizer et al., 2009a). The project is on a mission to push the "Web of Data" forward by locating open data that currently exists on the internet and turn it into linked data, through the standards outlined by Berners-Lee (2006).

The extent of the project and LOD is continuously documented through the "LOD Cloud Diagram", which is a visual representation of linked open datasets, their size, and connections to other open data sets (Bizer et al., 2009a).

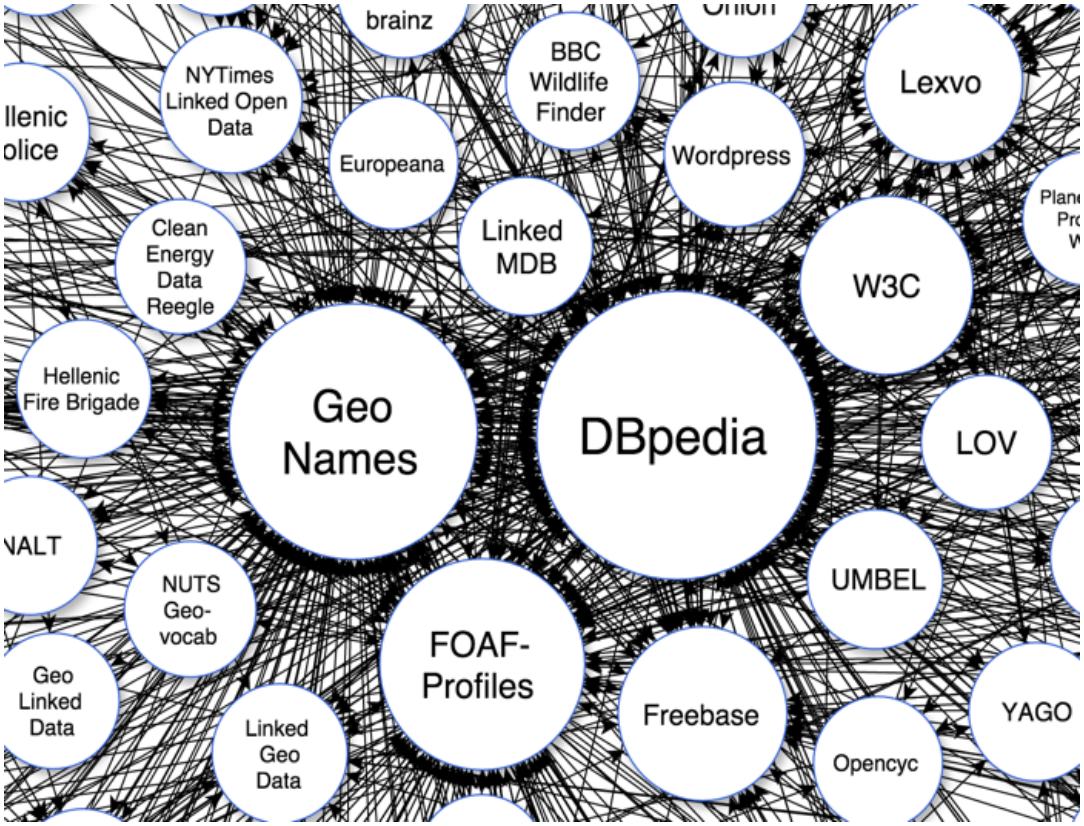


Figure 1: Center of the LOD Cloud Diagram
(See full diagram in Appendix 1)
(Cyganiak & Jentzsch, 2014)

A major data set and example of an initiative in the LOD cloud is DBpedia. DBpedia is a community effort to extract data from Wikipedia.org and its encyclopedia pages. In 2007 DBpedia contained more than 1.95 million data entities, such as people, places, music albums and films. All structured in accordance with the Linked Data Principles described by Berners-Lee (Auer et al., 2007). In 2014 this number grew to 4.58 million (DBpedia community, 2014) and DBpedia has become the central hub for LOD. Other data publishers are linking their data entities against DBpedia (Bizer et al., 2009a). This linking is compliant with Berners-Lee's fourth rule of linked data (Berners-Lee, 2006).

To promote quality linked open data, Berners-Lee (2010) introduced a rating-system to encourage open data providers, especially governments to improve their efforts. The rating system is a star-schema and is described in the figure below:

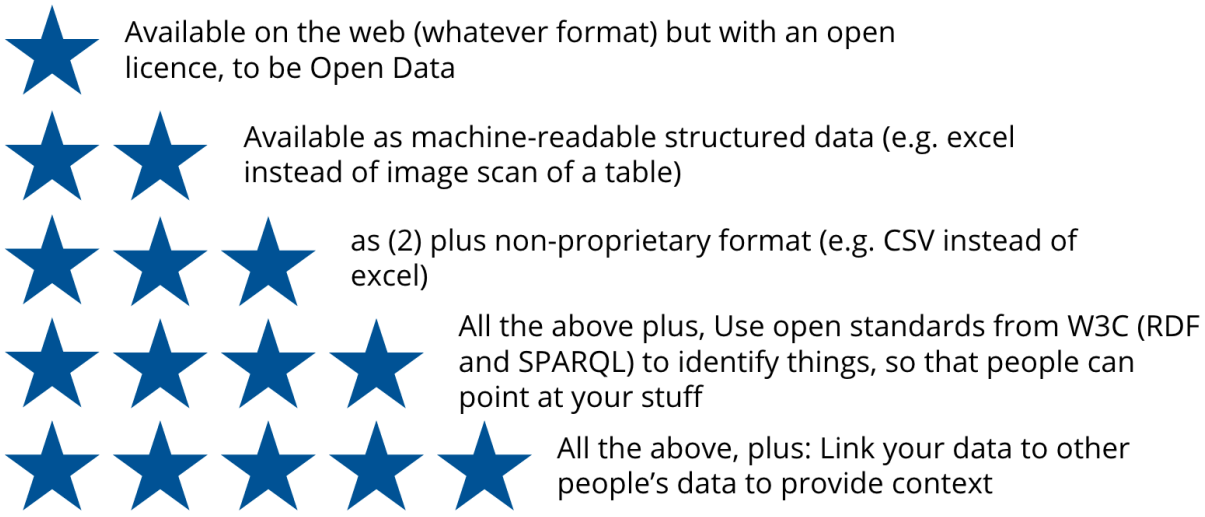


Figure 2: Linked Open Data 5 Stars Model
Adapted from Berners-Lee (2010)

This rating system has been widely acknowledged both by the community and in literature (Bauer & Kaltenböck, 2011; Bizer et al., 2009a; Hausenblas, 2012). With this rating system Berners-Lee hopes to promote better data quality and connectivity of open government data, by turning it into linked open data.

Linked open data is a domain in the open data literature less concerned with the challenges of making data open, and more with the opportunities and challenges of turning web pages into

data and interlinking that data so that it can be utilised fully by everyone across the World Wide Web.

Open Government Data

Open government data (OGD) has had a catalysing impact in the open data research domain. OGD is a government policy that is increasingly being adopted by governments worldwide (Huijboom & Van den Broek, 2011). Throughout the literature open government data plays a central role as a result of the many recently launched open government initiatives.

Governments produce vast amounts of data to accomplish their daily activities (Ubaldi, 2013). The collection and production of data by governments make them excellent cases for opening up their data to the general public in order to pursue the ambition of a transparent government (Huijboom & Van den Broek, 2011).

When Obama entered as President of the United States in 2009, his administration announced a transparency strategy, to ensure extraordinary openness in government: the Open Government Directive (Attard et al., 2015). This announcement was the first of many to come, as many western countries (United Kingdom 2009, Australia 2010, Denmark 2010) have soon after similarly declared their support of transparent governments through means such as open data (Huijboom & Van den Broek, 2011).

Jetzek et al. (2013) find open government data particularly interesting due to the fact that datasets have already been collected, and *"offer value beyond what is captured from the originally intended use."* (p. 2) and then argue that when *"opened up, government data become a common, shared resource (i.e., public good) that is provided by the government."* (p. 2) and thereby share the same view on data as Davies & Bawa (2012, p. 2), who also envision *"data as a common resource."*

Researchers of OGD has particularly shown interest in opportunities and challenges of publishing and using OGD (Janssen et al., 2012; Ubaldi, 2013; Davies & Bawa, 2012; Attard et al., 2015), as well as opportunities for value creation enabled by OGD (2014). These topics are addressed below:

Opportunities and Challenges

The literature on Open Government has identified an array of opportunities for opening data. Those include huge economic growth through a data-driven society and unprecedented levels

of transparency. In this section, we present the opportunities and primary hindrances for realising the potential of open government data.

Opportunities

The primary drivers for opening government data is transparency and economic growth (Janssen et al., 2012; Ubaldi, 2013). Janssen et al. (2012) find that the overarching expected benefits from open data are stimulation of innovation, economic growth, and growth of the knowledge economy. Janssen et al. (2012) also suggest that data in itself is useless. It is how data is used that makes it valuable. The broad scope of data use scenarios thus makes it difficult to calculate the economic returns of investing in open data initiatives.

Transparency

The opening of data is seen as an altruistic act as well as a contributing factor to transparency, which in turn will result in improved accountability, trust, and satisfaction (Janssen et al., 2012). Attard et al. (2015) suggests that transparency hinders corruption, which can have severe economic costs. Ubaldi (2013) suggest that *"OGD can be used to help the public better understand what the government does and how well it performs, and to hold it accountable for wrongdoing or unachieved results"* (p. 4).

Economic growth

Ubaldi (2013) argues that OGD can foster improved decision making, increase public awareness. Additionally, she sees OGD as a significant contributor to economic growth, which has the potential to unlock new opportunities for entrepreneurs and businesses that find new use of the data.

Open data contributes to *"collective intelligence of the public"* (Janssen et al. 2012, p. 10) and empowers the wisdom of the crowd. The wisdom of the crowd is the idea that groups can collectively generate better results than what even the smartest person can achieve on his own (Surowiecki, 2005).

Technical

In addition to the political and economic benefits, open data also brings operational and technical benefits. Such as easier access to data and the ability to integrate public and private data. It is costly to collect and store large amounts of data. The transparency from open data can help to reduce duplicate data and the cost and work effort of multiple parties collecting the same datasets (Janssen et al., 2012).

Challenges

While opening data brings many opportunities, it also implies many challenges. Data providers and data users face both shared and individual challenges (Janssen et al., 2012).

Different categorisations of challenges are found in the literature, but the underlying problems identified are similar. Challenges are amongst others are characterised as political, technical, economic, organisational, cultural and legal (Ubaldi, 2013; Janssen et al., 2012).

Policy challenges

A significant problem relates to policy; lacking procedures and standards on how to deal with open data in governments. These standards and procedures need to be implemented to ensure who owns data, to eliminate issues of copyright and reuse of data (Janssen et al., 2012).

Furthermore, many countries adopt short-term strategies that look for quick wins, opening up small datasets of high quality and neglect future data releases with deteriorating data quality assurance and lack of data updates. Ubaldi (2013) suggests that policymakers should employ long-term strategies, which ensures high availability of open government data over the long run.

Technical challenges

"Government data are often un-harmonised as every public agency has its own set of data, formats and standards." (Ubaldi, 2013, p. 31).

The interoperability of data is a key concern in the domain of open data. Other key technical challenges: include technology infrastructure, privacy measures, information security and integration of OGD tools and applications (Janssen et al., 2012).

To mitigate the issue of un-harmonised data, Ubaldi advises a cautious approach to building centralised OGD portals. *"It is important that this is developed through a collaborative approach, creating ownership and sustainability."* (Ubaldi, 2013, p. 31). Further suggesting that creating a single standardised portal for OGD can be detrimental for governments due to the trade-off between standardisation and allowing for experimentation. A standardised portal implies preparation of the data instead of giving raw access to developers, which Ubaldi (2013) argues is important to support experimentation. Janssen et al. (2012) also propose that merely opening a portal to centralise access to datasets is not sufficient and that broader perspective and a more collaborative approach need to be taken by governments.

Economic & financial challenges

Governments are facing economic challenges that are consequently slowing the pace in the development of OGD initiatives. Governments are concerned with the costs of data collection, provisioning and conversion of large data volumes into re-useable formats. The latter also includes anonymising personal data and converting from proprietary to open standards, which are both resource intensive tasks (Ubaldi 2013).

Organisational challenges

Governments need to develop the right institutional structures to ensure the quality of data, accountability and responsibility, which implies hiring the right people and setting up a governing body. Furthermore, it is important for governments to reach out and collaborate with the wider community to ensure the data with the highest probabilities of being used are released (Ubaldi, 2013). The publishing parties of open data lack policies of publicising data and resources for publicising data (Janssen et al., 2012).

Cultural challenges

Governments have to seed a culture of openness. According to Janssen et al. (2012), a risk-averse culture is a barrier for governments.

> *The belief that making data public disempowers public officials, or makes them more vulnerable as they risk unveiling faults, can at times create an environment among civil servants, or even policy makers, which does not fully support implementation of OGD initiatives.* (Ubaldi, 2013, p. 36).

Further proposing the need of developing government programmes that serve the purpose of changing the attitudes of government officials.

Legal challenges

The opening and sharing of government data imply complex legal issues which need to be addressed. Legal challenges include privacy violation, security requirements, and no or unspecific licenses. (Janssen et al. 2012). Ubaldi (2013) suggests governments develop guidelines and handbooks to cover legal issues.

Value Creation from Open Government Data

A related issue in the research domain concerns how to create value from open data and particularly open government data. A study by Manyika et al. (2013) sought to quantify the value of open data and suggests that $3 trillion in value can be unlocked through open data. Which is largely due to the promise of a data-driven economy that enables growth opportunities.

The available research focused on value creation from open data is based on predictions and hypothesis. Since many open data initiatives are in their infancy, the evidence to support theories is still limited (Jetzek et al. 2014).

While the benefits of open data are recognised Janssen et al. (2012) highlights that: *"The main challenge is that open data has no value in itself; it only becomes valuable when used."* (p. 9). Furthermore suggesting that ROI is impossible to estimate due to the potential applications being extremely hard to predict.

Jetzek et al. (2014) have proposed a model that further conceptualises the realisation that value does not come from data in itself. They suggest four generative mechanisms for realising value: *information transparency, collective impact, data-driven efficiency* and *data-driven innovation.* The value generating mechanisms are driven by a number of enablers: *incentives*, *open access to data*, *data governance*, *capabilities* and *technical connectivity*.

Their proposed framework of archetypical generative mechanisms that transform data into value is illustrated in Figure 3.

The framework conceptualises how value can be derived from open data in an open system, a sharing society: *"an open economic and social system in which information technology is leveraged to empower individuals, corporations, non-profits and governments with data that are shared, reused and transformed to sustainable value through different mechanisms."* (Jetzek et al., 2014, p. 65).

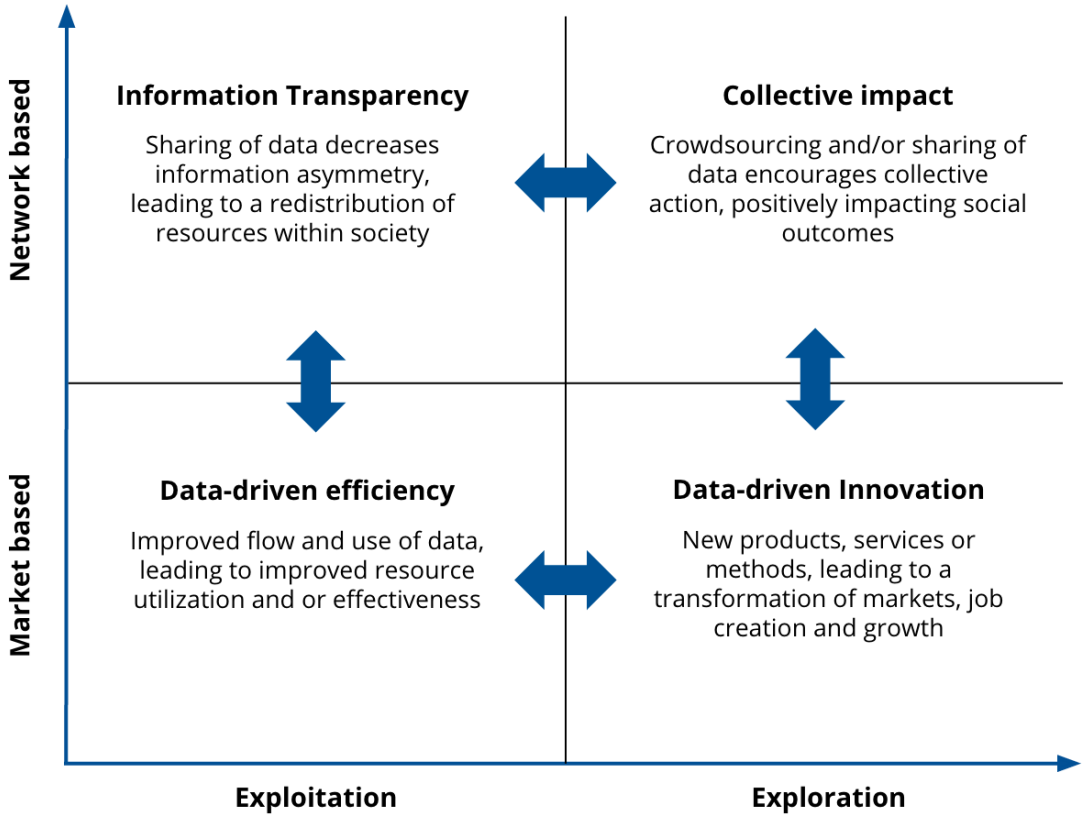


Figure 3: Archetypical Generative Mechanisms of Open Data
Adapted from Jetzek et al. (2014)

The archetypical mechanisms are elaborated below:

Information Transparency: If data is available, accurate, accessible & trustworthy, it generates value by letting individuals acquire knowledge that results in actions.

Collective Impact: When large groups of individuals contribute to a common cause, with positive social outcomes.

Data-driven Efficiency: When stakeholders use data to improve productive efficiency resulting in cost savings and improved quality

Data driven Innovation: When stakeholders use data to improve productive efficiency resulting in cost savings and improved quality

Jetzek et al. (2014) propose that these four mechanisms interact within the sharing society. Furthermore, they suggest that the value generation mechanisms are dependent on the private and public sector working together to provide incentives and opportunity to generate value.

They employ theory from behavioral economics, to explain the drivers of sustainable value generation with open data and suggest that motivation, opportunity, and ability (MOA) positively impacts the outcome of value creation from open data, elaborated in the following:

> *The motivation, opportunity and ability of individuals to use data for value generation are influenced by: the incentives provided; the level of technical and legal openness of data; the maturity of resource (data) governance; the general data-related capabilities in society; the technological maturity and prevalence.*
> (Jetzek et al., 2014, p. 79).

In addition to the above, Manyika et al. (2013) define the three value levers as *1) Improved decision making, 2) New offerings* and *3) Accountability.* These propositions share traits with Jetzek et al.'s (2014) proposal of archetypical mechanisms: Improved decision making caused by data-driven efficiency, new offerings resulting from data-driven innovation. And information transparency and collective impact contributing to accountability, which is especially relevant in the context of value generation from Open Government Data.

The proposed perspectives highlight the significance of not only opening government data but also ensuring mechanisms for value creation through the opened data.

Business Models for Open Data

As established, there is enormous economic potential hidden in open data. Several authors from the practising and scholarly communities have described and suggested possible business models for open data. In this section, we present predominant and differing views on Open Data Business Models (ODBM).

In 2014 a systematic review was made by Zeleti et al. (2014) to analyse and conceptualise the existing business models made possible by open data. They argue that most work done on open data business models has been carried out in the practise community. Their review resulted in the identification of 15 open data business models from the literature. These 15 business models are grouped into five categories based on the value discipline, defined as *"the ways in which businesses can differentiate itself from competitors"* (Zeleti et al., 2014, p. 9). The five categories are: 1. Razor Blade, 2. Indirect Benefit, 3. Cost Saving, 4. Premium & 5. Freemium.

Ferro & Osella (2013) offers another categorisation of open data business models. Their categorisation is oriented towards business models based on the re-use of open Public Sector Information (PSI). The ODBM identified are listed below:

Business Model	How it works
Premium	Service / product offering in exchange for payment
Freemium	Basic service/product offered for free; payments for a higher feature set
Open Source Like	Free and open data sets and payments for supplementary (value adding) services.
Infrastructural Razor and Blades	Data sets stored free & accessible through API. Payment cross-subsidization: revenues through offering value-added services. Payments occur on an "on demand" basis based on computing power needs for buying party.
Demand-Oriented platform	"One-stop all-in-one data shop". Payments for advanced services & refined data sets.
Supply-oriented platform	Requires intermediaries; Public Sector Information (PSI) holders are charged; not developers
Free as branded advertising	Providing useful data for free as a means of persuading an audience towards a brand / company. No direct revenue, serves as a promotional channel.
White label development	Companies who want to use PSI as an attraction medium, but don't have the in-house competencies to do so; utilisation of advertising factory to get a turnkey solution.

Table 1: Open Data / PSI Business Models
Adapted from Ferro & Osella (2013) and Janssen & Zuiderwijk (2014)

Janssen & Zuiderwijk (2014) argue that the above eight archetypical ODBM are viewed from a revenue point of view, and *"[...] do not capture the user point of view, networks, and other aspects of business models"* (p. 698) and can better be labeled as revenue models.

On this basis Janssen & Zuiderwijk (2014) performed an empirical study on open data infomediaries in Netherlands to analyse existing business models. They propose six infomediary business models, for connecting data providers with users:

1. Single-purpose apps
2. Interactive apps
3. Information aggregators
4. Comparison models
5. Open data repositories
6. Service platforms

Janssen & Zuiderwijk (2014) propose that the business models differentiates on two main variables:

1. Level of access to data: raw vs. refined data (trade off between complexity and ease of use)
2. Level of dialogue: simple presentation of data vs the opportunity of user-generated content & dialogue with users and providers.

Suggesting that their models can be used as abstract reference models, and detailed analysis of business models have to be made to gain better insights into open data business models.

The Open Data Institute (2016b) propose that the publishing of open data can be supported through three general business models, which is represented in Table 2:

Business Model	How it works
Freemium	Basic service is provided for free & value added services are charged for. Example access to information is free; analytics combined with other data costs (Bonina, 2013; Open Data Institute, 2015).
Cross subsidy	Charging different groups different prices. Implies using revenue from existing data services, to develop new services (Bonina, 2013; Open Data Institute, 2015).
Network effects	Through collaboration with partners costs of storing and maintenance are reduced, and value is gained from the collective extended use of data (Bonina, 2013; Open Data Institute, 2015)

Table 2: Business Models for Publishing Open Data
Adapted from Open Data Institute (2016b)

Placr and Mastodon C are examples of companies that employ the freemium and the demand-oriented platform business models respectively (Bonina, 2013).

Placr[6] a UK based data company. They have utilised OGD to provide a centralised source of transport information (busses, trains, etc.). They employed a freemium model, "cash and carry" and a "retail model". By publishing a single API[7], they offer data-as-a-service and their retail model implies selling authorities a consumer facing web app, that allows customisation (Bonina, 2013).

Mastodon C[8], a UK data-expert company founded in 2012. Offer two offerings 1) a self-managing auto-scaling big data platform platform-as-a-service to process big data analytics. 2) Consulting: helping organisations evaluate their use of data (Bonina, 2013).

Open Data Business Models are evolving, and many have still not seen the light of day. Bonina (2013, p. 24) suggests that *"from an economic perspective, it is not clear how to monetise or create economic value from an open, public good type of resource, that will be sustainable over time."*

From our review of open data business models, we find that business models are constructed at two levels (see Figure 4).

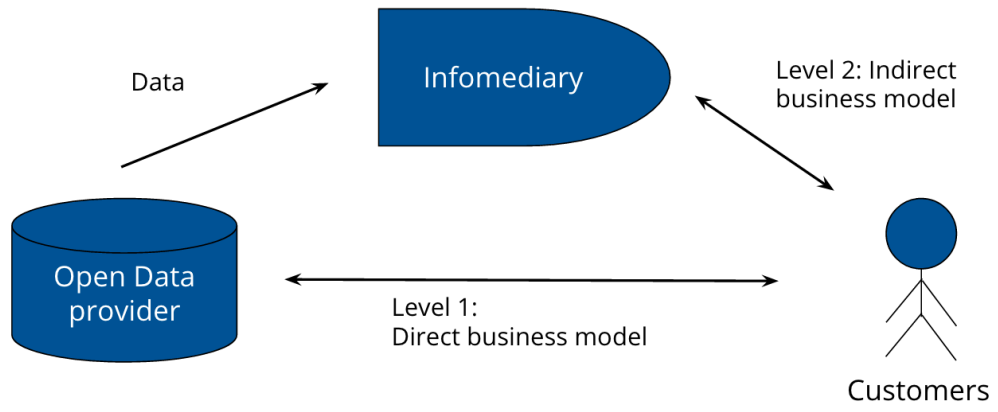


Figure 4: Direct and Indirect Open Data Business Models
Authors' interpretation

[6] Placr: placr.co.uk
[7] API: Application Programming Interface
[8] Mastodon C: http://www.mastodonc.com/

The first level is where the data publisher attempts to base a business on a data offering. The second level is the infomediaries, leveraging the vast amounts of open data and place themselves between data providers and users. The majority of research addresses the second level business models (Janssen & Zuiderwijk, 2014; Ferro & Osella, 2013; Zeleti et al., 2014). While what we view as first level business models, has primarily been addressed by the Open Data Institute (2016b) and Zeleti et al. (2014).

Research Domain Summary

Throughout the literature the focus can be divided into three broad areas: 1) finding value in open government data, 2) basing business models on open data and 3) addressing technical challenges in providing data in an easily accessible and processable format.

Investments in open data have been addressed in various ways: opportunities, challenges, value generation and business models derived from open data.

We have identified that the literature within the open data domain has been scoped mostly towards open data in governments (OGD), the use of open government data and the mission to structure open data on the World Wide Web through linked open data (LOD) projects.

Recurring issues within open data can be split into two categories:

- How to extract value, through new business models and innovation
- How to institutionalise open data and build open data infrastructures

Research schools that have contributed to the domain include information systems, computer science, economics and e-government research.

Gap in Research

Research within open data is yet very limited and few papers are based on actual empirical studies. This leaves significant room for research and new learnings within the open data domain.

Our review of the open data research, reveals that issues of using open data and issues of providing open data have been addressed. Most research has dealt with scenarios where governments open data and companies in society make use of it. This study explores new

territory and examines companies as the supplier of open data. In a majority of articles, reports and strategy papers open government data is used as a synonym for open data (Heimstädt et al., 2014). However, Heimstädt et al. (2014) highlights that open data may as well originate from commercial, academic or other sectors. We explore the commercial origin of open data and refer to this as O*pen Company Data.* Although we term it "open" company data, we perceive "open" as an ideal state, and that there is a spectrum from closed to open. In this research, we focus on why companies would approach the open end of such a spectrum.

Literature on open data business models, opens the discussion of basing a business on a free and public resource. In many cases businesses in the open data domain are infomediaries that creates a business through utilisation of open data. The Open Data Institute (2016b) suggests that businesses should consider opening some of their data, to benefit from one of the three business model options: freemium, cross subsidy or network effects. We find this prompt particularly interesting and seek to examine the potential of opening data for companies. Thus we carry out this study within the open company data direction. Governments find their justification for opening data in the potential for transparency, economic growth and technical advantages, but how do companies relate to these potential upsides? Governments are rushing to open their data. We question if businesses will find a reason to do the same and take the first tentative steps in the research of companies as data publishers.

Analytical Framework

Opening data is an act of sharing. It is an act of taking part in a network of surrounding stakeholders and partners. To gain an understanding of potential incentives and hindrances for companies to open data, we employ theoretical concepts such as the data spectrum, business models, and platform theory. These theories play a central role in our analysis of the case company, Trustpilot.

The data spectrum framework (Open Data Institute, 2016a) allows for analysis of data sharing initiatives, to determine the openness, ranging from closed to open data and contributes to a nuanced discussion of data accessibility.

The application of business model theory (Osterwalder & Pigneur, 2010) offers an understanding of how a business works. These insights advance an understanding of the implications, data sharing has on value creation and value capture.

Substantial value can be generated in the network surrounding a data publisher. To get an understanding of business in the perspective of business ecosystems, we employ platform theory (Iansiti & Levien, 2004a; Baldwin & Woodard, 2009; Gawer, 2010; Cusumano 2010; Hagiu & Wright, 2015; Shapiro & Varian, 1998). Platform theory offers important concepts such as network effects and complementors, that helps to understand the justification of data sharing. This perspective complements business model theory with a broader and more strategic perspective on data sharing.

Data spectrum

We adopt the data spectrum developed by Open Data Institute (2016a) to understand and nuance the way open data is discussed. The data spectrum suggestively *"[...] helps you understand the language of data"* (Open Data Institute, 2016a).

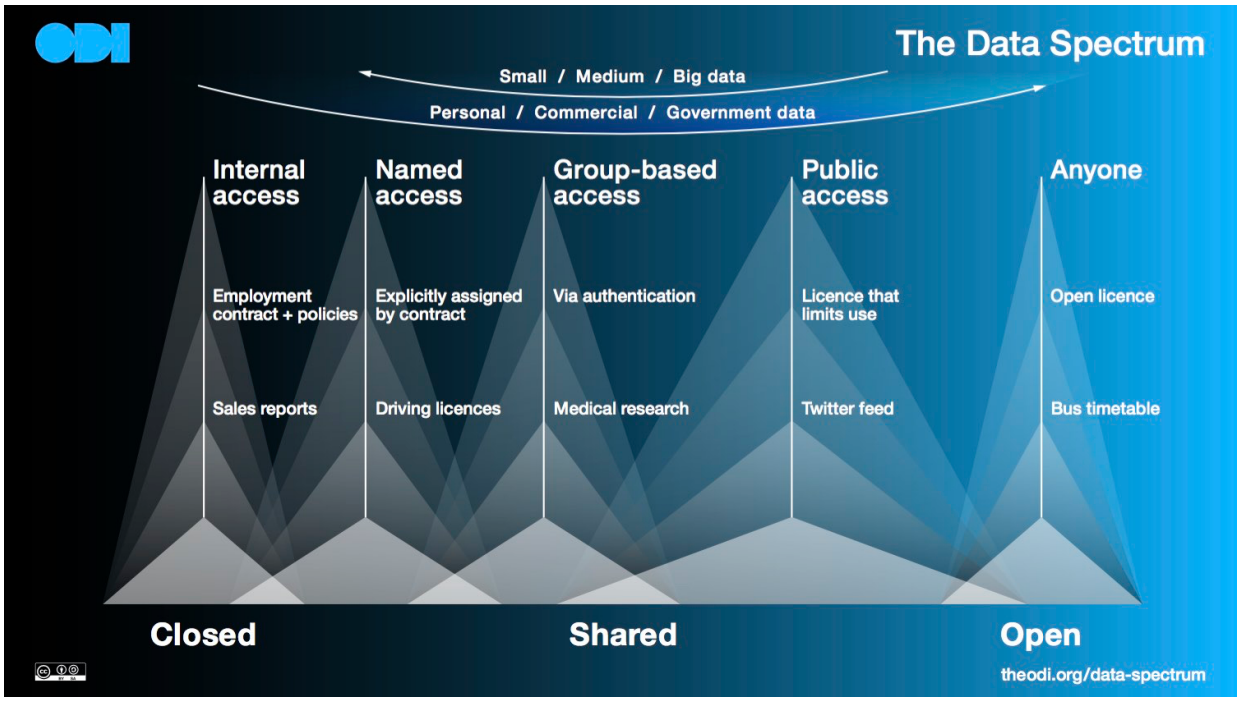


Figure 5: The Data Spectrum
Open Data Institute (2016a)

To understand businesses and their approach to open data, we need not to just discuss whether open data is offered or not. There is a need for a more refined identification of to what degree they move from closed data towards open data. Along the data spectrum, there are three different terms that categorise data access. Those are: closed, shared and open. Besides offering these three terms for data access, the data spectrum suggests that data is not easily categorised into one of these three buckets (Broad, 2015). Closed, shared and open data has shades, meaning that sometimes a dataset is accessible under conditions somewhere in between.

Data with closed access refers to *"Data that can only be accessed by its subject, owner or holder."* (Broad, 2015).

Data with shared access refers to data which is neither completely closed nor completely open. That is, data which is not only accessible by its original owner, but neither adheres completely to the definition of open data. Shared data might only be available to specific people or organisations, which refers to *named access* (Broad, 2015). Or it might only be available to groups which meet certain criteria, which is referred to as *group-based access.* Finally, it might be available to everyone, but not under a license which complies with the "open" definition, thus it is referred to as *public access.*

Even though this framework suggested by the Open Data Institute (2015) has not been backed by scientific research, the Open Data Institute is an organisation deeply involved with open data initiatives, back by the government of the United Kingdom. They further express how the framework has been co-developed and evaluated through workshops with their audience (Broad, 2015). Before encountering the framework, we found ourselves looking for a way to describe what lies between the two extremes of closed and open data. In this study, the framework is applied in an empirical setting, where Trustpilot, a commercial organisation, is offering external actors certain data access privileges. Additionally, we assess the usefulness of the applying the data spectrum in this empirical setting.

Business Models

This section introduces the notion of a business model, value creation and value capture. Hereafter we present a specific framework for analysing business models: the business model canvas.

A business model is the logic of how value is created, delivered and captured (Osterwalder & Pigneur, 2010). It describes how a business is able to serve its customers and provide them with value to generate revenue. Most mature businesses are settled on their business model, which means they have found a contemporary formula of how to create and capture value. In today's digital business environment, data plays a significant role in the business. However, it is unclear how data is embedded in the business model and whether it plays a critical role in creating value for customers. Posing questions such as: Is the possession and usage of data an essential component in the business model? How would a business model destabilise if the company released its data to the public? Which customer segments could benefit from the release of that data? Could new customer segments emerge? Could it lead to new revenue streams?

To address the implications data sharing brings to business, it is necessary to understand the business model. Therefore, we apply a business model frame to gain deeper insight into the effects and implications data sharing has on the business model.

Different definitions of a business model have been proposed throughout the theoretical literature. Theorists have in recent years based their definitions on: collaborative transaction, revenue sources, product value architecture or value propositions (Al-Debei et al., 2008).

In an extensive literature review, Zott et al. (2011) argue that no consensus in business model literature has yet been found, implying that no industry-wide definition has been adopted.

Zott et al. (2011) highlight the significance of having common grounds for understanding a business model, as it is invaluable to understand, not only how businesses capture value, but also create value. An idea that is significant in relation to open data, as open data has the potential to create significant value that might not be readily captured by businesses.

Value Creation and Capture

The belief that open data is a potent value creator, makes it necessary to understand the notions of value creation and value capture. Opening data in accordance with the definition of open data has inherent implications (see Research Domain, p. 14). In essence opening data corresponds to giving a raw material away for free and with unrestricted application, consequently no direct value is immediately captured.

Pagani (2013, p. 618) interprets Bowman and Ambrosini's (2000) proposal of value creation and value capture as *"the contribution to the utility of the final good or service to end users and distinguish it from value creation defined as the difference between revenue and cost retained by the firm"* (Pagani, 2013, p. 618). A perspective adopted in this paper.

In the digital world, where Open Source Software (OSS) is becoming more and more popular, companies and developers of OSS are creating much more value than they are capturing (O'Reilly, 2012). Another example is Google: by making all the world's information accessible they are creating tremendous amounts of value, but only capturing a tiny fraction of the value through their ads and other means of monetisation. "Tiny" in this context is of course relative.

Tim O'Reilly (2012), populariser of the term Open Source, argues that the Internet pioneers have created vast amounts of value, and have not retained much of it. He argues that the pioneers behind the technological protocols of the WWW, who have made the "Internet"

possible - have generated an enormous amount of value, however, they have not captured any significant value from it. Opening up data could bring the same implications; companies opening data are at risk of missing out on capturing the value creation they contribute to.

This suggests that there is a fundamental clash between open data in governments and open data in the private sector: the governments are interested in society at large. It therefore makes sense to create more value, as governments will capture it through the value generated by the companies, for instance through taxes. This is in contrast to companies, who are interested in maximisation of shareholder value (Brandenburger & Stuart, 1996; Hillier et al., 2011).

The application of a value creation vs value capture lens, will allow us to better distinguish between where the case business is creating value and capturing value. Which in turn will allow better explanation and discussion in regards to the implications of opening company data.

Business Model Framework

Since Zott et al.'s literature review in 2011, Osterwalder and Pigneur (2010) have received much attention for their new conceptualisation of a business model: the Business Model Canvas. A conceptualization that has been applied broadly both in literature and amongst practitioners.

Being widely understood and accepted, we employ Osterwalder & Pigneur's (2010) definition of a business model: *"A business model describes the rationale of how an organisation creates, delivers, and captures value."* (Osterwalder & Pigneur, 2010, p. 14). The motivation behind their framework is explained as:

> *We need a business model concept that everybody understands: one that facilitates descriptions and discussion. We need to start from the same point and talk about the same thing. The challenge is that the concept must be simple, relevant and intuitively understandable, while not oversimplifying the complexities of how enterprises function.* (Osterwalder & Pigneur, 2010, p. 15)

They propose that a business model is comprised of 9 inter-related components that allow a company to generate revenue, as illustrated in Figure 6. The components are referred to as the business model's "building blocks".

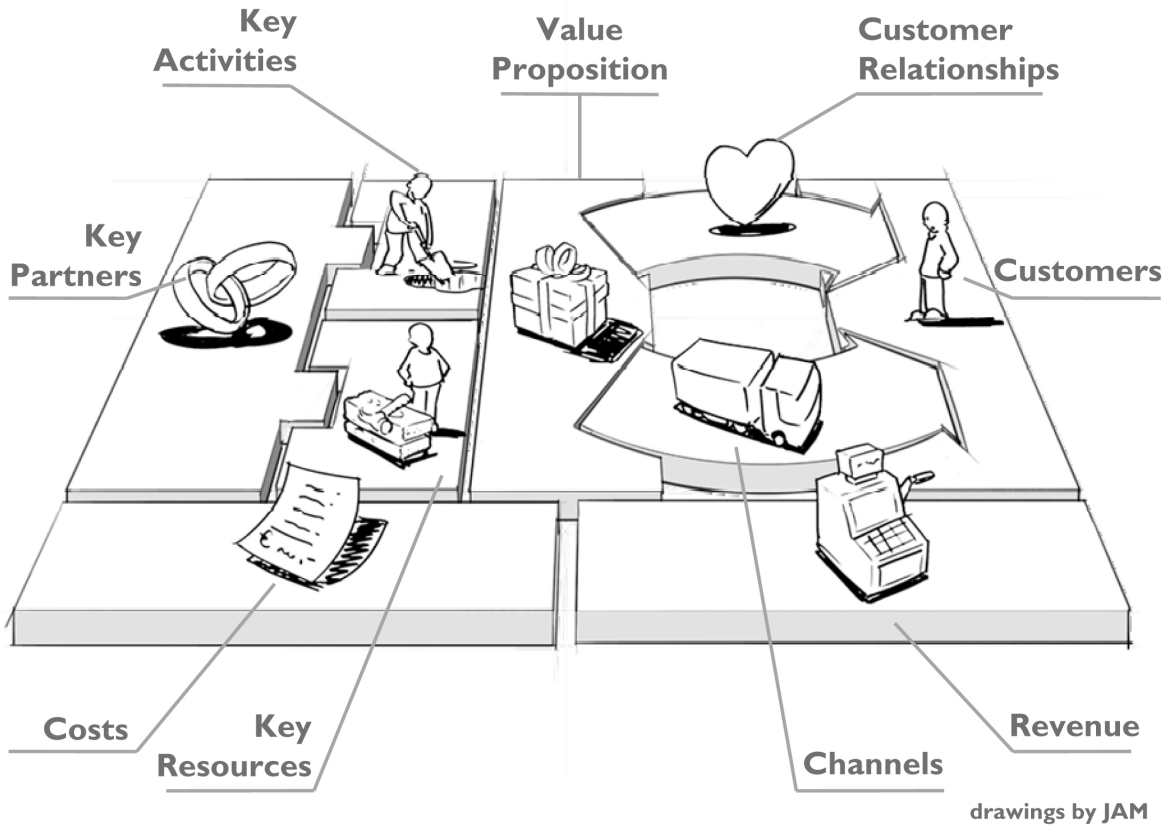


Figure 6: Business Model Canvas
(Osterwalder & Pigneur, 2010)

1) Customer Segments

"The customer segments defines the different groups of people or organisations an enterprise aims to reach and serve." (Osterwalder & Pigneur, 2010, p. 20)

Osterwalder and Pigneur argue that organisations must make *"a conscious decision about which segments to serve and which segments to ignore."* (2010, p. 20). And design their business model on the understanding of the segments they are serving.

For example, consumer electronics companies often target *mass markets* with their projects, whereas the watchmaker Rolex targets a *segmented* market of high net worth individuals.

2) Value Proposition

"The value proposition describes the bundle of products and services that create value for a specific customer segment" (Osterwalder & Pigneur, 2010, p. 22)

Convenience, design and brand amongst others are the factors that constitute the value proposition of companies (Osterwalder and Pigneur 2010). According to Osterwalder and Pigneur a value proposition *"seeks to solve customer problems and satisfy customer needs with value propositions."* (2010, p. 16)

A value proposition is the value delivered to the customers. An example hereof is Salesforce's utilisation of the SaaS (Software-as-a-Service) distribution model, that allows clients to host CRM software online. A model that prevents the need for on-premise infrastructure around Salesforce solutions. This is a value proposition that enables cost savings for customers (Osterwalder & Pigneur, 2010).

3) Channels

"The channels describes how a company communicates with and reaches its customer segments to deliver a value proposition." (Osterwalder & Pigneur, 2010, p. 26).

Osterwalder and Pigneur (2010) propose the channels as how the company interfaces with its customer: the touch points. They propose that channels consist of 5 distinct phases and highlight the significance of finding the right mix, to bring a value proposition to market. The 5 phases are 1) Awareness, 2) Evaluation, 3) Purchase 4) Delivery and 5) After sales. These 5 phases can either be owned or operated by partners, for example, a company might outsource its customer support to a partner, and furthermore the awareness phase is in many cases outsourced to marketing and PR agencies. They argue that the most important factor is to find the right mix to optimise the customer experience and maximise revenues (Osterwalder and Pigneur, 2010).

4) Customer Relationships

"The Customer Relationships describes the types of relationships a company establishes with specific Customer Segments" (Osterwalder & Pigneur, 2010, p. 28).

Osterwalder & Pigneur (2010) propose that the choice of a customer relationship model has great influence on the overall customer experience. For example, a business may choose a self-service experience, where the customer is provided with the means to help themselves. Or

they might choose dedicated personal assistance, which implies a great emphasis on building relationships between representatives of the company and their customers. In B2B sales, personal assistance is often seen in the form of key account managers and customer success managers, who build and maintain customer relationships. Different customer relationships can co-exist within a single business and thus, are not mutually exclusive.

5) Revenue Streams

"The Revenue Streams Building Block represents the cash a company generates from each Customer Segment" (Osterwalder & Pigneur, 2010, p. 30)

Osterwalder and Pigneur (2010) use the metaphor of a human body to describe revenue streams: *"If customers comprise the heart of a business model, Revenue Streams are its arteries."* (p. 30). They propose that a company must constantly ask itself how much customers are truly willing to pay. Further proposing that a business model can involve two different types of revenue streams:

1) Transaction revenue through one-time customer payments.
2) Recurring revenues resulting from ongoing payments

Revenues can, for instance, be generated through selling assets - which implies selling ownership rights to a physical product. Amongst other ways are usage fees, licensing fees and subscription fees. The latter being a revenue model employed by services such as Netflix and Spotify. In such an arrangement a user is granted access to their service through payment of a monthly subscription fee.

Osterwalder and Pigneur (2010) suggest that companies can employ two types of pricing mechanisms:

1) **Fixed Menu Pricing:** Predefined prices based on static variables.

- Example: A mobile data plan, where pricing is fixed and depending on the amount of data the customer wants to pay for.

2) **Dynamic pricing:** Prices change based on or market conditions through negotiation, yield management, real-time market and auction.

- Example: Auctions on eBay.com; the pricing is determined through bidding from potential customers. Another example is airline fares, which are dynamically based on supply and demand (e.g. availability of seats, season).

6) Key Resources

"The Key Resources Building Block describes the most important assets required to make a business model work." (Osterwalder & Pigneur, 2010, p. 34)

Osterwalder and Pigneur (2010) propose that all business models need key resources, as they *"allow an enterprise to create and offer a value proposition, maintain relationships with customer segments and earn revenues."* (p. 34)

Further suggesting a typology of resources: *physical*, *financial*, *intellectual* and *human*.

Physical resources are amongst others referred to as buildings and manufacturing facilities. Intellectual resources include brands, trademarks and customer databases.

An example of a company that has designed its business model around its intellectual resources is Qualcomm. An American chipset manufacturer that revolves around patented microchip designs that are licensed to manufacturers, earning them billions in revenue without being the ones who produce the physical chips. The patents owned by Qualcomm are, however, made possible by their human resources. Thus, Qualcomm is a type of company that is heavily reliant on its human and intellectual resources.

7) Key Activities

The seventh building block is described by Osterwalder and Pigneur (2010, p. 36) as *"the most important things a company must do to make its business model work."* They suggest that key activities can be divided into three categories:

1) Production
2) Problem solving
3) Platform/Network

Arguing that a PC manufacturer such as Dell must produce hardware, hence supply chain management is a key activity. Whereas for a consulting company such as Deloitte, a key activity is problem-solving. Lastly, a company like eBay has a business model with key activities associated with development, management and marketing of their marketplace (Osterwalder & Pigneur, 2010).

8) Key Partnerships

"The Key Partnerships Building Block describes the network of suppliers and partners that make the business model work." (Osterwalder & Pigneur, 2010, p. 38)

Many businesses see partnerships as an essential part of their business. Osterwalder and Pigneur (2010) distinguish between 4 types of partnerships:

1) Strategic alliances between non-competitors,
2) Coopetition: strategic partnerships between companies,
3) Joint ventures for development of new businesses,
4) Buyer-supplier partnerships to assure supplies.

Furthermore, motivations for entering partnerships are presented as: optimisation and economy of scale, risk reduction and acquisition of specific resources or activities. Partnerships thus allow businesses to focus on what they are best at and allows proper and optimal resource allocation.

An example of a partnership that reduced risk and allowed optimal resource allocation, is the case of the TPCA alliance, that Peugeot, Citroen and Toyota formed in 2002. These automobile manufacturers partnered up to reap the benefits of a shared platform for their entry level micro cars, known as the Toyota Aygo, Citroen C1 and Peugeot 107.

9) Cost Structure

"The Cost Structure describes all costs incurred to operate a business model." (Osterwalder & Pigneur, 2010, p. 40). In Osterwalder and Pigneur's (2010) terminology, a distinction of whether the organisation is cost-driven or value-driven has to be made. Ryanair is an example of a business that has based its entire infrastructure around the accommodation of a low-cost structure.

A business cost structure can either be:

a) **Cost-driven**
   This cost structure aims to keep costs down, which in many cases imply automation, low price value propositions and high levels of outsourcing.

b) **Value-driven**
   The value-driven cost structure implies that companies are less concerned with the cost, but has a strong focus on value maximisation. In contrary to a cost-driven structure, such infrastructure would focus on providing a high level of service or a product of high quality. Osterwalder and Pigneur (2010) exemplify this cost structure in the case of luxury hotels, lavish facilities and exclusive services.

Whether a cost structure is value driven or cost driven depends on the characteristics of the cost structure, which according to Osterwalder and Pigneur (2010) are defined by fixed costs, variable costs, economies of scale and economies of scope.

Critique of the business model canvas

Although Osterwalder and Pigneur (2010) are widely recognised for having brought a common terminology to the world of business models, they have received criticism for their conceptual model. Kraaijenbrink (2012) criticises the Business Model Canvas for ignoring external factors such as competition. Further suggesting that the building blocks are based on different layers of abstraction, which can lead to focusing too much on certain areas of the canvas. Lastly, Coes (2014) in a dissertation that addresses the strengths and weaknesses of the business model canvas suggests that the strategic objective of the organisation is disregarded - which also contributes to the neglection of external factors.

In this study, the concept of a business model is used to describe the inner workings of Trustpilot. By establishing the current business model, we will analyse the role of data in the business model. This analysis will offer basis for an understanding of how data sharing is justified in business.

Although we are not conducting a strategy analysis of Trustpilot, we mitigate the issues of the missing strategic perspective of the business model canvas, by employing platform theory.

Platform theory

To expand the view of the business beyond the mechanisms addressed in business models, we complement our analysis with an externally focused strategic perspective: platforms in business ecosystems. Platform theory serves as a theoretical concept for understanding the strategic aspects of data sharing towards the surrounding actors in a network of businesses.

We are curious to learn how businesses can extend their thought of business beyond traditional sales of products (Gawer, 2010), thus moving from a product strategy to a platform strategy (Cusumano, 2010). We question if data can be the core of a platform and opening it can help strengthen the data suppliers position in a business ecosystem.

Cusumano (2010) describes how businesses used to simply build products and deliver services. Today businesses still build products and deliver services, but an increasing number of companies are now successfully running businesses based on platforms. To stay competitive and maintain a competitive position Cusumano (2010) suggests that companies should not just focus on product, but also focus on platforms.

Platforms can be described as *"[...] reuse or sharing of common elements across complex products or systems of production*" (Baldwin & Woodard, 2009, p. 22). This is just one of many definitions of platforms. Researchers have diverse understandings of what a platform is. Product development researchers, technology strategists and industrial economists among others have used the term in different ways (Baldwin & Woodard, 2009; Gawer, 2010). We combine the mentioned perspectives for the purpose of this thesis.

Types of Platforms

Baldwin and Woodard (2009) suggests that platforms exist in the form of product lines within a firm, across firms as multi-product systems and in the form of multi-sided markets. However, they see common characteristic in the architecture of different types of platforms: *"certain components remain fixed over the life of the platform, while others are allowed to vary in cross-section or change over time"* (Baldwin & Woodard, 2009, p. 23).

Gawer (2010) attempts to clarify the diversity of the term "platform" through identification of four different types of platforms found through literature. Those are internal platforms, supply-chain platforms, industry platforms and multi-sided platforms. The focus in this study will be on industry and multi-sided platforms as we examine whether the motivation to

become an industry and multi-sided platform affects their justification for external data sharing.

Internal Platforms

Internal platforms are used in product development within a single company. Some term this type of platform as product platforms (Gawer, 2010). It is defined by products serving as a platform for modification. The products are designed to be modifiable to meet the needs of different customer groups. Variants of the product arise through additions, substitutions, or removal of features, which ultimately can enable mass-customisation (Pine, 1993).

Companies reportedly benefit from internal platforms through fixed-cost savings and product development efficiency. These benefits are achieved through re-usable components and the ability to produce a large number of derivative products (Gawer, 2010).

The same benefits can be achieved across a supply-chain in what Gawer refers to as Supply-Chain Platforms (Gawer, 2010).

Industry Platforms

Industry platforms are much like internal platforms and supply-chain platforms, in the sense that they are the foundation for creating and providing various products, services or technologies from the same matter at core. Industry platforms are regarded in the context of an industry or business ecosystems (Iansiti & Levien, 2004a). Business ecosystems can be viewed as an array of businesses (Gawer, 2010), that co-exist and offer related products or services. The term business ecosystems describe the business world we see today, where businesses are no longer operated along traditional supply chains or limited to a single industry. According to Iansiti and Levien (2004b) companies, collaborate across industries and thousands of companies coordinate, to deliver value propositions to end customers. A business ecosystem includes customers, competitors, makers of complementary products, companies that are outsourced to, firms providing the underlying technology of the business, regulatory agencies and media outlets (Iansiti & Levien, 2004b). Gawer suggests that in such ecosystems the products and services must function together as part of a technological system (Gawer, 2010).

Industry platforms also referred to as external platforms (Gawer & Cusumano, 2014) are leveraged across business ecosystems. An industry platform is usually developed by one or a few firms and offered to other firms to build products, services and technologies upon (Gawer, 2010). Firms building on top of the platform does not necessarily serve the same end

customers and can provide unique offers which are channelled towards a very particular customer group. A large number of loosely coupled firms, not necessarily competitors, thus create innovations that are based on the platform (Gawer, 2010). Microsoft's Windows is an example. It is an operating system which allows other firms to produce software for computers based on a common platform: Windows. A platform like Windows enables software developers to focus on their applications and not on building the backbone system running their application.

Gawer (2010) highlights a key feature of industry platforms to be the enablement of innovation, through a wide network of firms participating and creating value in the industry ecosystem. A platform provider is surrounded by multiple complementary businesses, that base their products on the platform.

Through increased participation in the industry ecosystem and use of the platform, barriers to entry are made for new platforms (Gawer, 2010). As more firms participate and build upon the platform, structures and dependencies are created. It leads not only to switching costs but also entail network effects (Shapiro & Varian, 1998) as elaborated in the next section.

In essence, platforms provide the means for innovation and are highly valued in technological domains, where the reuse of components has shown extremely valuable. Since platforms are often greatly affected by network effects, it leads to few big players at the top of the ecosystem as seen in many technology domains. The consequence of platforms thus becomes limitations in competition, since new platforms are facing an uphill battle in turning the masses of the ecosystem from one platform to another, also known as *tipping* (Gawer & Cusumano, 2002).

In this thesis, we see data sharing through the lens of an industry platform. Thus, data is seen as the common element, utilised in various digital products and services.

Multi-sided Platforms

Industry platforms and multi-sided platforms (MSPs) often two sides of the same coin. In this section, we present mechanisms necessary to understand multi-sided platforms. We employ definitions proposed by Shapiro and Varian (1998) and Hagiu and Wright (2015).

Hagiu and Wright (2015) argue that in today's literature definitions of multi-sided platforms are either too vague or excessively specific. They propose that a multi-sided platform at its most fundamental level, implies interaction between two distinct groups and suggest that

MSPs have two key features 1) *"They enable direct interactions between two or more distinct sides", 2) "Each side is affiliated with the platform"* (p. 163). This implies that the two parties have direct access to each other through the platform and that they are affiliated through platform-specific investments, which include monetary fees or an opportunity cost of participation.

Network effects

A key characteristic of multi-sided platforms is the existence of network effects. Network effects amongst others exist as a) direct network effects and b) positive indirect network effects. Direct network effects arise when consumers value the product more as a result of similar consumers using the product as well (Shapiro & Varian, 1998). Positive indirect network effects entail that the value of the platform amongst distinct user groups increase, as the overall number of participants in the network increase (Hagiu, 2006).

Switching costs

Burnham et al. (2003) define switching costs as: *"the onetime costs that customers associate with the process of switching from one provider to another."* (p. 110). And propose eight types of switching costs: 1) economic risk costs, 2) evaluation costs 3) set up costs, 4) learning costs, 5) benefit loss costs, 6) monetary loss costs 7) personal relationship loss costs and 8) brand relationship costs. These eight types of switching costs can all be placed into three distinct high-level categories: 1) procedural switching costs, 2) financial switching costs and 3) relational switching costs. It is important to note that switching costs are not limited to financial costs, but include intangible costs perceived by the customer.

Lock-in

The creation of customer lock-in is in many companies a desirable outcome. Lock-in implies that customers of a business will find it difficult to stop their engagement or switch to a new product. Another form of lock-in is vendor lock-in. Shapiro and Varian (1998) suggests that lock-in *"arises whenever users invest in multiple complementary and durable assets specific to a particular information technology system."* (p. 12) They propose that lock in arises from 1) contractual commitments, 2) durable purchases, 3) brand-specific training, 4) information

and databases, 5) specialised suppliers, 6) search costs and through and 7) loyalty programmes (Shapiro & Varian, 1998, p. 117).

The above-described platform dynamics of network effects, switching costs and lock-in allows for a closer examination of the consequences of data sharing in a business that serves several distinct customer groups - as is the case in this study.

The Analytical Framework

Our analytical framework covers theoretical concepts, which offers perspectives through which business can be understood. This analytical framework is applied to understand business and reasons for granting external parties access to data. It offers valuable perspectives in understanding and discussing how data sharing is justified in business.

The data spectrum offers a common terminology when discussing the openness of data. Business model theory will contribute to the understanding of how a business works and the relation to data in the creation and capture of value. A platform perspective will expand the understanding of the business and its participation in a coordinated effort to deliver value within a business ecosystem.

Open data is a novel domain, and it is interesting to learn how these existing business theories apply in the context of data sharing.

Research Direction

Research Problem

Open data has been a growing topic in the public sector for several years (Huijboom & Van den Broek, 2011). Governments around the world are making various datasets from their governmental departments available for anyone to use. Today it would seem like a prestige project for governments to be on top of the Open Data Rankings (Open Knowledge Foundation, 2015; World Wide Web Foundation, 2015). While expectations are high, the true return on investment has still not been fully examined (Jetzek et al., 2013).

Open data is expected to have substantial value-generating potential. McKinsey estimates the potential of open data to an astonishing $3 trillion (Manyika et al., 2013). Innovation is highlighted as a key driver for turning open data into economic value. Free access to data is expected to foster innovative and data-driven products and services and enable overall performance enhancements in society at large. Data can be costly to collect, store and govern. Open data, therefore, enables anyone to access data without being hindered by the related costs. With closed data, only the entity who possesses the data or have the means to collect data will be able to explore the potential of its application. When data is opened, the crowd is empowered to contribute and pursue innovative ideas (Surowiecki, 2005).

The literature on open data has primarily focused on governments and non-profit grass-root organisations as data providers. Governments share their datasets to society. Meanwhile, grass-root organisations are working to make publicly available information on the World Wide Web accessible as linked open data (Auer et al., 2007).

If the proposition that the availability of data can enable enhancements to our society at large, is true, then how do we make sure to unleash this potential to the fullest?

Large datasets reside inside governments and across the World Wide Web's shared realms such as Wikipedia, yet there is also enormous data volumes inside companies. If these datasets were publicly available for everyone to use, it would seem that only the imagination would limit the innovative applications. But why should companies consider offering data for the outside

world? Doing so could potentially put their competitive position and existence at risk and create additional costs.

Despite the immediate concerns and contradictions from a business standpoint, tentative signs are seen from certain businesses who share some of their data with more than just close strategic partners.

Amongst companies that share their data with a broader non-specific audience, we have encountered Walmart, Facebook, Twitter, Google and Twitch (Facebook Inc., 2015; Google Inc., 2016; Twitch Interactive Inc., 2015; Twitter Inc., 2016; Walmart Stores Inc., 2016).

Walmart for instance shares data about their products, product reviews, product categories and stores (Walmart Stores Inc., 2016). While Facebook shares data about almost every kind of entity on their service through the "Graph API" (Facebook Inc., 2015).

Although most of these data sharing initiatives do not comply with the definition of open data, it is important to learn what motives causes such actions and which hindrances that limit further openness in a business setting.

Open data has several forms, amongst others: Open Government Data, Linked Open Data and Linked Open Government Data. Data resides not only in governments or not-for-profit organisations but inside profitable businesses. We refer to these datasets, when opened, as Open Company Data, which is the focus of this research. Figure 7 has been adapted from Attard et al. (2015), to illustrate our introduction of company data and our focus on the intersection between open data and company data.

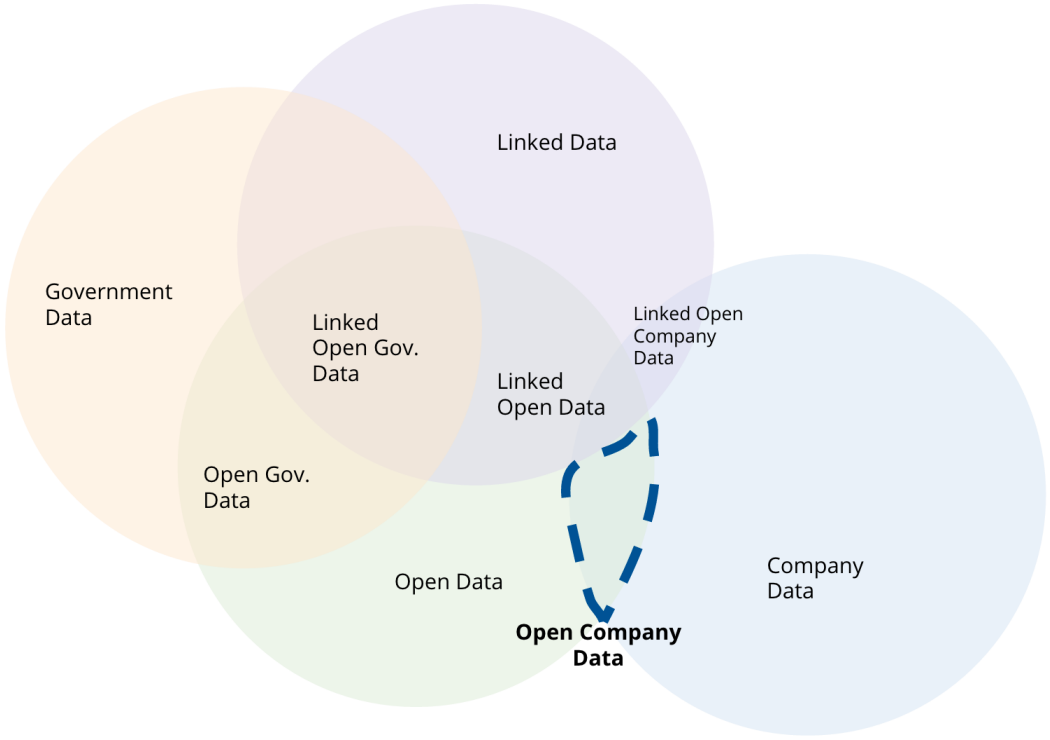


Figure 7: Positioning our research in the Open Data Domain (1)
Based on an adapted version of Attard et al.'s (2015) illustration of Relationships between Open, Government and Linked Data.

Further, we see a distinction in the research domain of analysing either the data provider or the data user (Janssen et al., 2012). In this research, the focus is on companies providing data for external use as highlighted in the illustration below.

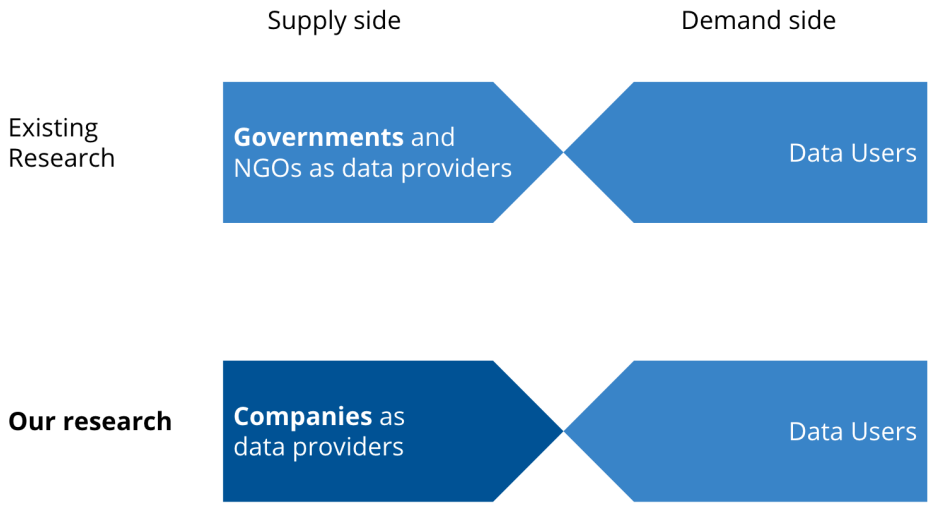


Figure 8: Positioning our research in the Open Data Domain (2)

Our position in research is further elaborated in the research question.

We are yet to observe significant research on the motivation for a private company to employ open approaches to data sharing. The problem we address is a lack of research on open data supplied by companies. We propose a break up with the trend of viewing open government data as synonymous with open data (Heimstädt et al., 2014). Data, if transformed, has the potential to generate substantial value for society, and thus, it is problematic, that a significant part of the world's data is isolated inside companies.

Research Question

The purpose of this thesis is to provide a deeper understanding of why some companies choose to share their data and discuss whether we will see companies offer open data.

The research question of this study is the following:

How is external data sharing justified in business?

This research question attempts to learn how companies view opening their data and what motivates them to take action in such a direction.

Since we realise that businesses might not be sharing their data as prescribed by the "open data" definition (OpenDefinition.org, 2015), we follow a more flexible phrasing: "external data sharing". Thus, a part of the answer is the determination of the current state of data openness in business.

We chose to look at the business justification for sharing data because it covers a broad range of possible outcomes. If we simply questioned: "How does external data sharing contribute to a profitable business?", we would limit our search and disregard potential intangible motivation. This question allows us to open the discussion of why open data is currently not justified in business.

To answer the research question and offer an in-depth understanding of the issue, we pose the following sub-questions:

- *How does business offer access to data and to whom?*
- *What is the role of external data sharing in the business model?*
- *How does external data sharing affect the business in a business ecosystem?*
- *What are the hindrances for business to open data?*
- *How can business move towards open data?*

RESEARCH DESIGN

In this section, we present the research design of this thesis. As proposed in Creswell's (2008) framework (see Figure 9), we see research design as consisting of three interconnected components. Those components are the philosophical worldview, strategy of inquiry and research method. In covering the methodology of this research, we will be referring to these three components.

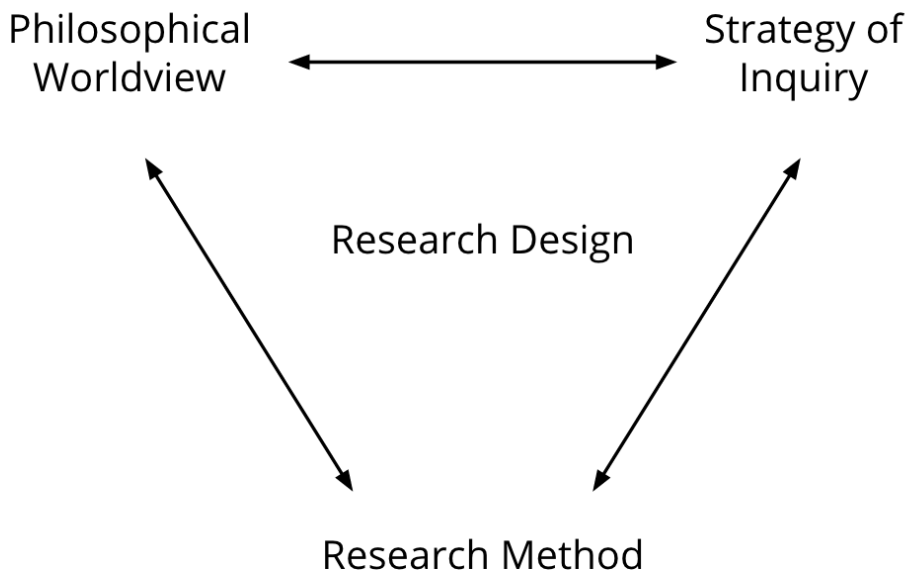


Figure 9: Simplified illustration of "A Framework for Design"
Adapted from Creswell (2008)

"Research Designs are plans and the procedures for research that span the decisions from broad assumptions to detailed methods of data collection and analysis." (Creswell, 2008, p. 3).

In short, this study is an exploratory case study on the company Trustpilot. A qualitative approach is applied with an underlying social constructivist worldview. Semi-structured

interviews are the primary source of data collection, while analysis and interpretation of data have been carried out through a mix of meaning condensation, meaning categorisation and meaning interpretation (Kvale, 1996). The quality of our research is assessed using the trustworthiness criterion suggested by Lincoln and Guba (1985).

Research Philosophy

In the following, we present our choice of paradigm: the underlying premises for conducting our study. When doing research it is of crucial importance to understand that the same study could have been viewed from multiple different worldviews, so our worldview acts as our lens and hence impacts the research by bringing both limitations and opportunities that other worldviews would not.

The selected worldview brings forth ontological, epistemological as well as methodological consequences, which will be presented in this section.

Guba and Lincoln (1994, p. 107) define paradigms as *"basic belief systems"* and states that, *"It represents a worldview that defines, for its holder, the nature of the "world," the individual's place in it, and the range of possible relationships to that world and its parts[...]"*.

The underlying worldview of this thesis is the constructivist worldview. Implying a reality that is constructed through social interactions; reality is created by people through interpretations. Since humans are different because of our values and assumptions, realities are also different. The social constructivist worldview implies that meanings are constantly constructed on the basis of human beings personal, cultural and historical experiences (Creswell, 2008).

We as researchers, therefore, interpret the meanings individuals in our study have. Our job is to make sense of and to interpret. We are not in the search for an 'objective' truth, as positivists believe there is. Instead, our role is to find meaning in a complex world of intersubjectivity. The decision of whether to open data or not is directed by the people in an organisation. Thus, it is their interpretation of it being beneficial to do or not, which leads to their decisions and the reality we experience. It is necessary to examine their understanding of the world, in order to understand their actions.

The social constructivist worldview implies the following ontological and epistemological consequences:

The ontology: the nature of reality is, that realities are multiple, intangible mental constructions, that are socially based and dependent on individual persons who hold the constructions. Thus, reality is subjective. (Guba & Lincoln, 1994).

The epistemology: the study on how knowledge is created, is transactional and subjective, which means that the object of study and the investigators are so interactively linked, that the *"'findings' are created as the investigation proceeds"* (Guba & Lincoln, 1994, p. 111). Knowledge is being constructed together by the researcher and participants.

Lastly, the methodology in the social constructivist worldview is hermeneutical and dialectical, with the final aim to reach a consensus construction, which is more informed and sophisticated than any predecessor (Guba & Lincoln, 1994). The effect of a hermeneutical and dialectical methodology, in this study, is that constructions are continuously reconstructed through our interactions with the case in study.

Logic of Inquiry: Abduction

Following our philosophical worldview, we employ an abductive logic of inquiry. The abductive approach entails a spiral process of conducting a scientific inquiry (Blaikie, 2007). Blaikie (2007) argues that there is no such thing as a clean cut "pure" induction or deduction. Proposing that inductive researchers cannot approach data generation and interpretation with a "blank mind", it will be influenced by assumptions deductively arrived at from previous work in their field. Correspondingly deductive researchers testing hypotheses will draw on theories that have been inductively arrived at.

Deductive inferences are *valid*, as long as the premises are *true*. In induction, inductive inferences are *probable,* when the premises are true. In an abductive approach, we infer best explanations (Douven, 2011). Fox (1998) describes the process of abduction as:

> *Abduction is inference to the best explanation. It is a form of problem solving used in a diverse number of problems, from diagnosis to story understanding, to theory formation and evaluation, to legal reasoning, to, possibly, perception.* (Fox, 1998, p. 1)

Our abductive research strategy entails that we use our participants discourse; their language and meanings (first order constructs) - to "abduct" to our own categories in our research (second order constructs) (Blaikie, 2007).

Consequently, the ontological assumptions of our abductive logic of inquiry mean that reality is constructed by social actors, with no single reality but multiple and changing realities. This proposition supports the study's social constructivist philosophy.

The method of abduction: *Inference to the best explanation*, is highly relevant for case study research. Dubois and Gadde (2002) suggest that *"Most textbooks on research methodology fail to take account of the opportunities offered by an intertwined research process enabled by case research. They tend to describe case studies as a linear process."* (p. 555). During this study we have gone back and forth between our research activities, theory and empirical observations to continuously expand our understanding of the phenomena, as suggested by Dubois and Gadde (2002).

Thus being abductive has implied that we do not view the case study as a linear process. Instead as a cyclical process, where empirical data and theory is continuously revisited as the investigation continued: a process referred to as systematic combining by Dubois and Gadde (2002).

Little prior research has been done on open data in the private context. An abductive logic of inquiry has helped us gain insight into the complex reality of business, ask *how* questions and understand the underlying human accounts and patterns in relation to opening or not opening data.

Deductive or inductive approaches would not allow us to approach the phenomena in such depth that we require.

We as researchers acknowledge that not one logic of inquiry cannot be followed rigidly. However, with the purpose of the study in mind, we find the abductive logic most appropriate in describing our way of arriving at conclusions. As previously stated our view is that reality is complex, which requires the application of an approach that can embrace such reality.

We are not trying to falsify a theory as a deductive approach would imply and neither are we collecting empirical evidence without any preconceptions about the phenomenon to generate new theory. Instead, we acknowledge that we have preconceptions and theories in mind when we conduct our study but are still open minded when collecting data since we might not be able to describe the phenomenon through existing theories and perspectives. In this sense, we attempt to arrive at a conclusion, which is synthesised from existing theoretical understandings and new empirical insights.

Type of Research Design

The research approach of this study is qualitative, which aligns with the social constructivist philosophical worldview.

Qualitative research stands in contrast to quantitative research. A qualitative research approach is generally focused on the analysis of words and builds on open-ended questions, whereas quantitative research is focused on numbers and closed-ended questions (Creswell, 2008). In between qualitative and quantitative research is the mix method approach, which combines qualitative and quantitative forms (Creswell, 2008).

Qualitative research is particularly useful when exploring and understanding social or human problems of great complexity (Creswell, 2008), which has been the purpose of this study. Finding an explanation for external data sharing in business requires an in-depth understanding of how employees of the organisation see their business. Thus, the reality they are constructing. Such understanding is best reached through qualitative research. The purpose of this study is not to quantify the impact of different factors on the decision of whether to open data or not, but to provide an in-depth understanding of how different parties in the business view the situation and which issues of opening data they focus on.

Adopting a qualitative research approach implies the interpretation of data and not the search for a statistically significant answer, which can be turned into a universal law or theory. We further elaborate on generalizability of our study in the following section.

Research Strategy

In the search for an answer to our research question, we did an exploratory case study on Trustpilot.

Case study

Case studies are a popular strategy of inquiry in social science research (Yin, 2009) and can be defined as *"a research strategy which focuses on understanding the dynamics present within single settings"* (Eisenhardt, 1989, p. 534). Case studies are not exclusively used in qualitative research. However, that is often the case (Creswell, 2008; Eisenhardt, 1989).

A case study is particularly useful in finding answers to questions of "why" and "how" (Oates, 2006; Yin, 2009). Which corresponds with the posed "how" question: "How is external data

sharing justified in business?". Other strategies of inquiry which finds answers to "why" and "how" are experiments and history. However, history does not focus on contemporary events, and experiments require control of behavioral events (Yin, 2009). Since we want to understand the situation in Trustpilot from a natural setting, an experiment is not a suitable strategy of inquiry. Thus, we are not attempting to isolate certain variables in a laboratory and manipulate the research in a systematic way as experiments by positivistic research could typically imply.

Besides the considerations mentioned above, the choice of a case study is motivated by the search for an in-depth understanding of a business that shares data. Understanding which can only be gained through the acknowledgement of the business context. According to Yin (2009, p. 18), that is the kind of understanding a case study provides: *"you would use the case study method because you wanted to understand a real-life phenomenon in-depth, but such understanding encompassed important contextual conditions [...]".*

In providing answers for this study, a case study strategy is particularly suitable since its nature is to focus on depth rather than breadth (Oates, 2006). Instead of surveying a broad range of companies about their stance on opening data, we do an in-depth analysis of a single instance, to gain a deeper understanding of the complex mechanisms which has lead to the current situation. Since empirical research on businesses opening their data to its surroundings is limited, we can only guess which questions are relevant to ask. Performing a survey thus would limit the answers through our prescribed understanding of the issue, that would form the questions posed. Through a case study a deeper understanding can be reached through continuous questioning of "how" and "why". This approach can lead to insights beyond the obvious. Additionally, the case study is characterised by its holistic view on real-life events (Oates, 2006). Thus, such an approach has allowed us to understand the complexity of the phenomenon.

Even though a case study can provide deep insights into a specific instance, it is questioned by critics whether those insights can be generalised or even contribute to scientific development (Flyvbjerg, 2006; Yin, 2009). Flyvbjerg (2006) attempts to answer this critique and suggests that scientific development is not only reached through generalisations, but a case can play an important role in driving science forward. *"[...] formal generalizations is only one of many ways by which people gain and accumulate knowledge"* (Flyvberg, 2006, p. 10). Adding that "*The force of example"* is often underestimated (Flyvbjerg, 2006). He further suggests that case studies can often be generalised, depending on the strategic choice of case. Case studies that do not attempt to generalise can still contribute to collective knowledge contribution: *"A purely*

descriptive, phenomenological case study, without any attempts to generalize can certainly be of value in this process and has often helped cut a path toward scientific innovation" (Flyvbjerg, 2006, p. 10).

Yin (2009) suggests that generalisations can be categorised into analytic generalisation and statistical generalisation, where qualitative case studies focus on *analytic generalisation*; to expand and generalise theory. Contrarily, statistical generalisation implies enumeration of frequencies to make inference about a population or universe (Yin, 2009). In analytic generalisation, theory developed previously is used as a template with which to compare the empirical results of the case study. This study aimed for analytical generalisation, by applying both business model theory and platform theory as complementing theories in explaining reasonings behind a data openness stance.

In this particular case study, we are attempting to take early steps in the scientific understanding of open data in business. Our findings from the case of Trustpilot is not necessarily applicable to other businesses but provides a deep understanding of a specific case. Whether the complex situation in other businesses is alike is to be explored in further research within the field. Thus, this study focuses on understanding unexplored territory holistically in a single setting, rather than concluding on the broader spectrum of businesses.

Yin (2009) proposes three types of case studies. Those are exploratory, descriptive and explanatory. The purpose of our research informs the type of case study (Zucker, 2009). This study is exploratory in nature. The state of open data research, particularly the direction this study has taken is yet an unknown area and calls for exploratory research. Saunders et al. (2007) describe exploratory research *"a valuable means of finding out what is happening; to seek new insights; to ask questions and to assess phenomena in a new light"* (p. 139). While addressing existing speculations, we generated new questions and propositions through the use of a pilot case.

Case selection

In selecting a case for this study, we searched for a company who already offered their data to parties external to the organisation and with a physical location in Denmark. Trustpilot fulfilled these criteria.

Unit of analysis

An important element in a case study is the *unit of analysis*, which helps to define the "case". The unit of analysis can, for example, be an individual, an organisation, a community, a system or a project (Yin, 2009). The unit of analysis of this case study is a business. In specifying the unit of analysis, it becomes clear what we are trying to learn, and from where we will be looking for the answers to the research question. Since we looked for a for justification for external data sharing in business we studied the business itself. Alternatively, if we had tried to learn how others find use of their data, the unit of analysis could have been the ecosystem in which Trustpilot is involved.

Settling on a case

In selecting a business as a case for this study, we searched for private Danish companies that enable others to use their data. Most businesses are exchanging data with their closest business partners, like suppliers, distributors and financial institutions with specific purposes. We were looking for businesses not only sharing their data with existing stakeholders but who offer it publicly in a loosely coupled manner. This entails that other businesses, individuals or organisations can make use of their data. The choice to look for a business in the Danish business landscape, was a convenience criterion to allow for rich face to face interviews in their natural setting.

Through a list constructed by a group of Danish programmers (Mauran et al., 2016) we located a number of businesses in Denmark, who actively offer some of their data to others. Those were:

- Dansk Supermarked A/S[9]
- Saxo.com A/S[10]
- Rejseplanen A/S[11]
- Trustpilot A/S[12]

Since Rejseplanen is closely related to the government, we concluded it was not the best fit for this research: addressing private organisations in contrast to governmental organisations.

[9] https://developer.dansksupermarked.dk
[10] http://api.saxo.com/
[11] http://labs.rejseplanen.dk/labs/data__brug/
[12] https://developers.trustpilot.com/

After having been in touch with the different case companies, we managed to settle on a case study on Trustpilot.

	Private Company	Location in Denmark	Public Data Interface
Dansk Supermarked A/S	Yes	Yes	Yes
Saxo.com A/S	Yes	Yes	Yes
Rejseplanen A/S	(Yes)	Yes	Yes
Trustpilot A/S	**Yes**	**Yes**	**Yes**

Table 3: Identified companies that fulfilled case selection criteria

Trustpilot is an online business, which offers webshop reviews through their website. Here consumers can share their online purchasing experience with others and look up businesses to learn how they have been rated, before placing an order with them.

Trustpilot is particularly interesting since they share data from their core product, which is reviews. The digital nature of their business makes them a great case since they share the digital characteristic with businesses mentioned earlier, that also share data: Facebook, Twitter, Google and Twitch.

Research Method

This section presents the chosen empirical methods of this study. Our methods of collecting data and creating new knowledge are presented. The paper is mainly based on primary data, which refers to data collected by us as researchers for the purpose of this paper. Primary data is collected through interviews and Trustpilot's website. Supporting secondary data has been collected through external media coverage and reports on Trustpilot.

Empirical method

Primary data is collected through semi-structured interviews with Trustpilot employees, which are related to their external data sharing activities and strategy. We conducted six

interviews with key employees of Trustpilot in the period from 24th February to 20th April 2016. Below is a list of interviews conducted:

#	Referred to as	Role	Rank	Date	Duration (h:mm)	Transcript
1	Respondent 1	Engineering	Very high	24th Feb. 2016	1:44	Appendix 7
2	Respondent 2	Marketing	High	11th Mar 2016	1:10	Appendix 7
3	Respondent 3	Strategy	Very high	29th March, 2016	1:18	Appendix 7
4	Respondent 4	Market and Sales	Medium	18th April, 2016	0:43	Appendix 7
5	Respondent 5	Technical Specialist	Medium	18th April, 2016	0:31	Appendix 7
6	Respondent 6	Integrations	Medium	20th April, 2016	0:39	Appendix 7

Table 4: Respondent Interviews Overview

The utilisation of semi-structured interviews has allowed us to gain an in-depth understanding of the reality within the case organisation. In contrast to structured interviews, semi-structured interviews enabled us to control the direction of questions, to encourage an open exchange and exploration into unexpected directions (Creswell, 2008).

Our interviews were conducted through the use of interview guides, which were customised towards the respondent in question. We learned about the informants from what their colleague could tell, the information they provided themselves in initial communication, and what we could learn from their Linkedin profile. Interview guides consisted of themes derived from our analytical framework, as well as supporting questions. The questions and themes of our interview guide originated from our research question coupled with themes from the theory that could help in answering the research question. A process that is inspired by Kvale's (2007) 7 steps to interviews suggesting to thematise and design guides before interviewing. Interview guides can be found in Appendix 7.

The high-level structure of our interview guides were as follows:

- Build rapport
- Introductory questions
- Questions divided into themes

After each interview, we did a post interview debrief to gather our thoughts and new ideas. These ideas were then used to evolve our next interview guide and drive our research.

Our abductive logic of inquiry has meant that our frame of understanding evolved through the process of the study. Periods between interviews allowed us to synthesise our findings and continuously revisit our research questions and interview guides throughout the study.

To ensure richness, all interviews were conducted face-to-face and are documented through transcriptions. We are aware that our presence as researchers can affect the respondents responses, which we have attempted to mitigate through building rapport with the respondents before the actual interviews. Since the case company is global and the office has a multitude of nationalities present, we conducted most interviews in English. However, if interviewees felt more comfortable speaking Danish, we would agree to do so. Both researchers were present at all interviews. Reference to interview recordings and transcriptions is provided in Appendix 7.

In addition to interviews, we also had informal encounters with people from the organisation, the informal nature was beneficial as under these conversations, the respondents did not feel pressured to be "politically correct" as all formal interviews were recorded. These conversations were summarised in notes and processed in our post data collection debriefs.

Documents in the form of web pages have also been used as a means of collecting data; this includes screenshots of Trustpilot's consumer website, business app, support articles and API documentation. The use of documents allowed us as researchers to further understand the case company in their own language and vocabulary (Creswell, 2008). The use of multiple data sources has served as a means of triangulation, which is important for the acquisition of knowledge about the multiple and diverse set of realities involved (Johnson, 1997).

Selection of respondents

The selection of respondents was based on a non-probabilistic basis and employed snowball sampling: *"a technique for finding research subjects. One subject gives the researcher the name of another subject, who in turn provides the name of a third, and so on."* (Vogt, 1999, p. 368).

Respondents were recruited as informal research assistants to help gain access to people amongst their acquaintances, whom they believed could contribute on the topics covered in interviews (Atkinson & Flint, 2001). In guiding our respondent's referral to the next respondent, we presented the themes and issues we wished to cover and requested alternative points of view, e.g. people from other departments of the organisation. The process of recruiting informants has been visualised in the figure below:

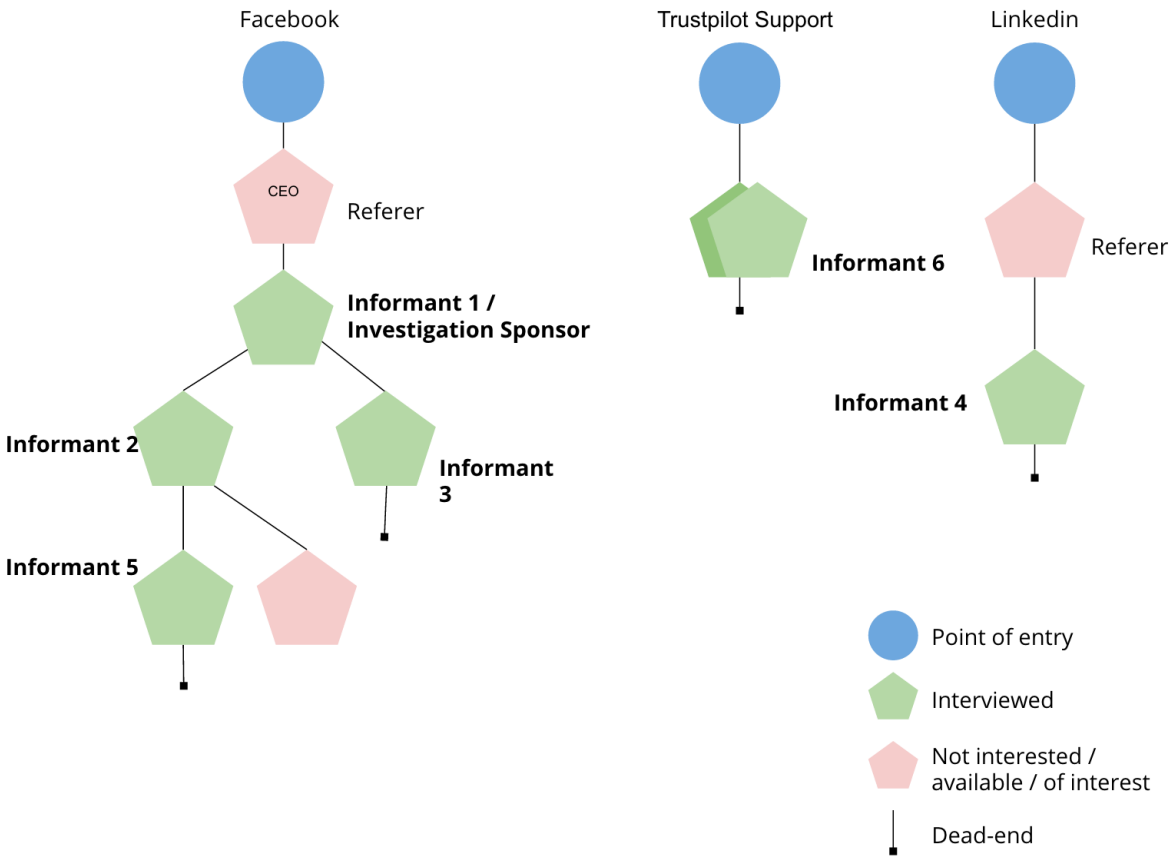


Figure 10: Illustration of respondent sampling process

This method of locating subjects was useful since we did not know whom in Trustpilot that possessed the knowledge we were looking for. We leveraged the employees insights about each other's responsibilities, activities and knowledge, to find the most interesting respondents.

A downside of this sampling technique is that the respondents can refer us colleagues whom they agree with, while excluding those they disagree with. Another issue to snowball sampling is premature dead ends, as we experienced. Along the way, the snowball sampling technique

lead to dead ends as seen in Figure 10. In this scenario, we searched for additional respondents, which we located from Trustpilot's Linkedin company profile and Trustpilot's support inbox (see Figure 10). Getting in touch with new employees in this manner lead to two additional informants recruited purposely with the focus on broadening the variety of informants based on their domain.

Analytical Method

Learning from collected data requires analysis and interpretation. During this study, we analysed and interpreted interviews and documentation collected from Trustpilot. We looked for patterns which could help to answer the research question at hand.

"The process of data analysis involves making sense out of text and image data" (Creswell, 2008, p. 183).

In preparing the data for analysis, we fully transcribed each interview and organised interviews and documentation in folders by source. Starting the analysis, we read through the entire data material to write short summaries, also known as meaning condensation (Kvale, 2007). This helped gain an overview of the data and the responses, before deeper interpretation.

From here we thoroughly reviewed the data for themes to group text segments by predetermined and emerging codes (see Appendix 8). *"Coding is the process of organizing the material in chunks or segments of text before bringing meaning to information"* (Creswell, 2008, p. 186). The process of coding interviews is also referred to as meaning categorisation (Kvale, 2007). In this regard, we looked to categorise the transcripts by codes from the analytical framework, codes from the open data domain and unexpected codes from the setting and context which reappeared throughout the data. Thus, we embrace an approach which both focuses on analysis through the perspectives of the applied theories as well as looking for new themes in the data. This process aligns with our abductive approach; moving back and forth from theory and empirical data. The entire process of analysing data has been a circular and iterative process, where meaning condensation, meaning categorisation, interpretation, theoretical reflection and further data collection has been interwoven activities.

In practice, a software tool called Saturate App[13] was used to add codes to our transcriptions and automatically group text segments. Using such a tool proved valuable in the analysis process in contrast to manual pen and paper.

To understand the case of Trustpilot and to familiarise readers with the case, we constructed a description of Trustpilot as a company. Such a description is advised by Creswell (2008, p. 184): *"Case study and ethnographic research involves a detailed description of the setting and individuals, followed by analysis of the data for themes or issues"*.

The description of Trustpilot leads to an analysis of the business model of Trustpilot, adhering to the business model framework presented by Osterwalder and Pigneur (2010). The purpose of applying the business model canvas is to understand the mechanisms at play in Trustpilot, which can help to understand their perspective on data. The role of data in the business model is analysed to determine why an external data sharing initiative is in place and how that helps Trustpilot create, deliver and capture value.

Furthermore, Trustpilot's data sharing initiative is analysed through the perspective of platforms and ecosystems to provide an alternative view of how to understand the motivation for sharing data in a private setting.

Research Quality

In evaluating the research quality of this paper, we discuss the trustworthiness, which is suggested as the primary quality criterion in constructivism (Guba & Lincoln, 1994). Four criteria indicate, the trustworthiness of a study: *credibility*, *transferability*, *dependability* and *confirmability* (Lincoln & Guba, 1985). These four criteria can be seen as the qualitative researchers' equivalent to the popular criteria internal validity, external validity, reliability and objectivity (Marton, 2013; Shenton, 2004).

Credibility

A central criterion in conducting research of trustworthiness is that of credibility (Lincoln & Guba, 1985). It can be boiled down to a simple question, is this research credible? Lincoln and Guba (1985) propose multiple conventions of ensuring high credibility. Amongst others, those include the importance of getting early familiarity with the culture of the participating organisation, the use of triangulation through the usage of different sources, and employment

[13] http://www.saturateapp.com

of strategies to ensure honesty from respondents. Furthermore, thick descriptions of the context and contact with a range of respondents are all factors that contribute to credibility in research (Shenton, 2004).

Throughout this study, we have employed several strategies to ensure credibility. Examples thereof include our choices of well-established research methods, such as semi-structured interviews, theming and systematic coding of interviews for analysis. We placed a high emphasis on eliciting honest responses from our respondents through rapport building. Furthermore, we attempted to build rapport early on in each interview and took further precautions to ensure our respondents were honest by offering anonymity. Variety in respondents was assured through the inclusion of employees of different seniority and domain of responsibility. This also helped triangulate responses by checking statements against one another, as well for the purpose of ensuring variety in our data sources. Additionally, we attended a keynote by Peter Mühlmann, the CEO and founder of Trustpilot. He talked about the culture and history of Trustpilot all of which contributed to an early familiarity with the organisation of Trustpilot. To improve our familiarity with Trustpilot, we requested permission to observe the case company. However, it was not an option.

Transferability

Transferability refers to the extent in which findings of the study can be applied to other situations (Shenton, 2004). Had this been quantitative research, we would have emphasised demonstrating large enough sample sizes to generalise to wider populations. *"[...] the naturalist cannot specify the external validity of an inquiry; he or she can provide only the thick description necessary to enable someone interested in making a transfer to reach a conclusion about whether transfer can be contemplated as a possibility"* (Lincoln & Guba, 1985, p. 316).

In extension of the credibility criterion of providing thick descriptions, we again emphasise our focus on providing as much contextual information as possible. This is to enable other researchers to distinguish and evaluate whether our results can be transferred to their study and thus giving the thick description necessary to allow for transferability to other contexts. Thick descriptions are provided through thorough elaboration of our methodological underpinnings and providing contextual information throughout the analysis of Trustpilot (Shenton, 2004).

Dependability

The dependability criterion is comparable to the conventional reliability criterion (Marton, 2013; Seale, 1999; Shenton, 2004), that suggests other researchers should be able to carry out the same study and end up with the same result. However, we acknowledge that findings are a result of social constructions and that social phenomena are dynamic. Thus, the exact same result can not be expected.

Instead of focusing on repeatability, the dependability criterion concerns to which degree the researchers has been accounting for the likelihood of stability of the findings over time (Anney, 2014). In practice, this involves accounting for the ever-changing context of the study (Anney, 2014).

Ways to provide such an account include an audit trail. An audit trail accounts for all research decisions and activities to describe how data was collected and analysed. Such an account has been given in the previous sub-section, *Research Method*. Additionally, documentation from the inquiry process is included in appendixes, such as references to interview recordings, interview guides, post-interview debriefs, transcriptions, interview summaries and categorisation procedures in the form of codes.

Confirmability

Confirmability concerns the degree to which results or findings can be confirmed or corroborated by others than the researchers themselves (Anney, 2014). Thus, the focus is on establishing trust in that the results are derived from the data and not from the researcher's imagination.

As social constructivists, our findings are second-level constructs as described earlier. That means we do not attempt to claim our findings as objective truths, as would be the objective of a positivistic inquiry. Instead, we provide readers with explicit descriptions of the researcher's backgrounds so judgements of bias, interests and motivation can be made (see About the authors, p. 125).

In practice, we address this criterion by supporting our claims in the analysis with reference to data sources. This way readers can confirm findings, by investigating referenced data themselves. The full data material is found in appendixes (Appendix 7).

The clarification of research quality criteria offers a way for the reader to assess the quality of our research.

Analysis

In this section, we analyse Trustpilot to explore their justification for granting external parties access to their data.

We start the analysis with a case description of Trustpilot. From here, the analysis consists of three parts. In the first part, we classify Trustpilot's data sharing initiative in the data spectrum. In the second part, the dynamics of Trustpilot's business model are examined to understand the role of data sharing. Lastly, we explore Trustpilot as a platform, rather than a product to further extend the investigation of their rationale for sharing data in a business ecosystem.

Case Description: Trustpilot

Trustpilot was founded in 2007 by Peter Mühlmann, who set out to help his parents make better online purchasing decisions. Trustpilot is a review community, where consumers can rate and review their experiences with online businesses. Other consumers can learn from their experiences, before an eventual purchase. Since Trustpilot's entry in 2007, it has grown significantly and now employs close to 500 employees (Respondent 1, personal interview, Feb. 2016). Trustpilot has raised $117 million in funding and has offices in Copenhagen, London, New York and Melbourne (Crunchbase, 2016; Trustpilot.com, 2015). According to Crunchbase, a new review is posted every 5 seconds, which gives an indication of a rapidly growing review community.

Trustpilot has two key offerings: One is the review community, where anyone can write reviews and score their experience with the respective business they dealt with. The result of a review effects the *TrustScore* of the respective business. Second is their business offering, which amongst others allows companies to embed their TrustScore on their own website. Furthermore, the business product allows businesses to communicate with the reviewers through the platform and gain deeper analytical insights into customer review data (Trustpilot.com, 2015).

Trustpilot's competitors in the review space are services such as Yotpo.com, BazaarVoice.com, Reviews.co.uk and Google Trusted Stores. Although their different products vary in scope, they exist in the same space of customer reviews and product reviews (Respondent 3, personal interview, Mar. 2016).

Position in the Data Spectrum

From the short description of Trustpilot as a case and its context, we now move into the analysis of the specific phenomenon of interest: external access to Trustpilot's data. As explained earlier, we see open data as an extreme state along the spectrum from closed to open; a perspective supported by the Open Data Institute (2016a). In this section, we present an account of the external access to Trustpilot's data, compare it to the open data definition and analyse the degree of openness in a spectrum from open to closed data. We find that Trustpilot does not offer open data, but shared data.

External data access interface

Trustpilot grants data access to external parties, such as customers (Respondent 1, personal interview, Feb. 2016, Respondent 5, personal interview, Apr. 2016, Respondent 6, personal interview, Apr. 2016) and various types of partners (Respondent 1, personal interview, Feb. 2016, Respondent 2, personal interview, Mar. 2016). The primary outlet of Trustpilot's data is their APIs.

Trustpilot offers two levels of data access through what they call "Public APIs" and "Customer APIs" (Trustpilot.com, 2016d; Respondent 1, personal interview, Feb. 2016). Public APIs provides access to data otherwise available through the user interface on Trustpilot.com. None of the data in Public APIs is thus secret and could equally be obtained by crawling[14] the website (Respondent 1, personal interview, Feb. 2016).

Data exposed in the Customer APIs is private and belongs to the customer (Trustpilot.com, 2016d). By protecting data owned by customers in separate APIs, Trustpilot is dealing with data privacy (Respondent 1, personal interview, Feb. 2016).

In order to retrieve data from any of the APIs, an API key is required as illustrated in Figure 11. The API key is retrieved from Trustpilot (Respondent 5, personal interview, Apr. 2016;

[14] Crawling: A technique for extracting data from a website's underlying markup.

Trustpilot.com, 2016d). The API key authentication mechanism enables Trustpilot to control who can access the APIs.

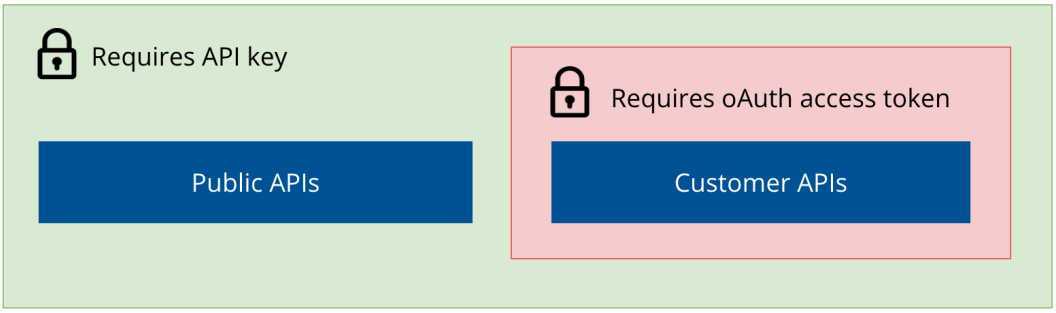


Figure 11: Illustration of authentication mechanisms for Trustpilot's APIs
Constructed by the authors, based on information from Respondent 1, Respondent 5 and Trustpilot.com (2016d)

For the Customer APIs, a second level of authentication is added (see Figure 11). If a customer wants to access their own data through the APIs, they will be obtaining an access token prior to making requests (Respondent 5, personal interview, Apr. 2016; Trustpilot.com, 2016d). This is achieved through an authentication protocol known as oAuth2. Customers can also allow third party services or apps to access their data from within the Customer APIs. This mechanism is similar to those of Facebook and other social media, that we experience when we give apps permission to use our information (Respondent 1, personal interview, Feb. 2016). Particularly, customers must authenticate third party apps through the oAuth2 protocol, which provides an access token. This access token allows third party developers to retrieve data on behalf of the customer.

Obtaining access

API keys are not distributed to anyone and everyone. Trustpilot offers API keys for its top-tier customers. More specifically those with a "Pro" or "Enterprise" subscription (Respondent 1, personal interview, Feb. 2016; Trustpilot.com, 2016f), whereas customers with a "Free" or "Lite" account cannot access the APIs. Furthermore, only Enterprise customers are allowed access the Customer APIs (Trustpilot.com, 2016f).

	Public APIs	Customer APIs
Free Customers	Disallowed	Disallowed
Lite Customers	Disallowed	Disallowed
Pro Customers	Allowed	Disallowed
Enterprise Customers	Allowed	Allowed
Partners	Allowed	Allowed (with permission from customer)
Everyone else	Disallowed	Disallowed

Table 5: Trustpilot API Access Overview
Constructed by the authors based on information from Trustpilot.com (2016f), Respondent 1 (personal interview, Feb. 2016), Respondent 2 (personal interview, Mar. 2016) and Respondent 5 (personal interview, Apr. 2016)

Besides top-tier customers, partners can be granted access to Trustpilot APIs through the Partner Program (Respondent 1, personal interview, Feb. 2016, Respondent 2, personal interview, Mar. 2016; Trustpilot.com, 2016e). Third parties looking for partnership access are required to send an application through an online form (Respondent 1, personal interview, Feb. 2016). Applicants must fill the form with their basic information such as name, email, website, country etc. (Respondent 2, personal interview, Mar. 2016). If the applicant plans to build something with the Trustpilot APIs, a description of how APIs and data will be used is required, as well as information on where the data is planned to be displayed. The form prompts: *"How do you plan to use Trustpilot's API and data?"* and *"Which domain(s) do you plan to display Trustpilot data on?"* (Trustpilot.com, 2016e).

Partnership applications

Trustpilot assess partnership applications to determine whether the third party is intending to use the data in an acceptable manner (Respondent 2, personal interview, Mar. 2016). The focus is here on whether the applicant is capable of building a functional and value creating product of high quality, or whether the applicant is a customer attempting to gain free access or if the applicant is likely to abuse the data and mislead consumers (Respondent 2, personal interview, Mar. 2016). Additionally, this process enables Trustpilot to filter out and evaluate if applicants are serious or just 'playing around' (Respondent 1, personal interview, Feb. 2016).

If an applicant is accepted for the partner program, Respondent 2 says that a contract is signed and highlights the terms of use:

> *So we have written guidelines that we put on our support site. But essentially you need to credit the data source. There needs to be an actual link to the profiles. Just as proof, that anybody can go in and check, so they say: "This company is rated 8,9 out of 10 on Trustpilot.com with 300 reviews" [...] You are not allowed to mess with the data and alter it in any way. You are not allowed to mix Trustpilot scores with some other scores and create a 'Frankenstein score' because that is just too confusing.* (Respondent 2, personal interview, Mar. 2016, 16:28)

Respondent 2 refers to the guidelines outlined on the Trustpilot Support Center (2016a). These requirements include content requirements, SEO[15] requirements, and branding requirement.

Content requirements dictate that: Ratings must be displayed as star rating, star label or TrustScore. Display of rating or review must include a link back to Trustpilot. If a subset of reviews is displayed, it must clearly be stated.

SEO requirements dictate that reviews cannot be included directly in a website's source code. Instead, it must be inserted dynamically so that Google's search engine cannot index it.

Branding requirements force data users to display the Trustpilot brand every time they display reviews or ratings.

Accessible Data

Trustpilot offers access to a lot of data through their API. API users can access data about business units, categories, consumers, reviews, invitations and resources (Trustpilot.com, 2016d). A complete overview of data accessible through the Public and Customer APIs is available in Appendix 4. Accessible data has primarily been examined through the Trustpilot API documentation (Trustpilot.com, 2016d). The following gives an outline of data accessible through Trustpilot APIs.

Business Units: Business Units are the businesses, which can be found and reviewed on Trustpilot.com. Business Units are what Trustpilot refers to as "*[...] one main domain name (primary name), and it can be associated with other domain names or sub-domains (referral names).*" (Trustpilot.com, 2016d). An example of a business unit could be Amazon.com.

[15] Search Engine Optimization

Among data accessible about business units are the website URLs, TrustScore, stars, the number of reviews and associated reviews and categories.

Categories: A category is another entity in Trustpilot's data. A business unit has one or several business categories. Categories could be *Travel* and *Vacation, Transport* or *Electronics* (Trustpilot.com, 2016c). Information accessible about categories includes category name, parent category, the number of businesses in the category and associated businesses.

Consumers: API users can also access information about the consumers who are writing reviews on Trustpilot. Information accessible on consumers includes link to Trustpilot profile, display name, city, country, gender, birth year and about text. Through the Customer API, the email of the consumer is also accessible.

Reviews: Trustpilot is primarily about reviews of business units, but has also started to support product reviews by request from the customer (Respondent 3, personal interview, Mar. 2016).

Information about reviews of businesses include who the reviewing consumer is, review title, review text, stars given, language, date, the reply from the business unit and verification status. With access to the Customer APIs, additional information is accessible, like referral e-mail, tags, source of the review, a status and some report data.

Roughly the same information is given about product reviews. However, further information about product reviews is also accessible through the Customer API, such as product brand, product name, and universal product identifiers.

Invitations and Resources: Trustpilot enables their business customers to send automated invitations to consumers, inviting them to write a review about their experience with their service. In this context, certain emails based on email templates are sent to the consumer. Information about available templates can be retrieved through the Customer API. Trustpilot also offers a number of resources, which include links to Trustpilot image assets, data about languages supported by Trustpilot, data about countries and translations of star ratings into verbal expressions such as "Good" in different languages.

The external facing APIs are the exact same versions Trustpilot use to develop internal products (Respondent 1, personal interview, Feb. 2016; Respondent 3; personal interview, Mar. 2016; Respondent 6; personal interview; Apr. 2016). However, they also have additional APIs, they only use internally (Respondent 3; personal interview; Mar. 2016).

Openness of Trustpilot data

Trustpilot's external data access has now been outlined. In the following, we assess the openness of Trustpilot's data by applying and positioning it on the Data Spectrum.

Open data is, as previously stated, defined as: *"[...] data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike."* (Open Knowledge Foundation, 2013). This definition makes it clear that the data Trustpilot offers through their APIs is not open. Trustpilot does not allow anyone to use, share or build on their data. The allowance to use and build on their data is restricted to paying customers and approved partners. Additionally, these parties are not allowed to use the data for any purpose they desire. Trustpilot set a number of guidelines for what is acceptable use of their data and what is not acceptable use.

If Trustpilot's data is not open - is it then closed? As defined by Broad (2015) closed data is: *"Data that can only be accessed by its subject, owner or holder."*, which is also not the case for the data found in Trustpilot's APIs. The data is accessible not only to Trustpilot. Customers and approved partners can even access data about other businesses and consumers on Trustpilot.

It's Shared Data

Trustpilot's data is neither in line with the definition of open nor closed data. Instead, Trustpilot's data can be categorised somewhere in between the two states, termed as "Shared Data" in the Data Spectrum as highlighted in Figure 12.

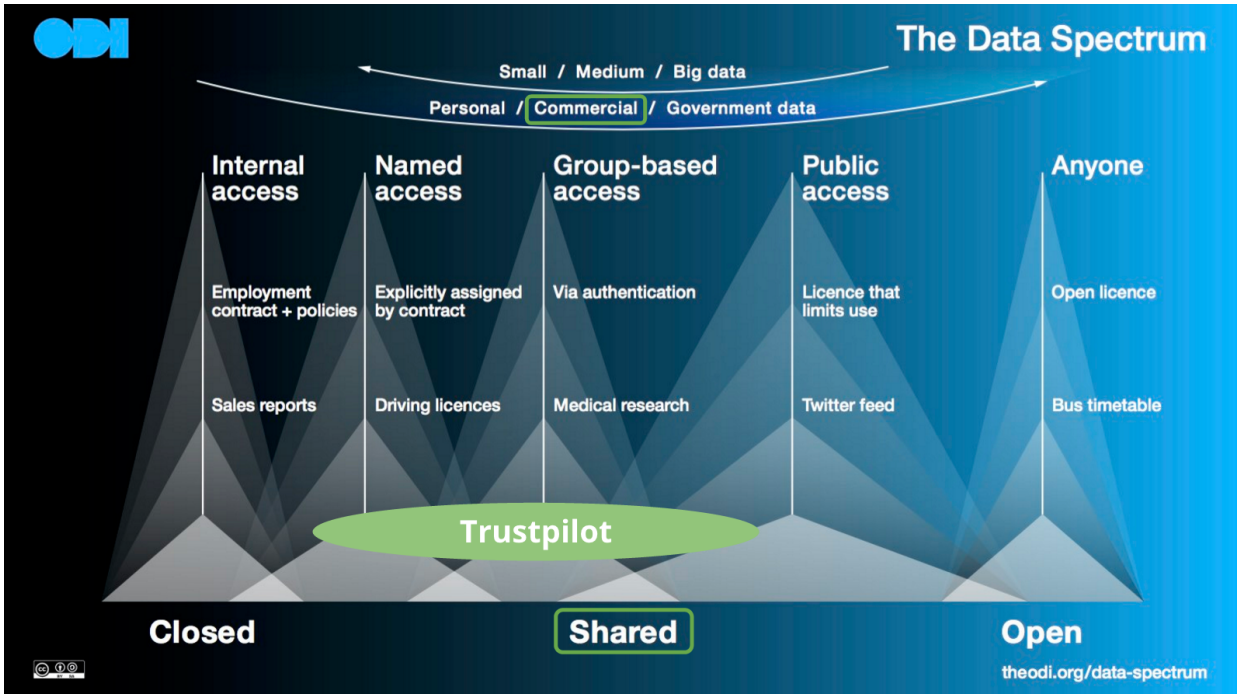


Figure 12: Positioning Trustpilot's Data in the Data Spectrum
Data Spectrum by Open Data Institute (2016a)

A more precise position in the Data Spectrum is difficult to determine. Trustpilot does sign contracts with partners looking for access to APIs (Respondent 2, personal interview, Mar. 2016). However, access is obtained via authentication. Partners and paying customers of Trustpilot are two different groups, that can access the same data through the Public APIs. Customers gain access to more data which is specifically provided for their business. Thus, the access to this data is tied to their identity.

Even though Trustpilot term some of their APIs as Public APIs, it is not publicly accessible and thus cannot be categorised as *Public access*. However, the information exposed in the Public APIs is not secret, but publicly available through the consumer-facing part of their website, Trustpilot.com.

Trustpilot's external data initiative can be categorised as shared data. More precisely it is positioned in between *Named access* and *Group-based access.*

Summarisation of Data Sharing in Trustpilot

In summary, Trustpilot provides two levels of API access. The first is the Public APIs, which allows the consumer of the API to programmatically access all information available on Trustpilot.com. The second is the Customer APIs, which allows the API user to access more personal information. This API is intended for business users to gain deeper analytical insight into the profiles of their customers.

Lastly, we have established that Trustpilot does not conform to the definition of open data and neither satisfies the definition of closed data. Therefore, we classify data in Trustpilot's APIs as shared data. Placing it in the center of the data spectrum. Shared data can broadly be categorised as data accessible by certain audiences and with limited use authority.

Business model

To understand how Trustpilot sees data sharing and why they have settled on their current data sharing policy, it is important to first understand the core of their business - the inner workings of Trustpilot and what makes it successful.

The analysis of the business model provides insights, which helps to understand the role of data and data sharing in the business. Further allowing for an understanding of the underlying background for the actions and the current position in regards to external data sharing.

To analyse the business model of Trustpilot we address the nine interrelated building blocks of the business model canvas as proposed by Osterwalder and Pigneur (2010).

Customer Segments

Trustpilot is serving two groups of customers. Those are businesses and consumers (Respondent 2, personal interview, Mar. 2016). The two groups engage with each other on Trustpilot's website in what can be categorised as a multi-sided platform (Hagiu & Wright, 2015). The duality is illustrated in Figure 13. Consumers are writing and reading reviews, while businesses receive feedback, build credibility and respond to reviews (Respondent 4, personal interview, Apr. 2016).

Figure 13: Illustration of Trustpilot as a multi-sided platform.

Consumers are not seen as customers by all respondents but as an important target audience. Respondent 4 describes consumers as: *"People that are using reviews to make buying decisions"* (personal interview, Apr. 2016; 3:23).

Businesses are seen as the actual customer since they are the ones paying for Trustpilot's offerings. Respondent 4 describes customers: *"I would say the customers on Trustpilot are businesses... Large and small. Any business where reviews, would be important."* (personal interview, Apr. 2016, 3:23). Even though Trustpilot serves largely all kind of businesses, businesses are not simply businesses in the perspective of Trustpilot. Respondents refer to different attributes of businesses which help them distinguish business customer groups, those are: online vs. offline, products vs. services, size and industry.

The initial focus of Trustpilot was on e-commerce stores, but now the focus has expanded to additional segments such as more traditional and offline businesses. Examples of such are found within the financial sector, banks, insurance companies (Respondent 1, personal interview, Feb. 2016) and law firms (Respondent 4, personal interview, Apr. 2016). However, a large portion of customers is still e-commerce businesses (Respondent 1, personal interview, Feb. 2016). Trustpilot offers their services to all company sizes: small, mid-sized and larger enterprises (Respondent 2, personal interview, Feb. 2016). A large share of customers are small businesses, but Trustpilot is more focused on serving the larger companies. We return to this issue in the building block, *customer relationships*.

Value Proposition

Trustpilot is about *"Building trust between consumers and companies"* (Trustpilot.com, 2016a).

Being a multi-sided platform, Trustpilot delivers different value propositions to consumers and businesses. Consumers and businesses trade with each other and face a twofold problem: *"there is a twofold problem with buying things on the Internet today. It is hard for companies to figure out how to show people that they are a good company and it is hard for the consumer to figure out which companies are the good ones."* (Respondent 2, personal interview, Feb. 2016, 2:40). The customer problems mentioned here by Respondent 2 is the core of what Trustpilot base their business model on.

Consumers

For consumers, Trustpilot offers an easy way to evaluate businesses, their service and offerings. This enables consumers to make more informed buying decisions, and ultimately reduces their risk of placing orders with businesses that are unlikely to deliver what is expected. As described by Respondent 2 (personal interview, Feb. 2016) checking satisfaction ratings has become an integrated part of making purchasing decisions, when evaluating alternative vendors.

More specifically Trustpilot offers their TrustScore, a numeric score between 0-10 tied to the company's website domain name. The more satisfied customers of a certain business are, the higher the score that business will achieve. The TrustScore is presented visually in an equivalent star-rating from 1 to 5 stars, which is supported by actual review statements from other consumers (Trustpilot Support Center, 2016b). Essentially consumers are enabled to: *"[...] find the good companies"* (Respondent 2, personal interview, Feb. 2016, 2:40).

Businesses

Trustpilot provides value for businesses in two overarching ways.

First of all, businesses can improve their conversion rates throughout the selling process by displaying the TrustScore at right times. This means that more consumers will make the desired action of placing an order and consequently revenue increase for the particular business (Respondent 1, personal interview, Feb. 2016).

Secondly, Trustpilot can help businesses receive feedback from their customers and if actions are taken accordingly, improve customer satisfaction (Respondent 1, personal interview, Feb. 2016). Better customer satisfaction ultimately leads to recurring purchases and a better brand reputation, which establishes trust with potential customers.

Trustpilot offers businesses a TrustScore and a display of customer reviews from a credible source. The TrustScore and consumer reviews collected by Trustpilot are used in various

contexts to increase the consumers awareness of the business. Key exposures of Trustpilot ratings are on Trustpilot landing pages, in Google Seller Ratings and on the business' own website through implementation of widgets offered by Trustpilot (Respondent 1, personal interview, Feb. 2016; Respondent 3, personal interview, Mar. 2016). Google Seller Ratings are Google's way of indicating the advertisers quality of service. Seller ratings are based on an aggregation of customer reviews from different sources including Trustpilot (Google, 2016b).

"We help companies proactively collect reviews and get real insight straight from the people who matter most, their customers." (Trustpilot.com, 2016b)

In essence, Trustpilot enables businesses to display their level of customer satisfaction to consumers and learn how to improve their customer relationships from customer feedback.

Channels

Trustpilot reaches their customers through direct sales. That is partly through their sales force and through their website, where businesses can sign up for an account themselves (Trustpilot.com, 2016g).

Trustpilot's sales force is both performing outbound sales calls and receive inbound calls from businesses, that want to learn more about their offerings (Respondent 1, personal interview, Feb. 2016; Respondent 4, personal interview, Apr. 2016).

Their sales force focus outbound calls to businesses that potentially could benefit from Trustpilot's offerings (Respondent 1, personal interview, Feb. 2016). Businesses can become a target for outbound calls if consumers have started to leave reviews for their domain name or if the business has signed up for a free account (Respondent 1, personal interview, Feb. 2016). They might also just do business in an industry where Trustpilot have a lot of experience (Respondent 4, personal interview, Apr. 2016).

Besides the outbound calls, Trustpilot receives inbound calls from businesses interested in learning more about what Trustpilot can do for them (Respondent 1, personal interview, Feb. 2016; Respondent 4, personal interview, Apr. 2016). Inbound calls comes from businesses that either have heard about Trustpilot by word of mouth, seen Trustpilot reviews on their competitor's site, in Google Seller Ratings or if an employee at a business has been on the consumer-end themselves to make a purchase, and then been invited to leave a review on Trustpilot (Respondent 4, personal interview, Apr. 2016). In these scenarios, businesses will call Trustpilot and say, *"What is that? I want that."* (Respondent 4, personal interview, Apr. 2016, 8:26).

A different channel which Trustpilot is practising is an indirect channel, where online, web or marketing agencies refer their clients to Trustpilot when helping them improve their online efforts. The agencies will sometimes suggest the use of Trustpilot's product when advising their clients. (Respondent 1, personal interview, Feb. 2016; Respondent 2, personal interview, Feb. 2016; Respondent 4, personal interview, Apr. 2016).

Customer Relationships

Trustpilot enables their customers to be self-serviced to a large extent through the "business app" (Respondent 1, personal interview, Feb. 2016; Appendix 2). Being a business using Trustpilot's free product does not necessarily involve personal assistance, whereas businesses interested in premium products will get in touch with a sales representative from the beginning (Trustpilot.com, 2016f). When sales representatives from the sales team have come to agreement with a new customer, they become the responsibility of the customer success team, a part of the overall customer retention team (Respondent 6, personal interview, Apr. 2016). Depending on the size of the account, the customer relation is maintained and developed by account managers and integration managers (Respondent 6, personal interview, Apr. 2016; Trustpilot.com, 2016f). The larger a business is the more resources Trustpilot put into building a strong relation. Thus, smaller businesses do not receive a lot of personal assistance, whereas large businesses, key accounts with enterprise plans, has a specific account manager maintaining the relationship. (Respondent 1, personal interview, Feb. 2016; Respondent 6, personal interview, Apr. 2016; Trustpilot.com, 2016f).

In addition to self-service and personal assistance, both consumers and businesses using Trustpilot can access the Trustpilot Support Center from where they can learn from frequently asked questions and submit requests to the support team (Respondent 5, personal interview, Apr. 2016; Respondent 4, personal interview, Apr. 2016; Trustpilot.com, 2016f).

Trustpilot has a strong focus on personal relationships with their larger business accounts, whereas consumers and small business accounts are primarily offered self-service and online support.

Revenue Streams

Trustpilot is primarily basing their revenue on businesses paying for subscriptions (Respondent 1, personal interview, Feb. 2016). In addition to the subscription revenue, they are generating revenue from consumer-facing advertisements on their website, but this revenue stream is not significant compared to the subscription revenue (Respondent 1,

personal interview, Feb. 2016; Respondent 2, personal interview, Mar. 2016). *"It's purely subscription, we have a little bit of ads, but that's not kind of adding much."* (Respondent 1, personal interview, Feb. 2016, 23:31)

Between the two customer groups, it is the businesses who keeps Trustpilot running financially. However, consumers contribute substantially through their contributions to the review base. It is unlikely that consumers will ever pay for Trustpilot's services according to Respondent 2 (personal interview, Mar. 2016). Trustpilot is creating most value for businesses and they are financially stronger, so they are the paying side of the multi-sided platform (Respondent 2, personal interview, Mar. 2016).

For businesses, Trustpilot offers three different subscription tiers and a free tier (Respondent 3, personal interview, Mar. 2016; Respondent 1, personal interview, Feb. 2016). Each tier offers additional features and an increased limit for how many review invitations, the businesses can send to its customers. The tiers are called Free, Lite, Pro, Enterprise (Trustpilot.com, 2016f). The price of each tier is not fixed. Trustpilot is practicing dynamic pricing, thus they intend to make larger customers pay exactly the price that matches the value they are delivered (Respondent 3, personal interview, Mar. 2016). Trustpilot customers are given a quote from the sales representatives based on their particular needs. Despite the negotiation of prices with each customer, Trustpilot is sometimes experimenting and show guiding price estimates for the cheaper tiers on their website (Trustpilot.com, 2016g). Lite is starting at $399/month, whereas Pro is starting at $799/month. Enterprise does not have a guiding price point - instead, a quote can be requested (Trustpilot.com, 2016g). The free tier essentially allows customers to try Trustpilot with limited features before an eventual purchase. Customers running a free tier will as they start to use the product continually run into feature barriers, which prompts them to upgrade their account (Appendix 3). Thus, the free tier can be seen as an additional way to reach customers, in extension to those mentioned in the previous building block (see Channels, p. 78).

Since the revenue streams of Trustpilot is mainly based on recurring revenue it is essential for Trustpilot to not only acquire new customers but also to retain existing customers with an active subscription (Respondent 6, personal interview, Apr. 2016).

Key Resources

The digital nature of Trustpilot means that their key resources are 1) human capital and 2) intellectual capital. Unlike business with a physical product, Trustpilot does not require manufacturing or any noteworthy physical resources.

Talented human resources are of key importance to Trustpilot. Respondent 1 (personal interview, Feb. 2016) especially emphasised the importance of people experienced within product development. The respondent highlighted the challenge of finding skilled product managers. A role that implies driving the product forward by engaging multiple relevant stakeholders. *"Clearly we need to have many hands-on engineers that are smart on cloud and scale and then product is definitely the toughest job, good product managers are almost impossible to find, so that's key."* (Respondent 1, personal interview, Feb. 2016, 32:57). The respondent also pointed to the importance of the sales team and customer success team, who ensure a successful onboarding of customers and the continuous subscription renewal each year. Further emphasising that people are a key resource at Trustpilot.

The intellectual resources are also resources of key importance to Trustpilot. Trustpilot's intellectual resources can be grouped broadly into three categories: 1) the brand capital, 2) their software and 3) their review base. These are the key resources Trustpilot need in order to deliver value to customers. Respondent 2 explicitly expressed their concern with protecting their brand: *"We don't want someone to create an app, that helps people fake reviews. That would be disastrous to our brand image."* (personal interview, 2016, 14:12).

Key Activities

In Trustpilot, the key activities are effective continuous product development, sales and customer retention. Respondent 1 (personal interview, Feb. 2016) expressed that they do not see many competitive review platforms on the market, and therefore it is necessary to ensure that their platform keeps developing. They must also continuously search for market fit on new markets and make sure that people keep using the product (Respondent 1, personal interview, Feb. 2016).

Respondent 1 (personal interview, Feb. 2016) expresses the importance of having a useful product that can grow Trustpilot's brand, customer and review base:

> *[...] we need to get people to use it. So if people buy it and don't use it, it doesn't benefit us very much. We might get some money in, but in the long run, we don't get the data. We don't get the reviews, which is what drives the consumers, so we need to get that wheel*

> *running. For us, it's about getting customers in, getting them to use it, collect reviews, display reviews and by displaying reviews the Trustpilot brand is out there more and more.* (Respondent 1, personal interview, Feb. 2016: 36:47).

The key activities can be grouped into two categories:

1. **Development and maintenance of the platform**
   The continuous improvement of the platform to capture more market share in new industry verticals, which extends the brand reach and market penetration. Approximately 95 out of 500 people are working with development of the Trustpilot products including the consumer-facing Trustpilot.com and the Trustpilot Business app. (Respondent 1, personal interview, Feb. 2016; Appendix 5).

2. **Sales and retention**
   The sales and retention side of Trustpilot's operations consists of approximately 350 people out of 500 (Respondent 1, personal interview, Feb. 2016; Appendix 5). Which indicates a strong focus on bringing in new customers, ensuring that customers adopt and are successful with the product (Respondent 6, personal interview, Apr. 2016).

Respondent 1 (personal interview, Feb. 2016) stated that Trustpilot had just surpassed one billion widget impressions per month. A feat that is made possible by customers integrating Trustpilot widgets on their websites. A key priority for the customer success teams is to ensure customers integrate Trustpilot correctly on their websites. Customers can amongst others include the TrustBox widgets that show the respective website's TrustScore directly on their website (Respondent 1, personal interview, Feb. 2016). These integrations exponentially promotes Trustpilot's brand. All of which ultimately results in more sales as described in the *Channels* building block (p. 78).

Product development, sales and retention support each other and are the most important activities to make the business model work.

Key Partnerships

Trustpilot engages in various forms of partnerships. The key amongst these are found to be Google and Amazon (Respondent 1, personal interview, Feb. 2016; Respondent 2, personal interview, Mar. 2016). Less significant are individual partners from their partner program. However, combined they add up. Trustpilot categorises partners in three groups:

1) Technology Partners: the external companies building apps and integrations between Trustpilot and other tools (Respondent 1, personal interview, Feb. 2016; Respondent 2, personal interview, Mar. 2016).

2) Syndication Partners: e.g. online aggregation services, such as Boligportal.dk, one of their active syndication partners. These partners integrate TrustScores as a means of improving click-through rates (CTR) and additionally extend Trustpilot's brand reach. (Respondent 1, personal interview, Feb 2016; Respondent 2, personal interview, Mar. 2016; Respondent 5, personal interview, Apr. 2016).

3) Referral Partners: primarily web development agencies and consultants, who introduce Trustpilot to their clients as part of their e-commerce strategies (Respondent 1, personal interview, Feb 2016; Respondent 4, personal interview, Apr. 2016).

The Google partnership is the most significant. This partnership implies that Trustpilot's reviews are used as an input to an aggregated review score when Google shows merchants in their search results (Google, 2016b). In this scenario, Trustpilot engages in a form of coopetition with Google. Trustpilot actually provides review data to a potential competitor. In this scenario Trustpilot bends its own data sharing guidelines to satisfy the requirements of Google. Respondent 3 (personal interview, Mar. 2016) highlighted the positive effects of partnering with Google even though they can be seen as a competitive force. A Google partnership is too big of an opportunity to miss (Respondent 3, personal interview, Mar. 2016).

A second significant partnership is Trustpilot's agreement with Amazon as a technical infrastructure provider. Respondent 1 (personal interview, Feb. 2016) described that they own no hardware to power their infrastructure, but instead power their entire platform on Amazon Web Services. Osterwalder and Pigneur (2010) classifies this type of a partnership as a buyer-supplier partnership. The key motivation is optimisation and economy of scale, in order to minimise costs, while maximising performance. For this reason, Trustpilot's partnership with Amazon is a key partnership, as it allows them to make their website available and scalable worldwide. Besides Amazon, Trustpilot has a number of smaller buyer-supplier partnerships, that allows them to deliver their products. Those include SendGrid, Apigee and Chart.io (Respondent 1, personal interview, Feb. 2016; Respondent 3, personal interview, Mar. 2016).

Cost Structure

The strong emphasis on sales and retention testifies to a value-driven cost structure. This implies that Trustpilot is more concerned with creating value for the individual customer than they are with cutting costs. (Respondent 6, personal interview, Apr. 2016).

As elaborated in the revenue streams section Trustpilot uses dynamic pricing. This pricing strategy effectively means that Trustpilot sales representatives engage in a process of negotiation for every new customer. This process entails high costs of customer acquisition and onboarding.

In theory, a SaaS business model should allow for high levels of self-service, which results in a low cost per customer and a scalable acquisition model. However, Trustpilot's model is heavily based on one-to-one selling, which results in comparatively high costs per customer and high fixed costs, in the form of salaries. That, however, does not point to Trustpilot being a cost-driven business (Osterwalder & Pigneur 2010). It is a value-driven cost structure, which implies that Trustpilot is not as concerned with the cost implications of individual customers, as they are with ensuring that customers get the most value from the product. This focus on customer value makes sure customers renew their contract each year, with the final aim to optimise the lifetime value of each customer (Respondent 6, personal interview, Apr. 2016). The value-driven focus explains their choice to have dedicated account managers for larger customer segments, which ensures that their priority customers receive personalised service. The primary costs associated with running Trustpilot are therefore the cost of human resources, including sales teams, customer success teams and developments teams.

Summary of the Business Model

Trustpilot combines two business model patterns as described by Osterwalder and Pigneur (2010): Trustpilot is a multi-sided platform and leverages a freemium model. A visual overview of the building blocks that constitute Trustpilot's business model can be seen in Appendix 6.

Being a multi-sided platform means that Trustpilot serves as an intermediary. In a multi-sided platform, value is directly created through the interaction between the customer groups as described by Osterwalder and Pigneur (2010, p. 77): *"The platform creates value by facilitating interactions between the different groups"*. Trustpilot is the intermediary between the critical consumer and the self-reflective business. Businesses and consumer segments are respectively

receiving value through 1) a review platform, which helps customers in evaluating businesses and making purchase decisions and 2) Trustpilot Business App, a SaaS product that allows businesses to optimise customer feedback collection and communication. In addition, value is created by allowing business customers to use the Trustpilot brand and widgets on their websites. This increases their conversion rates through a heightened level of trustworthiness.

Trustpilot's freemium model is evident in their free offer to consumers. Consumers can share their experiences and learn from others completely free of charge. Trustpilot subsidises the consumers by having the business segment bear the costs related to running the platform (Respondent 2, personal interview, Mar. 2016). This model is often seen in multi-sided platforms where one side of the platform is attracted through a free offer, while the other side is attracted by the presence of the first side (Osterwalder & Pigneur, 2010). The presence of the consumer side makes the platform much more valuable to Trustpilot's business segment.

Trustpilot acquires new customers through outbound calls to businesses that have received reviews on Trustpilot. The display of TrustScores on customers websites, customer usage of the Trustpilot automatic invitation system and display in Google search results are huge drivers of attention and extends Trustpilot's brand reach.

The relationships Trustpilot has with their customers varies depending on the size of the individual customer. More significant customers have dedicated account managers, whereas smaller customers service themselves mostly through self-service after the initial sale.

Trustpilot captures value from businesses through a dynamic pricing based subscription model for businesses, in an attempt to maximise the value captured from each customer. Trustpilot does not yet have a way to significantly monetize the traffic they drive, they show ad's on the website but that amounts to a very little share of the total revenue contribution. Respondent 3 (personal interview, Mar. 2016) expressed awareness of Trustpilot creating more value than it captures, as Trustpilot is publically available and anyone is able to base their buying decisions on their data without paying for Trustpilot.

The key activities required for Trustpilot to deliver value and to keep their business running is 1) product development and 2) sales and retention. The first refers to the continuous development of the website and technology platform, and the second refers to the sale and assuring that customers are getting value from the product and keep renewing their subscriptions. Trustpilot engages in key partnerships with Google and Amazon. The partnership with Google is a form of coopetitive relationship where Trustpilot is delivering review data programmatically to Google in order to extend their brand reach and value to

their customers. Their relation to Amazon is a buyer-supplier relationship, as Trustpilot buys its underlying IT infrastructure from Amazon. This infrastructure allows Trustpilot to deliver its value proposition worldwide.

Lastly, Trustpilot is characterised as having a value-driven cost structure, their primary aim is to maximise the customer lifetime value. This implies that they are not as concerned with the immediate costs associated with each customer but rather with keeping the customer. Running Trustpilot requires high fixed costs in the form of salaries.

The Role of Data Sharing in the Business Model

As established in the data spectrum section, Trustpilot shares data through APIs with customers and partners.

Trustpilot's APIs have a significant impact on the business model of Trustpilot. As elaborated in the *Obtaining access* section (p. 69) the APIs are provided to customers with a premium subscription tier, as well as partners who have been through a process of scrutiny, where they declare what they intend to build with the APIs and data.

Data access plays an important role in creating and capturing value for Trustpilot (Respondent 2, personal interview, Mar. 2016). Customers are enabled to utilise the data as desired, while Trustpilot captures more value through the premium that is paid for data access. Thus data sharing is an integral part of Trustpilot's business model. If Trustpilot were to open their data, it would also imply free data access for customers and thus undermine mechanisms of capturing some of the value created through the use of their data. Although Trustpilot captures some value from the use of their data, Respondent 3 (personal interview, Mar. 2016) still highlights the unmet challenge of figuring out how to maximise value capture from data sharing.

An important driver behind Trustpilot's APIs is the desire to the retain customers. This is partly possible through integrating deeply into business customers' system and processes, which creates barriers for customers to leave Trustpilot - effectively creating a lock-in on customers (Respondent 6, personal interview, Apr. 2016). However, businesses with deeper integrations are also able to extract more value from the product (Respondent 6, personal interview, Apr. 2016). Data sharing therefore broadens the value proposition and more value is created for the business customer, for instance through integrations to customer service software and e-commerce systems (Respondent 1, personal interview, Feb. 2016).

With technology partners, the specific returns from data access are less quantifiable, as they do not pay for the access (Respondent 1, personal interview, Feb. 2016). The hope is that technology partners help to make Trustpilot more useful for their customers, through integrations and apps. Integrations and apps allow salespeople to better address customer concerns when closing deals and limit the friction in onboarding new customers (Respondent 6, personal interview, Feb. 2016).

Respondent 2 and 3 (personal interview, Feb. 2016; personal interview, Mar. 2016) suggested that integrations made by partners free up resources from their teams and allows Trustpilot to focus on their key activities such as the development of the core product. Respondent 3 (personal interview, Mar. 2016) sees the opportunity for partner integrations to help Trustpilot capture market share in new regions. The development of integrations to local e-commerce platforms is extensive and costly. Through the partner program, Trustpilot is able to leverage the expertise and resources from developers in their customers' local environments. Thus, if most Korean webshops use an e-commerce solution of Korean origin, Trustpilot can leverage local developers to build, maintain and support integrations. Trustpilot can only cover the largest e-commerce solutions with their integrations, which amounts for the majority of the market (Respondent 3, personal interview, Mar. 2016). However, the last share of the market is using thousands of different e-commerce solutions, which makes partners valuable for Trustpilot in increasing their global market share.

Although partners can free up resources for Trustpilot's development teams, sharing data with partners introduces a new set of activities (Respondent 2, personal interview, Mar. 2016). These activities include promotion, support and governance. Respondent 1 and Respondent 2 (personal interview, Feb. 2016) emphasises the importance of and heavy resource requirements of having an active approach to data sharing through their partner program.

> *[...] what I've seen other companies doing - well…. just opening data doesn't work. Everyone I've ever heard doing this, says that it doesn't trigger. You need to do something else. I think you need to do something on top. So you need to advertise it. Make it add value for people that would do it.* (Respondent 1, personal interview, Feb. 2016, 71:22)

Respondent 2 (personal interview, Mar. 2016) suggests that Trustpilot must work out a framework of how to deal with partners who use their data. Such a framework is needed to ensure that partners are offered technical support and that expectations are aligned. The framework also requires handling partners much like customers are handled in a CRM system (Respondent 2, personal interview, Mar. 2016).

In addition to the resource requirements, Trustpilot is concerned with how data consumers might alter the representation of the data (Respondent 1, personal interview, Feb. 2016; Respondent 2, personal interview, Mar. 2016; Respondent 6, personal interview, Apr. 2016). For example by altering the TrustScores or reviews to mislead consumers. Such actions have the potential to damage the integrity and trustworthiness of Trustpilot's brand. The Trustpilot brand is intellectual capital, which Trustpilot must govern since it plays an important role in delivering their value proposition to consumers (Respondent 2, personal interview, Mar. 2016; Respondent 4, personal interview, Apr. 2016). This means that monitoring the use of data and taking action to ensure compliance is necessary. An activity that grows as data is shared with more actors - ultimately raising the costs of the data sharing program.

Being a platform

We have determined that Trustpilot is sharing data with their top-tier customers as well as with approved partners. Additionally, Trustpilot's business model has been examined to understand how Trustpilot benefits as a business from sharing their data. The nature of their business model matches the pattern of a multi-sided platform. Trustpilot hopes to see more partnerships in the future so that their customers can receive even more value while limiting the development effort from Trustpilot's side.

We employ a platform perspective to further understand the mechanisms of driving a multi-sided platform and the attempt to move Trustpilot's value creation beyond a single product. Platform theory complements the findings from the business model perspective and sheds light to additional implications of data sharing. A platform perspective contributes further to the insight which is constructed about Trustpilot and their justification of data sharing.

Trustpilot and its Traits as a Multi-sided Platform

To better understand the characteristics of Trustpilot as a multi-sided platform, it is necessary to understand the dynamics of a multi-sided platform and how data sharing affects these. In the following, we take a closer look at network effects, switching costs and positive feedback loop.

In the case of Trustpilot, indirect network effects have a significant impact on business customer attraction. The indirect network effects arise from users reviewing businesses - which makes it valuable for the reviewed businesses to engage on the platform and interact with the dissatisfied consumers (Jensen, 2014).

Direct network effects, also known as same-side network effects, also apply. The users of Trustpilot publishing reviews, increase the value for other users on the platform. The reviews are the foundation of Trustpilot and essential for Trustpilot to be of value. Therefore, the direct network effects are significant. Respondent 1 (personal interview, Feb. 2016) stated that if no consumers have left reviews on a business profile, other consumers will not find the business profile of any value. Thus, it is crucial that a critical mass of reviews is in place, as that is paramount in attracting other consumers and find Trustpilot useful (Shapiro & Varian, 1998).

The continuous contribution of reviews from users increases the size of Trustpilot and makes it more valuable to consumers and businesses. The large review and user base enables Trustpilot to sell their product to businesses. All of which again generates more traffic for Trustpilot resulting in a positive feedback loop that implies that Trustpilot gets significantly stronger as the user base grows. Trustpilot is consciously pursuing a position as a market leader: *"we believe that there will not necessarily be a lot of these kinds of review platforms out there […] so for us it's about taking this market cut before anyone else comes in"* (Respondent 1, personal interview, Feb. 2016, 36:47). Trustpilot effectively exploits network effects to grow their market share and attract both ends of the platform.

The Effect of Data Sharing

As identified in the analysis of their key activities, Trustpilot has a strong emphasis on integrating deeper into the workflows and systems of their customers, in order to maximise the lifetime value of customers (Respondent 3, personal interview, Mar. 2016; Respondent 6, personal interview, Apr. 2016). These integrations happen in the form of custom integrations, Trustbox widget integrations and AFS[16], a service designed to automatically send out review invitations to end-customers after successful purchases.

Custom integrations between Trustpilot and customer systems are enabled by the data provided through Trustpilot APIs. Which implies that customers have to spend resources on training staff and building integrations, meaning that customers incur procedural switching costs. These increase the customer lock-in (Shapiro & Varian, 1998). Custom integrations combined with customer use of AFS implies that significant procedural and financial switching costs will incur, should a business desire to change for a competitor, as they are not allowed to transfer their reviews to a different review service.

[16] Automatic Feedback System

Data sharing means that Trustpilot business customers are increasing their own switching costs, by utilising the Trustpilot platform. The customers are thus making it harder for themselves to shift to a different review product.

Strong positive network effects are in most cases very desirable for a multi-sided platform like Trustpilot. Trustpilot's data sharing enables deep integrations, partly explains their incentive to share data with customers. The data sharing allows them to exploit and accelerate the direct network effects through the customers of their business customers, by integrating deeply into their websites and inviting their customers to leave a review. Their review contribution adds to the indirect network effects, as the larger consumer base attracts more businesses. Trustpilot leverages these network effects to attain their goal of becoming a market leader within the review space (Respondent 1, personal interview, Feb. 2016).

APIs as an Internal Platform

As most digitally advanced businesses today, Trustpilot base their software on an architecture where the user interfaces are decoupled from the underlying functionality. This kind of architecture gives flexibility for how and under which conditions the underlying functionality is used. The interface to the underlying functionality rarely changes, thus the customer-facing products or features simply interact, through APIs, with what Baldwin and Woodard (2009) would refer to as the "core". Respondent 3 highlights the modular features of their architecture in this statement: *[... you can actually build our entire product in Google Spreadsheet if you wanted - just by writing some kind of Google Script. Or you can write an entirely new interface for our product yourself, if that's what you wanted.]* (translated from personal interview, Mar. 2016, 7:18). Trustpilot's APIs can be seen as an internal platform as defined by product development theorists, where a platform is seen as a number of common elements which can be implemented across a range of products (Gawer, 2010). The common elements are the APIs which represents different components such as reviews, consumers, invitations and business units as described in Accessible data section (see p. 71). The individual APIs can be cherry picked and used as required by the context.

Trustpilot redesigned their architecture to an "API-first" architecture in the period from 2013 to 2015 (Respondent 3, personal interview, Mar. 2016). The purpose of re-engineering Trustpilot was to benefit from the advantages associated with modularity. Respondent 3 particularly mentions the benefits of an "API-first" architecture: *[... So it is nice because teams can work independently. It is possible to build smaller chunks of code, which a team can work on independently of releases.]* (translated from personal interview, Mar. 2016, 7:18). In principle,

Trustpilot APIs enables Trustpilot to work on many new product or feature ideas simultaneously (Respondent 1, personal interview, Feb. 2016). As Trustpilot's internal organisation grew, the number of development teams grew with it. It became hard to sustain an efficient way of developing different features simultaneously. The API architecture thus made it possible to scale the development organisation and keep it efficient in working concurrently (Respondent 3, personal interview, Mar. 2016).

Besides enabling Trustpilot's own development of new features, they were also aware of the possibilities APIs open, for others to hook into Trustpilot's features, functionalities and product (Respondent 3, personal interview, Mar. 2016). An idea covered in the next section.

Towards an Industry Platform

Trustpilot's APIs are the foundation of all products and features they offer to their business customers and end-consumers (Respondent 3, personal interview, Mar. 2016). But Trustpilot also has a vision of letting other companies use its APIs to build new products or to extend existing products. Such a setup has been attempted through the partner program, which enables external parties to integrate their products with Trustpilot's review engine or build entirely new products, which Trustpilot would not have thought of themselves (Respondent 2, personal interview, Mar. 2016; Respondent 3, personal interview, Mar. 2016). The respondents tend to agree that there is potential in including external parties and let them build on top of Trustpilot's APIs. Respondent 2 is positive of the opportunities and indicates that Trustpilot must be more than a tool if they want to succeed in the long run: *"But it is also the difference between being a platform and being a tool. Uhm... We want to be how people manage word of mouth on the internet. And you can't do that if you are just a tool."* (personal interview, Mar. 2016, 12:34). This perspective aligns with Cusumano (2010), who advocates for platform theory and suggests that businesses must see beyond product strategy and start to think more in the lines of platform strategy if they want to sustain their success in the long run - especially in turbulent industries.

Trustpilot expresses that turning their APIs into an external platform, that other businesses can leverage, requires mutual benefits. Partners must be able to see a way to benefit from integrating with Trustpilot and Trustpilot must be able to see a benefit from enabling them to do so (Respondent 1, personal interview, Feb. 2016; Respondent 2, personal interview, Mar. 2016). Respondent 2 puts it like this: *"Because when they win, we win. And my goal is to make it that, when we win, they win."* (personal interview, Mar. 2016, 53:20).

The businesses who are customers of Trustpilot does not only use Trustpilot's products when it comes to managing their customer interaction. They usually have an array of software, which supports activities related to customer interaction. Such systems include CRM systems, social media management tools, e-commerce systems, support ticketing systems and emailing systems. Therefore, the enablement of integration between Trustpilot and such systems allows businesses to interface with consumer reviews in the settings they require. For instance, Trustpilot recently launched an integration to Hootsuite (Respondent 3, personal interview, Mar. 2016). Hootsuite is a social media management tool for businesses, which enables scheduling of posts across various social media. This integration pulls the business' latest Trustpilot reviews into Hootsuite. A review can then be selected and easily posted to the business' social media. In this example, Trustpilot increases the usefulness of the Hootsuite application while promoting their brand in the Hootsuite App Directory and increase the distribution of their reviews on social media.

In addition to the system-to-system integrations, Respondent 1 and 3 (personal interview, Feb. 2016; personal interview, Mar. 2016) also acknowledge businesses need for combined data in analytics to get a holistic view across their different data sources. By giving partners who work with data aggregation access to Trustpilot's data, their customers can fully utilise such holistic analytical tools.

Even though there exists no clear path for how the partner programme creates a tangible return on investment, Trustpilot sees potential in intangible returns from co-marketing, customer crossover, better customer retention, customer subscription upgrades and freeing up developer resources (Respondent 2, personal interview, Mar. 2016). Integrations help to make Trustpilot more valuable for its customers and thus makes them more likely to stick with Trustpilot as a product (Respondent 2, personal interview, Mar. 2016). Whenever a partner includes Trustpilot's TrustScore or reviews in their product, it means Trustpilot gets exposure. And as suggested by Respondent 2 (personal interview, Mar. 2016), the more Trustpilot reviews are seen, the more valuable they become. Additionally, the exposure of Trustpilot's brand in other products function as a lead generator: *"So customers of theirs, who are not customers of ours, might become customers of ours. We can get leads that way."* (Respondent 2, personal interview, Mar. 2016, 10:44). Respondent 2 also sees the potential for innovative use of their APIs as a platform:

> *If you allow people to build things through you, then you can make your tool useful in ways that we would have never thought of. If anybody in the world can figure how they want to use reviews and customer service, that kind of tool, in any way that they want. I*

> *mean the rest of the world is much smarter than the people in this building. They are going to have ideas, we don't.* (personal interview, Mar. 2016, 12:34)

An important proposition of industry platforms is described here, which is that of innovative capabilities. Gawer (2010, p. 20) suggests that the objective of an industry platform is: "*To stimulate and capture value from external, complementary innovation*".

In addition to the mentioned benefits, Trustpilot also sees a benefit in reducing development costs (Respondent 2, personal interview, Mar. 2016). Through a platform, several markets and segments can be targeted with niche products and developing all these niche products would be costly. Therefore, Trustpilot wants to utilise partners, which in platform terminology can be referred to as complementors (Gawer & Cusumano, 2008).

Trustpilot can be categorised as a *platform-leader wannabe,* which refers to a company that wants to become a platform leader (Gawer & Cusumano, 2008). They want to be the future of how people manage word of mouth on the internet (Respondent 2, personal interview, Mar. 2016) and they acknowledge the need to include complementors to continue their success and build a business ecosystem around them. Trustpilot still find themselves in a position, where they must develop the complements for their platform themselves, and be complementors to other platforms (see Figure 14). Respondent 3 (personal interview, Mar. 2016) highlighted that they are in a power battle, where the bigger players, such as their Google partnership, can exercise their power over Trustpilot and contrarily Trustpilot can exercise their power over the smaller players in the ecosystem.

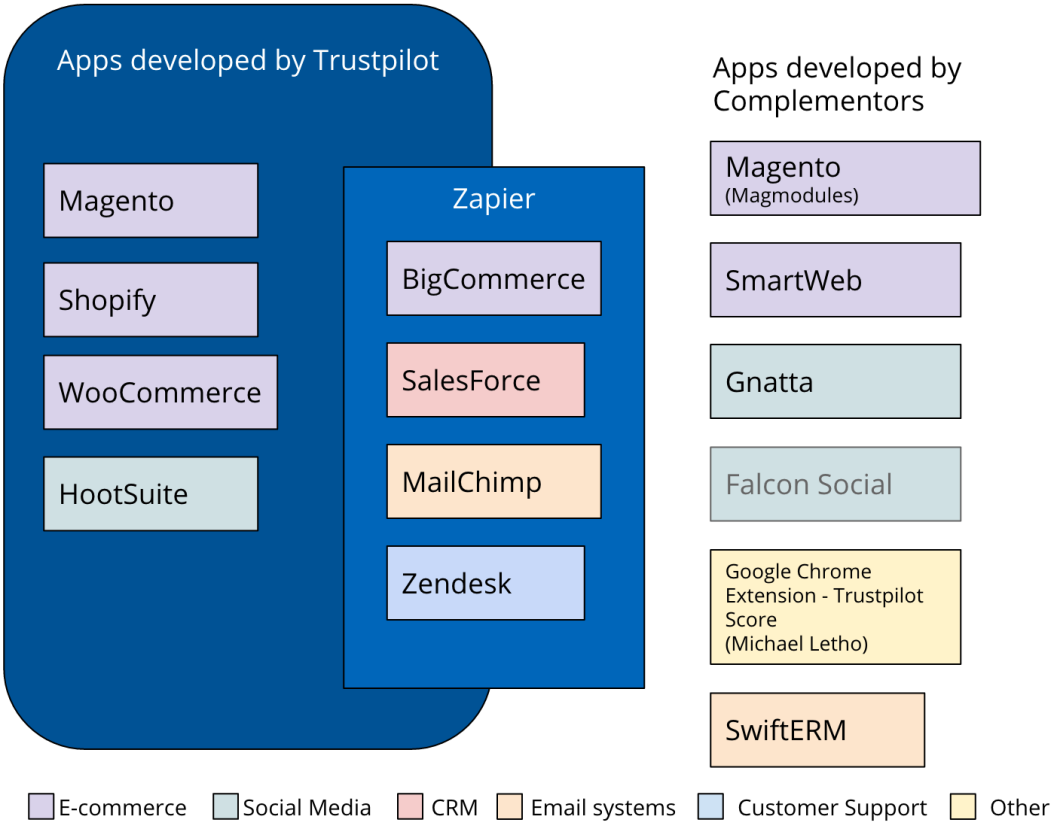


Figure 14: Trustpilot Complements by Complementor
Constructed by authors based on personal interviews and Trustpilot.com (Trustpilot.com, 2016h)

From the apps and integrations available, Trustpilot has yet developed most of them. However, Trustpilot starts to see more interests from complementors, who adopt the Trustpilot platform in their products. Gnatta and SwiftERM are examples of such (Respondent 3, personal interview, Mar. 2016; Respondent 4, personal interview, Apr. 2016). Trustpilot's sees the benefit of having complementors, which extend the use of Trustpilot beyond their own tool: *"[...] anything they are developing we don't have to. And in that sense, they do make us a lot stronger."* (Respondent 2, personal interview, Mar. 2016, 53:20), *"We believe these partnerships are valuable, and could be valuable, and the ones we have are valuable, but they are not big [...]* (Respondent 1, personal interview, Feb. 2016, 63:47).

It is evident that Trustpilot sees potential in sharing their data with partners through their API. When successful, the partner program helps to create additional value for Trustpilot's existing customers, in ways that Trustpilot could not have imagined themselves. In addition to

the value creation for existing customers, it offers opportunities of growing Trustpilot, through co-marketing with complementors and the exposure of Trustpilot as a brand which naturally follows when TrustScores and reviews are distributed across a larger array of complementing firms.

Respondent 2 sums the potential of sharing their data with partners through APIs: *"That review that used to live on Trustpilot now can go anywhere and you can spread it to so many more people and our customers get a lot more value out of that."* (Respondent 2, personal interview, Mar. 2016, 9:36).

Granting complementors access to Trustpilot's data also introduces the risk of direct competition between Trustpilot and its complementors. Respondent 2 (personal interview, Mar. 2016) sees products from partners that compete with the core features of Trustpilot Business App as a potential source of conflict: *"That would obviously be a huge problem if that happened."* (Respondent 2, personal interview. Mar. 2016, 26:20). However, he does not find it likely that such situations would occur, since Trustpilot engage in partnerships with businesses that are complementary in nature, such as social media tools. If it created a competitive situation, Trustpilot does not have measures in place but indicated that the partner could be cut off access or acquired.

Analysis Summary

In our analysis, we have found that Trustpilot offers data to the external world. However, this is limited to top-tier customers and approved Trustpilot partners. The use of the data is governed by a set of guidelines, which restricts application. Due to these restrictions on access and use, Trustpilot's data does not conform to the definition of open data. The access Trustpilot is offering to its data is better categorised as shared data in accordance with the data spectrum.

Data sharing plays an important role in Trustpilot's business model. Sharing data enables value creation for customers through application of reviews in the setting most appropriate for their customers' workflows and processes. Trustpilot partly captures the additional value added through the requirement of a top-tier subscription to gain access. Additionally, data sharing foster a network of syndication and technical partners, that display the Trustpilot brand, TrustScores and reviews. This exposure works as an additional channel for generating new customers.

We establish that Trustpilot is a multi-sided platform. Trustpilot leverages data sharing as a mean to gain more widespread use of their product, which further makes their platform more valuable through intensified network effects. These dynamics contribute to a positive feedback loop, where more consumers lead to more reviews, and more reviews lead to more consumers and paying Trustpilot customers. The positive feedback loop is made possible by encouraging business customers to integrate with Trustpilot through their data sharing initiatives.

Trustpilot is attempting to create an industry platform and involve the surrounding business ecosystem. We classify them as a platform-leader wannabe, since they are still building many complements themselves. Trustpilot sees potential in having complementors build unexpected innovative solutions based on their APIs. Additionally, Trustpilot sees several other benefits from sharing data with partners such as lead generation, customer retention and savings in internal development resources, while expanding to new and heterogeneous customer segments.

Discussion

In this section, we elaborate our findings to highlight its meaning. We compare and contrast our findings to what other researchers in the open data domain have addressed. Additionally, we discuss the limitations of the study, implications and suggest further research.

In this study, we posed the question of *"How is external data sharing justified in business?"*. Through analysis of Trustpilot, we have categorised their data sharing initiative and identified business justifications for granting certain customers and partners access to data. In the following, we recap the results and discuss their meaning.

Justifying external data sharing

Examining Trustpilot, we have found that justification for external data sharing in business can be grouped into four overarching categories:

Increase Customer Utility
Data sharing has shown to have a compelling impact on the business model. Data sharing enables deep integration between Trustpilot and their business customers, allowing for more value creation through better customer workflows. Additional value creation allows for higher pricing and value capture through more revenue.

Extend Brand Exposure
External data sharing is found to be a way to increase brand awareness. The exposure of data in combination with attribution to the data source creates extra marketing channels for the business. Customers' and partners' display Trustpilot brand, TrustScores and reviews on websites, search engines, in apps and app stores, which significantly boosts Trustpilot's exposure to new and existing users.

Amplifies Platform Dynamics
External data sharing is found to help increase the intensity of network effects. Through integration with customers, Trustpilot leverages their customer bases. Integrations also lead to increased switching costs for Trustpilot customers and creates barriers to leave the platform.

Leverage Complementors

Sharing data with the business ecosystem extends product development beyond internal development teams. Thus, Trustpilot's customers can benefit from a wider range of integrations and innovative applications of Trustpilot reviews. Complementors, thus, create significant value through value adding complements. Additionally inclusion of external developer resources, allows Trustpilot developer teams to focus on core products and functionality.

Hindrances for openness

Additionally, our examination of Trustpilot have indicated a number of hindrances for further openness of Trustpilot's data:

Undermines Value Capture

Opening data could undermine the existing business model as customers currently are required to pay for a top-tier subscription to gain access to Trustpilot's data. If data was free for everyone, the incentive to upgrade could be reduced.

Loss of Control

Opening data has unpredictable outcomes. Thus, the use of Trustpilot's data could potentially hurt their business. Complementors in the business ecosystem could pose a threat. They can cannibalise the existing value proposition of Trustpilot, by building products that compete head to head with Trustpilot's revenue-generating offerings. Limiting data access is a way to stay in control of events.

Adverse Effects on Brand

Misrepresentation and misuse of data are big concerns for Trustpilot. Trustpilot's intellectual capital is a key resource. Opening data means that data users with bad intentions can alter and display reviews in a misleading way, which can potentially impact the integrity of Trustpilot.
Additionally, when third-parties build apps or integrations using Trustpilot's data, consumers might be confused with Trustpilot's involvement, and connect it to their brand. Solutions of poor quality have the potential to affect consumer's view on Trustpilot.

Open is not enough

Offering data to external parties is not an initiative which is carried out without an expectation of positive outcome. If others are to be motivated and successful in utilising

Trustpilot's data, it requires that Trustpilot promotes an eventual open data offering, support data users, while making sure data is not misused. These activities increase costs for the business.

Through our examination of justifications and hindrances for opening data, we have observed that the upsides of complete openness are not perceived to outweigh the challenges introduced. The choice of a limited data offer implies a controlled approach to external data sharing. Trustpilot does not immediately feel attracted to complete openness as they can balance their returns through a restricted data offering.

Why share data?

This study took a closer look at Trustpilot, an online business. Trustpilot is looking for ways to spread their review platform to all companies that sale products or services online. Furthermore, they are expanding into the space of offline businesses. As Trustpilot realise the importance of serving their customers as they require, they find it necessary not to restrict customers to only work with reviews within the boundaries of the Trustpilot Business App.

We have learned that through sharing data, Trustpilot is able to foster additional value creation for their customers. Some of that value is captured through a subscription based revenue model. Trustpilot has experienced that their customers request integration with business systems they use, which has lead to Trustpilot building a few of such integrations. However, Trustpilot realises that they do not have the resources to develop integrations to all business systems that their customers could possibly use, and only the largest enterprises have the competences and resources to do so themselves. Trustpilot also acknowledges that they will be unable to come up with the best ideas for how to find use of their data. The world will have ideas, that they do not. From these findings, it becomes clear that Trustpilot is particularly focused on their customers and how data can create more value for just that group of organisations.

When governments open data, there seems to be a less clear view of who the data user will be. It could be individuals, businesses, NGOs or others (Attard et al., 2015). If we analogue Trustpilot's focus on customer value creation to governments: The customer of a government would be its citizens and businesses - those paying taxes. Citizens and businesses are paying

taxes for the government to deliver a value proposition of a healthy, developed and well-regulated society. From this perspective, all citizens can be seens as customers of the government. This could explain why governments do not find it necessary to apply restrictions on data, where businesses would. A government cannot attract more taxpayers by limiting access to its data. On the other hand, Trustpilot sees a potential in limiting data access to increase the value captured from customers that use Trustpilot.

There is no doubt that Trustpilot sees the potential to generate value from granting external parties access to their data and to some extent they capture that value from customers through subscription payments. A clear return on investment from offering data access is not present, since customers can have many reasons to upgrade their subscription.

These insights into Trustpilot are important from the perspective of open data. If we hope to understand why companies like Trustpilot restrict access and use of their data, we must understand what they gain from doing so.

If a close look is taken at the data offered through Trustpilot's APIs and from what we heard from the Respondents, the data found in the Public APIs is also available on their website. Their data could potentially be crawled directly from their website. So why is it that they prevent outsiders from accessing that same data through the APIs? When discussing open data the focus is on setting data free for unrestricted use (OpenDefinition.org, 2015), but Trustpilot is not interesting in everyone making use of their data. They are looking for data users, that can create value for both themselves, Trustpilot customers and Trustpilot as a business. It would seem governments that open data are not worried about others benefitting from data without them benefitting. It seems governments believe that it is going to come back to them anyway through advancements in society at large. Governments could however find motivation in advancing their region through data in order to strengthen its region's position in the global economy. A global economy where we see countries and unions compete to be the strongest economy.

Why not Open Data?

The identified justifications for sharing data helps provide deeper insight into why we are see a business share data in a loosely coupled manner and sheds light to the reason to, why the ideal state *'open data'* is not found in business. Answers to the ultimate state still remains: could we

ever see truly open data from a business? We found that the undermining of value capture, loss of control and adverse effects on brand are significant hindrances that creates challenges Trustpilot cannot grasp.

To see valuable open data from businesses: that is data that can be leveraged to create value at large, the fundamental assumptions of successful business operation are challenged. Giving away something so potentially valuable as open data does not align well with the business objective, to maximise shareholder value (Brandenburger & Stuart, 1996; Hillier et al., 2011). Since open data creates significant value that might not readily be captured. A perspective that emerge from our study is the urge for a close to one-to-one relationship between value creation and value capture. Meaning that value created should also be captured. If data is published to the public then capture of value has to be justified.
Value creation is defined as *"the contribution to the utility of the final good or service to end users"* and value capture as *"the difference between revenue and cost retained by the firm"* (Pagani, 2013, p. 618). Assuming, businesses need to maximise value capture, to increase shareholder value, it is necessary to look at alternative options for value capture through open data. With value creation being an intangible and subjective construct. In the sense that, the customer decides what a good is worth to them.

It therefore makes it necessary to find answers to how intangible matters such as innovation through open data can be incorporated in a business, otherwise focused on tangible results.

With data being bits and bytes, data is close to free to replicate and often implies an almost non-existing marginal cost. The whole paradigm seems to be challenging the notions of scarcity. Thus, it is hard to grasp, as it challenges our traditional understanding of economics, where resources are scarce and trade of resources are based on economic transactions. Open data in itself is an unpredictable phenomenon, that business seem to see as a daunting subject. In a growth business like Trustpilot, where tangible results are measured and matter, directions cannot be based on intangible and unpredictable results.

We observed an artificially induced scarcity through governing the access to data, which further increases the complexity of addressing the question of why we are not seeing open company data. O'Reilly (2012) argues that the capitalistic focus on value capture in society measured through GDP does not take into account the health or wellbeing of the society. And that great innovations often are started by people, that were not driven by value capture and therefore O'Reilly (2012) argues, that value creation should be the primary concern. This also implies looking at value over the long term instead of short term gains, which are evident in investor behaviour we see today (O'Reilly, 2012). The concern with the immediate short term

was noticeable at Trustpilot through their inherent focus on capturing value from data right away. We therefore assert that it is necessary to look into new ways of measuring value and dealing with the intangible benefits is necessary, to motivate true open data from a private company.

Relation to Extant Research

We contribute to a new direction in Open Data research: Open Company Data. A new direction where businesses are studied to understand their potential as open data providers. To further expand our findings we contrast and compare our findings to existing directions in the open data literature.

Interlinking and Exchanging Data

As addressed in the Research Domain section (p. 15) a large body of Open Data literature is concerned with the technical challenges of linking data across the World Wide Web, some refer to this as the Semantic Web (Berners-Lee, 2006; Bizer et al., 2009b). From this study we have experienced that a business like Trustpilot is not concerned with the interlinkage between their data and data residing on other websites around the world. Doing so would require adherence to the standards of RDF, which includes assigning data identifiers into the markup for their website using URIs (Berners-Lee, 2010).

Rather Trustpilot is concerned with integrating between systems across the digital business environment. In their world data connection between organisations are based on APIs. Thus, Trustpilot would seem to be much more under influence of what is also mentioned as the API economy in practice (Holley et al., 2014; Willmott et al., 2013). There lies a potential challenge in coordinating and combining data sets across the business world and the grass root organisations contributing the the LOD Cloud (see Appendix 1). The LOD organisations are building one standard for organising and connecting data across the web, while businesses lean towards a data exchange via web services in the form of REST APIs[17]. Investigating differences and incompatibility between the two standards is beyond the purpose of this study, but it has become clear that Trustpilot is not concerned with LOD and the requirements of the 5 star model presented by Berners-Lee (2010). They strongly depend on APIs - what would seem a predominant standard in online software and enterprises, for exchange of data.

[17] REST APIs: Representational state transfer

Reasons for Sharing Data

Several contributors to the open data literature have addressed what we categorised as opportunities of Open Government Data (Attard et al., 2015; Janssen et al., 2012; Ubaldi, 2013). We found that the primary benefits associated with OGD could be grouped into transparency, economic growth and technical benefits. Governments and businesses have very different purposes and serve very different stakeholders, thus perceived benefits of opening data cannot be expected to be the same.

It is well known from the literature and numerous practical initiatives that governments find motivation to open data. However, businesses do not open data. Governments are essentially concerned with the overall wealth and wellbeing of citizens in society, while businesses are essentially concerned with satisfying stockholders through profitable results. We discuss the differences between what motivates governments to open data and what motivates businesses to share data.

Attard et al. (2015), Janssen et al. (2012) and Ubaldi (2013) all argue that governments open data partly because they feel obligated to offer transparency. In a democracy, society should know what is going on inside the public sector. This upside of opening data would seem rather intangible and is not directly related to an economic incentive, but rather an ethical incentive. However, transparency can help societies to counter corruption, which damages societies economically (Attard et al., 2015). Businesses like Trustpilot does not hold the same obligations towards society. Trustpilot could potentially benefit from offering transparency to address some of the media publicity (Clausen, 2016), where they are accused of improving TrustScores of their paying customers. Offering a full review dataset, that informs which reviews that have been included or denied by Trustpilot, could give transparency to their review verification process. However, Trustpilot would unlikely be interested in defending individual review approval assessments. Transparency to the outside world did not show to be an explicit concern in our study on Trustpilot. Thus, our study indicates, that the transparency argument for open data is not key for business.

Economic growth, including entrepreneurship and innovation is another upside of open government data highlighted in literature (Janssen et al., 2012; Ubaldi, 2013). Governments have a general interest in economic growth in society, especially if the economic growth of their particular region outpaces that of other regions. Opening data to the society supposedly offers an opportunity for both social as well as commercial innovation (Attard et al., 2015; Ubaldi, 2013). If data is utilised, it is expected to improve decision making and increase efficiency. While governments are focused on economic growth in society, we found that

Trustpilot is primarily concerned with their own economic growth as well as an indirect interest in their customers' economic growth. If they can offer data to their customers, in a way that enables them to improve their businesses, Trustpilot sees an opportunity to capture more value from those customers. While incentives for innovative use of data seems different between businesses and governments, there is a commonality in the wish for creating economic results. The governmental intent to empower entrepreneurs and other innovators in making use of their data to create economic growth is comparable to what Trustpilot is attempting through their partner program, where they want to leverage complementors. Trustpilot hopes to see these complementors make Trustpilot's review base even more valuable to its business customers so that they can benefit from the growing number of top-tier subscriptions. In the same way, governments hope to see entrepreneurs and existing businesses succeed in utilising government data to improve society and grow new businesses that can compete and generate jobs and economic and social wealth.

As for technical benefits (Janssen et al., 2012) of open or shared data, we have not seen a particular emphasis of such upsides in our case study on Trustpilot. Their customers see advantages in being able to easily connect directly to Trustpilot's data through APIs, and they are willing to pay for that. As Trustpilot's data is not open, it is primarily their customers that benefit from the operational and technical advantages, that their APIs offer.

Besides the benefits we just addressed, which originated from research on open government data, we have found that Trustpilot finds alternate benefits from sharing their data. Trustpilot sees potential in enforcing their position as a multi-sided platform as well as the hope to achieve more brand exposure. We question if these upsides of sharing data could also apply to governments or other institutions that share data. The open data definition is open for requirements of attribution, and thus, there is a potential for benefiting from brand exposure, when opening data. Governments are not particularly on a mission to market their "brand", which might explain why brand exposure is not mentioned as a potential upside of open government data. The various surveys (Open Knowledge Foundation, 2015; World Wide Web Foundation, 2015) that offer an overview of which countries that offer more open government data, definitely makes it apparent in which countries open data can be utilised. Companies that have an intent to base their business on open data might be attracted to start in or move to that particular country.

Although the motivation for sharing data with the surroundings varies between a business like Trustpilot and governments, there is also commonalities. There exist a common understanding that offering data to surroundings has the potential to foster innovation and

new solutions, products and services that create more value. Governments are less concerned with how to make sure they capture some of that value, while that is a primary concern for a business like Trustpilot.

Value Generation

Jetzek et al. (2013) argue that a significant contributor to value generation from open government data is the openness in itself. The release of open data means that data is reframed through a network of actors making sense of the data and extracting value from it in a multitude of different contexts. Further proposing a framework with four archetypical mechanisms that transform data into value: information transparency, collective impact, data-driven innovation and data-driven efficiency. The last two are particularly relevant for the case in question.

We observe data-driven efficiency and data-driven innovation as a significant value generating mechanisms and motivators for data sharing in Trustpilot. The data-driven efficiency occurs through sharing of data to their customers, which in turn creates efficiencies through customised integrations in workflows. This will in many cases lead to cost savings and additional value creation. A hypothetical scenario in practice: a customer service representative calls a dissatisfied customer. By having their recent reviews readily available in their customer service software, the representative can quickly understand the concerns through the reviews and address them, instead of retrospectively having to look up the name of the customer and find their review. This is only one possibility out of the many new efficiencies the data sharing from Trustpilot enables.

However, what is also noteworthy, is that the data Trustpilot provides is not open and yet customers can extract significant value from it. Trustpilot data is a small source of data in the big picture of customer communication. The context is, however, different from the context in which Jetzek et al. (2014) derived their mechanisms. They based their four archetypical mechanisms on the presumption of an open system: where people are not only concerned with their material well-being, but also the wish to become a contributing member of society (Jetzek et al. 2014). Therefore, transparency and collective impact are important value generating mechanisms, that contribute to the creation of beneficial social outcomes. We did not observe any socially driven motivation, in Trustpilot, a private company. We saw that they were concerned with optimising own utility, and less interested in social value. Thus, not finding value in transparency and collective impact as drivers. The paper from Jetzek et al. (2014) proposed a paradigm shift, where private organisations and governments work

together in a sharing society to create value at large. A paradigm that we did not readily observe at Trustpilot. Trustpilot was focused on what Jetzek et al. (2014) would propose as data-driven efficiency and data-driven innovation.

In this part of the discussion, the primary focus has been on the creation of value through open data. In the following, we look at how open data business models combine value creation and value capture.

Open Data Business Models

As earlier proposed opening data in a private company calls for the development of new business models due to value capture concerns. Are there any answers to be found in existing literature? We found that throughout literature many different open data business models have been examined. These offer different views on how to capture value through different business models. Janssen and Zuiderwijk (2014) examined open data business models, found in data infomediaries. These business models are made possible by the vast amounts of open data that has been published. Their identification of business models has contributed to the understanding of the *'middlemen'* effectively reselling data. These actors are not the data sources, giving them all together different incentives. A large part of the identified business models in literature, concern data infomediary business models, that enables end-customers to benefit from open data. Not much has been added to the discussion of how a private company, as a source of the data can extract or re-engineer their business models to capture value from open data. We are interested in adding to the discussion of how a business model can support value capture from open data. Ferro and Osella (2013) have proposed business models for PSI reuse, although still in a different domain, one of their business model archetypes is found interesting in the case of Trustpilot. That is, the proposed Free as Branded Advertisement business model (Ferro & Osella, 2013). It is an interesting business model, that could allow for true open data while still allowing the business to capture value. The business model implies that the data users are required to attribute and credit the original source (Ferro & Osella, 2013). This is still within the definition of open data suggesting that open data *"is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike."* (Open Knowledge Foundation, 2013). The Free as Branded Advertisement model could be used by Trustpilot as a strategic move, however, it would undermine their customers motivation to pay for premium subscriptions and would therefore require a substantial transformation of their business model. Such a model could improve channels, but hurt the revenue. Although the benefits are not readily measurable, we already saw that data sharing is used as a means of getting more exposure for Trustpilot.

Utilising this model could imply true open data from a business. It is necessary to further understand the business implications of such a business model, and how it can be successfully applied in a business. We do see these traits in their syndication partner program, where aggregator websites, such as price comparison sites are offered free data access.

Interesting business models are also proposed in the community by the Open Data Institute (2016b). They suggest cross-subsidization, that amongst others implies extension of brand reach through identifiers in data, or outsourcing of R&D through acquisition of companies making use of your data. They also suggest revenue can be generated by providing consultancy services, helping others make use of your data. This model shares traits with the Free as Branded Advertisement model proposed by Ferro and Osella (2013). Another model proposed is the freemium model. A model that implies, that businesses could give away subsets of data for free and charge for more detailed sets. An alternative freemium model is to charge data users for access through an API, instead of a data dump implying convenience in data access being the value added that the business charge for. Our results indicate that providing consultancy to data users could be a probable option. We learned that partners and customers are requesting support for the use of the API. Thus, it could be an option to offer free data access and instead charge for support, in using the data.

In the current reality of Trustpilot, they are not allocating a lot of resources to help API users. The reason being that the company is focused on growing their customer base. This suggests that, unless opening data is seen as a strategical activity and implemented with a clear framework of how to gain value from open data, it will be hard to embrace further openness, especially for a growth stage startup like Trustpilot (Respondent 1, personal interview, Feb. 2016; Crunchbase, 2016). We observe the Free as Branded Advertisement model, as a potential open data business model and put further emphasis on the need to examine the implications of this model in an empirical setting and also test how returns can be measured both intangible and tangible.

Ultimately providing open data requires a shift in how business models are designed.

Limitations

We acknowledge that a research study cannot be flawlessly carried out or comprise every aspect of an issue. In this section we present and discuss the limitations of this study to highlight potential shortcomings and unaddressed aspects.

Limitations of Analytical Frame

The analytical framework applied both guided and limited this study. Even though the analytical framework evolved throughout the study, it prescribed a certain direction of the study. Our analytical framework contributed to the design of interview guides, and thus, influenced the questions asked and answers provided by respondents. Had we applied an alternative analytic framework, we might have received different answers to differently formulated questions. Additionally the analytical framework impacted our approach to analysing empirical evidence, since our primary focus was on finding answers that matched the frame we had in mind.

Our analytical frameworks included the data spectrum, business model theory and platform theory.

The data spectrum proved particularly useful, although it is a novel framework emerged from practice. The data spectrum offer a vocabulary that refines data sharing beyond open vs. closed data. An exact position in the shared data range is still difficult to determine. The shades indicates that a clear cut between data access categories cannot be made. These shades could either be an expression of an immature framework or a degree of complexity which is difficult to cope with in a simple framework.

Our application of business models theory was limited to the interpretation offered by Osterwalder and Pigneur (2010). The conceptual model of a business model canvas offered a versatile analysis and insight into how Trustpilot creates, delivers and captures value. The role of data sharing was a dimension included by the authors. However, it is known that the business model is not designed to be applied in isolation. Osterwalder and Pigneur (2010) complements the framework with traditional strategic frameworks themselves to extend its usefulness. A noteworthy limitation of the business model framework, we experienced, is its lack of focus on the external environment, market and industry. Aspects that are deliberately beyond the scope of the framework.

In coping with the strategic shortcomings of the business model canvas we found platform theory suitable. Addressing Trustpilot's rationale through platform theory offered insights that complemented the findings from the business model canvas. It is questionable if platform theory is likely to be just as valuable in studies of other companies. It is intended primarily towards industries greatly influenced by technological evolution and businesses that address multi-sided markets. By applying industry platform theory, our research was limited to understanding the dynamics that affects Trustpilot, and thus, any theoretical suggestions for

how to succeed as a platform in a business ecosystem was omitted (Gawer & Cusumano, 2008). Since the purpose of this study was not to offer guidelines for businesses in successful data sharing, but rather identifying a rationale for actions. This delimitation was a conscious decision.

Overall the analytical framework proved useful in understanding the empirical setting, however, we see opportunities in understanding businesses rationale for data sharing through several other theoretical concepts and frameworks. Such frameworks include resource-based view (Wernerfelt, 1984), dynamic capabilities (Teece et al., 1997), shaping strategy (Hagel et al., 2008) and keystones in ecosystems (Iansiti & Levien, 2004a).

Methodological Limitations

The findings in this study offers a still picture of the situation in Trustpilot. The evolution of their data sharing initiative and rationales behind could develop over time, and thus, important findings might not be revealed in this study.

The case study on Trustpilot was not carried out without challenges, some of which limited the data collection process. The initial strategy for establishing contact to respondents was snowballing, however, we quickly faced dead-ends, where the respondents were unable to refer new subjects. From what we heard from respondents, very few people in Trustpilot were actually involved with and knowledgeable about their initiatives to share data through APIs. Contact was established with further respondents beyond the snowball. Their knowledge was not without relevance, but as expected very limited in regards to APIs and exposing data externally. Our empirical evidence was informed by different departments, but lacked insights from the top executives and the legal department. Although contact was attempted, they were unavailable for the study. We expect that top management could have offered additional inputs in regards to Trustpilot's strategic perspective on sharing data and that their legal staff could offer insights into legal concerns and efforts in sharing data.

Limitations of Scope

This study took a deep dive into a business to understand the underlying rationale for sharing data and not opening data.

The design of this study was an exploratory case study, and thus, leaves many questions open. The rationales identified could be specific to the case at hand, and thus, we cannot give

definitive answers as to how other businesses justify data sharing. Our emphasis with this case study has been on providing the reader with sufficient detail and insight, to make an individual assessment of whether it is applicable to their context or not.

Our understanding of Trustpilot's data sharing was formed through their internal view on the matter. Therefore external interpretations from customers, partners and other potential data users were not addressed. Since the focus of this study was on Trustpilot's rationale, we focused on their interpretations of their initiatives. Although, it would be interesting to see the issue from data the users perspective, it was not within the scope of this study.

Another limitation of this study was the focus on certain datasets. In this study we focused on data central to the Trustpilot offering and not data which concerns the inner workings of Trustpilot. An example of such data could be financial data about Trustpilot expenses. Such information is considered confidential and inappropriate for publishing. However, as suggested by Manyika et al. (2013) companies can benefit greatly by anonymously sharing such data to benchmark performance. In this study we decided to focus on the datasets which Trustpilot have already included in their APIs and thus shown willingness to share in a loosely coupled manner.

Since Trustpilot does not offer true open data, our research was limited to understand external data sharing. Open data was not part of their current reality, and, thus our findings are not based on a case study of a company with true open data.

Future Research

We brought forth the idea that there is a gap in the open data pool, restricting the full value extraction from data. We have taken early steps in the discussion of businesses as providers of open data - what we refer to as open company data. This brings forth a wide array of implications that requires further research to be fully understood. Through our engagement in open company data research, we faced several unaddressed questions, that we contribute as input to other researchers with interest in the field. Below are issues and questions that require further attention:

How can a business open data without undermining their value capture mechanisms? Business models need a significant revisit to accommodate the phenomena of open data. We

suggest a better understanding is need of how data can be offered by business, under an open license, while allowing for value capture through other means than sales of data access.

Which datasets in a business are most appropriate for an open license? This study has emphasised an understanding of underlying business rationale for sharing data. We did not distinguish thoroughly between which types of data sets might be more suitable for an open license. We suggest that additional research is done to address, which types of data sets could be more appropriate for sharing with unrestricted access and user rights - according to the open data definition.

What characterises a business that shares data externally? During our case selection we observed that businesses with digital products seemed to be better represented in those identified that share data externally. We suggest it is necessary to understand what characterises businesses that are more likely to share data externally. Such understanding could lead to a more nuanced debate and help distinguish between companies and examine the differences in motivation by business characteristics. We suggest that data sharing initiatives by businesses can be categorised in the data spectrum. Such a study could help to further mature the data spectrum framework.

What characterises datasets business shares externally? Data in business has very different origin, purpose and affinity. In this study the focus was on data already exposed by the business, however, a thorough understanding of what makes such datasets particularly suitable for sharing is needed. Such an examination could suggest which datasets are likely to be shared or opened in the future.

Besides these unmet issues, that we have encountered during this study, we propose that other studies of businesses that share data externally are carried out to further examine how our proposed justifications apply in those contexts.

Conclusion

This study set out to examine and understand businesses' attitude towards external data sharing, as well as to open the discussion of businesses as a source of open data. Through our research, we found that external data sharing in business does not align with the open data definition. The case study on Trustpilot provided four justifications for external data sharing in business and brought forth potential hindrances for what we term open company data. External data sharing is justified in business by the opportunity to 1) increase customer utility, 2) extend brand exposure, 3) amplify platform dynamics and 4) leverage complementors in a business ecosystem, while further openness is inhibited by 1) undermined value capture, 2) loss of control, 3) adverse brand effects and 4) the need for a proactive approach. We convert these findings into eight propositions for future researchers to examine in other business contexts.

Through a review of the open data research domain, we found that the majority of research on open data has been focused on open government data (OGD). We recognise the immense potential of open government data. However, we find it problematic to view open data as synonymous with open government data (Heimstädt et al., 2014) and, therefore, we suggest exploring companies as open data providers. We see a missing piece in open data research - open company data.

To examine this missing piece, we posed the research question: How is external data sharing justified in business?

By asking this question, an in-depth understanding of the state of external data sharing in business and the motivation for a business to share data has been established. To approach the extreme condition of open data, it is necessary to understand the underlying business justifications for why a company would contemplate data sharing with external parties. To examine this question empirically, we conducted an exploratory case study on Trustpilot. Empirical evidence was obtained through semi-structured interviews with Trustpilot employees, as well as, articles and documentation from Trustpilot's websites.

To better understand Trustpilot's business and their approach to data sharing, we addressed the following questions:

- How does business offer access to data and to whom?
- What is the role of external data sharing in the business model?
- How does external data sharing affect the business in a business ecosystem?
- What are the hindrances for business to open data?
- How can business move towards open data?

The analysis on how business offer access to data resulted in an assessment of the openness of data. Additionally, the data sharing initiative was positioned on the data spectrum. We found that Trustpilot does not offer open data, but shared data. Data is shared through APIs with premium customers, technology partners and syndication partners. Anyone wishing to access Trustpilot's data must apply for a partnership or pay for a premium customer account. We learned from Trustpilot, that restricting data access is a way for business to stay in control of how their data is used outside of the company.

Through use of Osterwalder and Pigneur's (2010) business model framework, we found that data sharing has a significant impact on Trustpilot's ability to create, deliver and capture value for customers. It was found that data sharing enables Trustpilot to offer customers better utilisation of their product and, thus, a stronger value proposition. The increased value creation also results in increased value capture, as access to data implies a price premium in the customer's monthly subscription.

Proposition 1: *External data sharing increases customer utility.*

Another positive outcome of external data sharing is extended brand exposure. Business customers, technology partners and syndication partners use Trustpilot's data to display the Trustpilot brand, TrustScores and reviews across websites, apps and integrations. The additional brand exposure creates channels, that help Trustpilot reach more customers.

Proposition 2: *External data sharing extends brand reach.*

By application of a platform theoretical frame, we established that Trustpilot is a multi-sided platform (Hagiu & Wright, 2015; Shapiro & Varian, 1998; Osterwalder & Pigneur, 2010). It is found that data sharing is an amplifier of platform dynamics. Customer integration utilisation of Trustpilot data implies that they are automatically sending review invitations to their customers. Automated invitations result in a greater review base and more Trustpilot users, further contributing to positive direct network effects, which increase the value of Trustpilot

for other consumers. Data sharing also increases financial and procedural switching costs for Trustpilot customers. Ultimately, the amplified platform dynamics contribute to a positive feedback loop, where Trustpilot is strengthened through intensified network effects and increased switching costs, heightening the barriers to leave Trustpilot.

Proposition 3: *External data sharing amplifies platform dynamics*

Trustpilot find themselves in a business ecosystem, where other software companies serve some of the same customers as they do. Trustpilot's customers often use more than one software tool to connect with their end-customers. Trustpilot's customers value integration between these systems. Through data sharing, customers can integrate Trustpilot data into other systems as desired. However, such integrations are costly to develop. Therefore, off-the-shelf integrations between the systems are valuable. Trustpilot views data sharing with technology partners, as an opportunity to leverage, what Cusumano and Gawer (2008) refer to as, complementors. Through data sharing with technology partners, Trustpilot attempts to establish an industry platform (Gawer, 2010), where others build innovative apps and integrations through the use of Trustpilot's review database.

Proposition 4: *External data sharing enables leverage of complementors*

Lastly, we discussed the hindrances for businesses in publishing open data. Through our examination of Trustpilot, we encountered objections towards further openness of data, indicating that the following hindrances inhibit open company data. Trustpilot is able to promote their top-tier subscriptions, by including programmatic data access. Consequently, a state where Trustpilot publishes open data for free and unrestricted access would imply undermining their existing value capture mechanism.

Proposition 5: *Sharing data under an open license undermines value capture*

Our study also indicates that Trustpilot fears a loss of control over their data. The loss of control entails that external parties are empowered to create competing products based on Trustpilot's data, which may impact Trustpilot's revenue streams significantly.

Proposition 6: *Sharing data under an open license entails unpredictable outcomes*

Another inhibitor of open data is the concern of data users misusing data. Misuse could include misleading end-consumer, through altered TrustScores or misrepresentation of reviews. Additionally, poor integrations of Trustpilot data and assets by customers and partners have the potential to damage Trustpilot's brand.

Proposition 7: *Sharing data under an open license entails risk of adverse effects on brand*

It is found that Trustpilot finds it unlikely that open data itself will result in valuable use of their data, which suggests that large investments are needed in proactively promoting and encouraging use of data.

Proposition 8: *Sharing data under an open license requires a proactive approach to achieve valuable use of data*

Our findings suggests that open data is not a straightforward decision. It entails unforeseeable impact on the current business model. We suggest that the business model and strategy need to accommodate the substantial impact from open data. Opening data has the potential to invalidate an existing business model, and new ways of generating revenue are needed.

We opened the discussion on whether we will find businesses move towards open data in the future. We discussed the relevance of our findings in the context of the open data research domain. The governmental incentives for open data were considered compared to private incentives. We noted that a significant differences between these entities exist. The premise on which governments opens data differ, in that governments capture a fraction of any value created in society through taxes, whereas businesses do not have such a measure in place. We discussed open data business models from literature against our findings, to uncover if there were new insights to be gleaned. Open data business models found in the literature were predominantly based on infomediaries basing their business on open government data. We highlight that business models, such as the Free for Branded Advertisement and Premium model, may have potential in the context of open company data. We observed from our study that the Free for Branded Advertisement model shared traits with what Trustpilot is currently doing and that the premium model, where data users for instance pay for support has interesting elements to it.

We have also put forth whether private investors' short term financial thinking may hinder the advent of business as open data providers. Based on the argument that some of the greatest innovations in society were not driven by value capture, suggesting that value should be seen over the long term, and capture should be secondary to value creation. Also, we opened the discussion of whether today's tangible nature in measuring returns also hinders open data, due to its returns not being readily measurable and capturable.

Final Remarks

This exploratory study of external data sharing in business, suggests that businesses do not find it necessary to apply open licenses to foster value creation in their business ecosystem. Businesses find opportunities to capture value created through the use of their data. Value capture is enabled by an artificial scarcity of data, simulated through limited access. Instead of offering data widely and freely, a business like Trustpilot selectively assesses the possibilities of collaboration with individual data users. The exploratory nature of this case study means that our findings are not conclusive. We suggest that more open data research is conducted, within business to further advance knowledge of how companies can become contributors to the cumulative pool of open data and advance society at large.

References

Al-Debei, M. M., El-Haddadeh, R., & Avison, D. (2008). Defining the business model in the new world of digital business. In Proceedings of the Americas Conference on Information Systems (Vol. 2008, pp. 1-11).

Anney, V. N. (2014). Ensuring the quality of the findings of qualitative research: looking at Trustworthiness Criteria. Journal of Emerging Trends in Educational Research and Policy Studies, 5(2), 272-281.

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. Government Information Quarterly, 32(4), 399-418.

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007). Dbpedia: A nucleus for a web of open data (pp. 722-735). Springer Berlin Heidelberg.

Baldwin, C. Y., & Woodard, C. J. (2009). The architecture of platforms: A unified view. Platforms, markets and innovation. 2009, pp. 19-44 (pp. 19-44)

Bauer, F., & Kaltenböck, M. (2011). Linked open data: The essentials. Edition mono/monochrom, Vienna.

Berners-Lee, T. (2006). Linked data. Retrieved from http://www.w3.org/DesignIssues/LinkedData.html

Berners-Lee, T. (2010). Linked data > is your linked open data 5 star? Retrieved from http://www.w3.org/DesignIssues/LinkedData.html

Bizer, C., Heath, T., & Berners-Lee, T. (2009a). Linked data - the story so far. International Journal on Semantic Web and Information Systems, 5(3), 1-22.

Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., & Hellmann, S. (2009b). DBpedia - A crystallization point for the web of data. Web Semantics: Science, Services and Agents on the World Wide Web, 7(3), 154-165.

Blaikie, N. (2007). Approaches to social enquiry: Advancing knowledge. Polity Press.

Bonina, C. (2013). New business models and the value of open data: Definitions, challenges and opportunities. London School of Economics and Political Science, Department of Management.

Bowman, C., & Ambrosini, V. (2000). Value creation versus value capture: towards a coherent definition of value in strategy. British Journal of Management, 11(1), 1-15.

Brandenburger, A. M., & Stuart, H. W. (1996). Value-based business strategy. Journal of Economics & Management Strategy, 5(1), 5-24.

Broad, E. (2015). Closed, shared, open data: What's in a name? Retrieved from https://theodi.org/blog/closed-shared-open-data-whats-in-a-name

Burnham, T. A., Frels, J. K., & Mahajan, V. (2003). Consumer switching costs: A typology, antecedents, and consequences. Journal of the Academy of Marketing Science, 31(2), 109-126.

Clausen, M. (2016). Ekspert: Trustpilot vildleder brugerne. Retrieved from http://www.dr.dk/nyheder/penge/kontant/ekspert-trustpilot-vildleder-brugerne

Coes, B. (2014). Critically assessing the strengths and limitations of the business model canvas. University of Twente

Creswell, J. W. (2008). Research design (3rd edition)

Crunchbase. (2016). Trustpilot | CrunchBase. Retrieved from https://www.crunchbase.com/organization/trustpilot

Cusumano, M. A. (2010). Staying power: Six enduring principles for managing strategy and innovation in an uncertain world. Oxford University Press.

Cyganiak, R. & Jentzsch, A. (2014). The Linking Open Data cloud diagram. Retrieved from http://lod-cloud.net/

data.gov. (2016). The home of the U.S. government's open data. Retrieved from https://www.data.gov/

Davies, T. G., & Bawa, Z. A. (2012). The promises and perils of open government data (OGD). The Journal of Community Informatics, 8(2).

DBpedia community (2014). Facts & figures. Retrieved from http://wiki.dbpedia.org/about/about-dbpedia/facts-figures

Dubois, A., & Gadde, L. (2002). Systematic combining: An abductive approach to case research. Journal of Business Research, 55(7), 553-560. doi:10.1016/S0148-2963(00)00195-8

Facebook Inc. (2015). Graph API reference. Retrieved from https://developers.facebook.com/docs/graph-api/reference

Ferro, E., & Osella, M. (2013). Eight business model archetypes for PSI re-use. Open Data on the Web Workshop, Google Campus, Shoreditch, London,

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. Qualitative Inquiry, 12(2), 219-245.

Fox, R. (1998). Layered abduction and abductive inference. Computer Science On-Line,

Gawer, A. (2010). Towards a general theory of technological platforms. Imperial College London Business School: Druid.

Gawer, A., & Cusumano, M. A. (2002). Platform leadership. Boston, Massachusetts: Harvard Business School Press.

Gawer, A., & Cusumano, M. A. (2008). How companies become platform leaders. MIT Sloan Management Review, 49(2), 28-35.

Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. Journal of Product Innovation Management, 31(3), 417-433.

Google Inc. (2016a). Google APIs explorer. Retrieved from https://developers.google.com/apis-explorer/

Google Inc. (2016b). Understanding your seller ratings. Retrieved from https://support.google.com/adwords/answer/2375474

Guba, E., & Lincoln, Y. (1994). Competing paradigms in qualitative research

Hagel, J., Brown, J. S., & Davison, L. (2008). Shaping strategy in a world of constant disruption. Harvard Business Review, 86, 80-89.

Hagiu, A. (2006). Multi-sided platforms from microfoundations to design and expansion strategies.(115)

Hagiu, A., & Wright, J. (2015). Multi-sided platforms. International Journal of Industrial Organization, 43, 162-174.

Hausenblas, M. (2012). 5 ★ open data. Retrieved from http://5stardata.info/en/

Heimstädt, M., Saunderson, F., & Heath, T. (2014). Conceptualizing open data ecosystems: A timeline analysis of open data development in the UK. Conference for E-Democracy and Open Governement, 245.

Hillier, D., Clacher, I., Ross, S., Westerfield, R., & Jordan, B. (2011). Fundamentals of corporate finance. McGraw-Hill Education.

Holley, K., Antoun, S., & Arsanjani, A. (2014). The power of the API economy. IBM Corp.

Huijboom, N., & Van den Broek, T. (2011). Open data: an international comparison of strategies. European journal of ePractice, 12(1), 1-13.

Iansiti, M., & Levien, R. (2004a). Keystones and Dominators: Framing Operating and Technology Strategy in a Business Ecosystem. Harvard Business School.

Iansiti, M., & Levien, R. (2004b). Strategy as ecology. Harvard Business Review, 82(3), 68-81.

IDC Research. (2014). Data growth, business opportunities, and the IT imperatives. Retrieved from http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. Information Systems Management, 29(4), 258.

Janssen, M., & Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. Social Science Computer Review, 32(5), 694-711.

Jensen, J. (2014). Online reviews even the bad ones are good for brands. Retrieved from http://www.cmo.com/articles/2014/5/21/online_reviews_even_.html

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014). Generating sustainable value from open data in a sharing society. IFIP Advances in Information and Communication Technology, 429, 66-82.

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013). Generating value from open government data. The 34th International Conference on Information Systems. ICIS 2013

Jetzek, T. (2015). The Sustainable Value of Open Government Data: Uncovering the Generative Mechanisms of Open Data through a Mixed Methods Approach. Copenhagen Business School, Department of IT Management.

Johnson, R. B. (1997, December 22). Examining the validity structure of qualitative research. Education, 118, 282.

Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures and their consequences. Sage.

Kvale, S. (1996). Methods of analysis. Interviews: An introduction to qualitative research interviewing (pp. 187-191)

Kvale, S. (2007). Doing interviews (1. publ. ed.). London: SAGE.

Kraaijenbrink, J. (2012). Three shortcomings of the business model canvas. Retrieved from http://kraaijenbrink.com/2012/07/shortcomings-of-the-business-model-canvas/

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE Publications.

LOD Community. (2015). SweoIG/TaskForces/CommunityProjects/LinkingOpenData. Retrieved from http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Doshi, E. A. (2013). Open data: Unlocking innovation and performance with liquid information, McKinsey Global Institute.

Marton, A. (2013). Purposive selection and the quality of qualitative IS research.

Oates, B. (2006). Case studies. Researching information systems and computing (pp. 140-153) Sage Publications Ltd.

Open Data Institute. (2016a). Data spectrum. Retrieved from http://theodi.org/data-spectrum

Open Data Institute. (2016b). How to make a business case for open data. Retrieved from https://theodi.org/guides/how-make-business-case-open-data

Open Knowledge Foundation. (2013). What is open data? Retrieved from http://opendatahandbook.org/guide/en/what-is-open-data/

Open Knowledge Foundation. (2014). Open knowledge: What is open? Retrieved from https://okfn.org/opendata/

Open Knowledge Foundation. (2015). Global open data index. Retrieved from http://index.okfn.org/

OpenDefinition.org. (2015). Open definition 2.1. Retrieved from http://opendefinition.org/od/2.1/en/

O'Reilly, T. (2012). Value creation vs. value capture: Musings on the new economy. Retrieved from https://www.linkedin.com/pulse/20121112204533-16553-value-creation-vs-value-capture-musings-on-the-new-economy

Osterwalder, A., & Pigneur, Y. (2010). Business model generation A handbook for visionaries, game changers, and challengers (1st ed.). US: John Wiley & Sons Inc.

Pagani, M. (2013). Digital Business Strategy and Value Creation: Framing the Dynamic Cycle of Control Points. MIS Quarterly, 37(2), 617-632.

Pine, B. J. (1993). Mass customization: The new frontier in business competition Harvard Business School Press.

Seale, C. (1999). Quality in qualitative research. Qualitative Inquiry, 5(4), 465-478.

Shapiro, C., & Varian, H. R. (1998). Information rules. Boston, Mass: Harvard Business School.

Shenton, A. (2004). Strategies for ensuring trustworthiness in qualitative research projects. Education for Information, 22, 63-75.

Surowiecki, J. (2005). The wisdom of crowds. Anchor.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal, 18(7), 509-533.

Twitch Interactive Inc. (2015). Twitch API v3. Retrieved from https://github.com/justintv/Twitch-API

Twitter Inc. (2016). REST APIs | twitter developers. Retrieved from https://dev.twitter.com/rest/public

Ubaldi, B. (2013). Open government data: Towards empirical analysis of open government data initiatives. Social & Environmental Accounting Journal, 31(1), 25-47.

Vogt, W. P. (1999). Dictionary of statistics and methodology: A nontechnical guide for the social sciences (2nd edition). Thousand Oaks: Sage.

Walmart Stores Inc. (2016). Walmart open API - welcome to the walmart developer network! Retrieved from https://developer.walmartlabs.com/

Wernerfelt, B. (1984). A resource-based view of the firm. Strategic Management Journal, 5(2), 171-180.

Willmott, S., Balas, G., & 3scale. (2013). Winning in the API economy

World Wide Web Foundation. (2015). Open data barometer - third edition.Open Data Barometer.

Yin, R. K. (2009). Case study research (4th edition)

Zeleti, F., Ojo, A., & Curry, E. (2014). Emerging business models for the open data industry. 215-226.

Zott, C., Amit, R., & Massa, L. (2011). The business model: Recent developments and future research. Journal of Management, 37(4), 1019-1042.

Zucker, D. M. (2009). How to do case study research

Case references:

Trustpilot Support Center (2016a). Trustpilot API and feeds, service review integration guidelines., 2016, from https://support.trustpilot.com/hc/en-us/articles/202319278-Trustpilot-API-and-Feeds-Service-Review-Integration-Guidelines

Trustpilot Support Center. (2016b). TrustScore explained. Retrieved from https://support.trustpilot.com/hc/en-us/articles/201748946-TrustScore-Explained

Trustpilot.com (2015). Trustpilot: The story so far... Retrieved from http://blog.trustpilot.com/blog/2015/6/1/trustpilot-the-story-so-far

Trustpilot.com. (2016a). About review. Retrieved from http://company.trustpilot.com/about-reviews

Trustpilot.com. (2016b). About trustpilot. Retrieved from http://company.trustpilot.com/about

Trustpilot.com. (2016c). The best companies in trustpilot's business categories. Retrieved from https://www.trustpilot.com/categories

Trustpilot.com. (2016d). Trustpilot API documentation. Retrieved from https://developers.trustpilot.com/

Trustpilot.com. (2016e). Trustpilot partnerships. Retrieved from http://apps.trustpilot.com/partners

Trustpilot.com. (2016f). Trustpilot plans. Retrieved from https://business.trustpilot.com/plans-pricing

Trustpilot.com. (2016g). The voice of the customer. Retrieved from https://business.trustpilot.com/

Trustpilot.com. (2016h). Trustpilot Apps. Retrieved from http://apps.trustpilot.com/apps

About the Authors

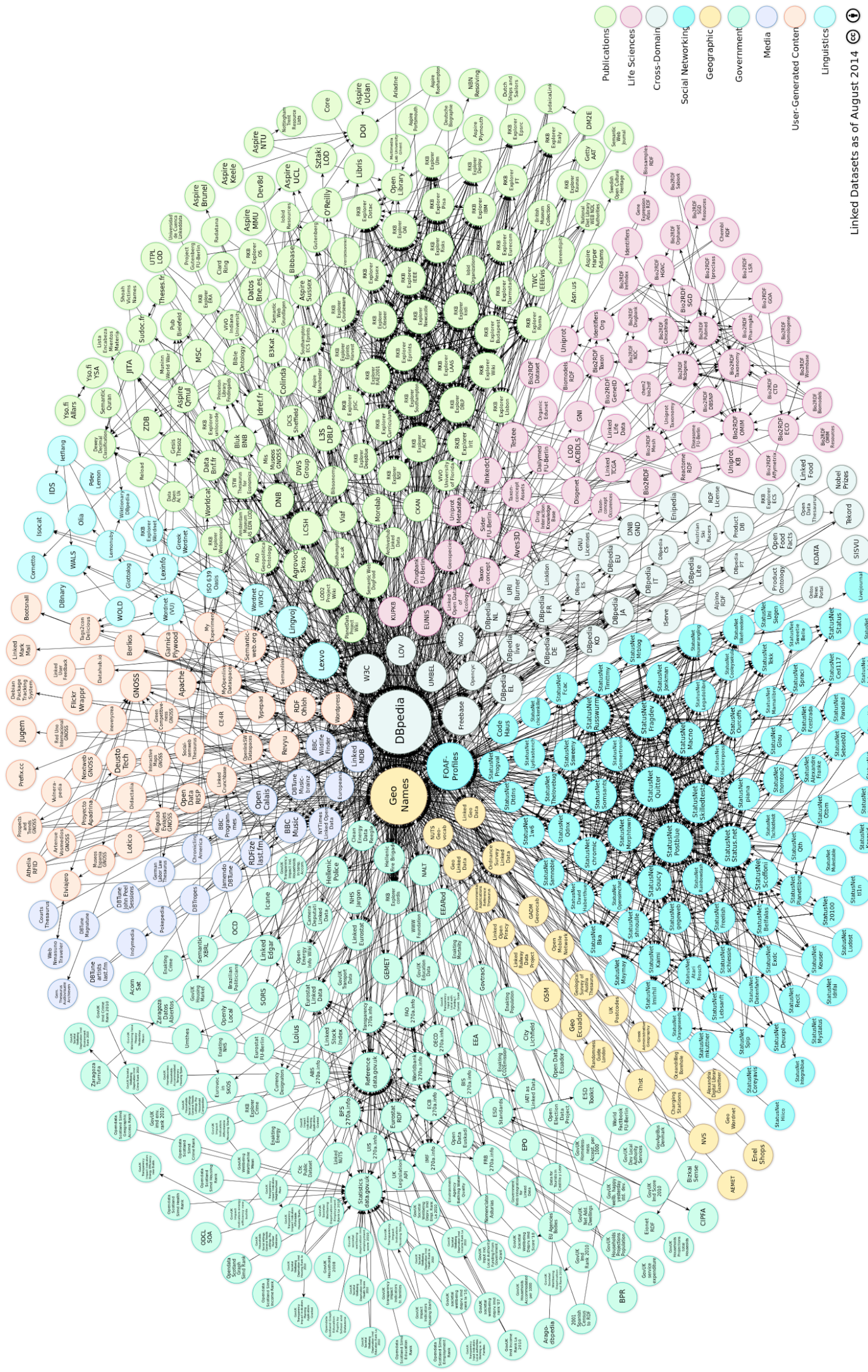
Reshal Katyal a digital entrepreneur and MSc. IT-Management and Business Economics student from Copenhagen Business School.
He has worked within the digital sphere for several years and has a strong interest in digital innovation and how data-driven innovation can transform society.

Bjarke Due Jensen, a MSc. IT-Management and Business Economics student and Business Intelligence Consultant.
Has professional experience from the entertainment and IT industry, with a primary focus on web technology, business information systems, business intelligence and data analytics.

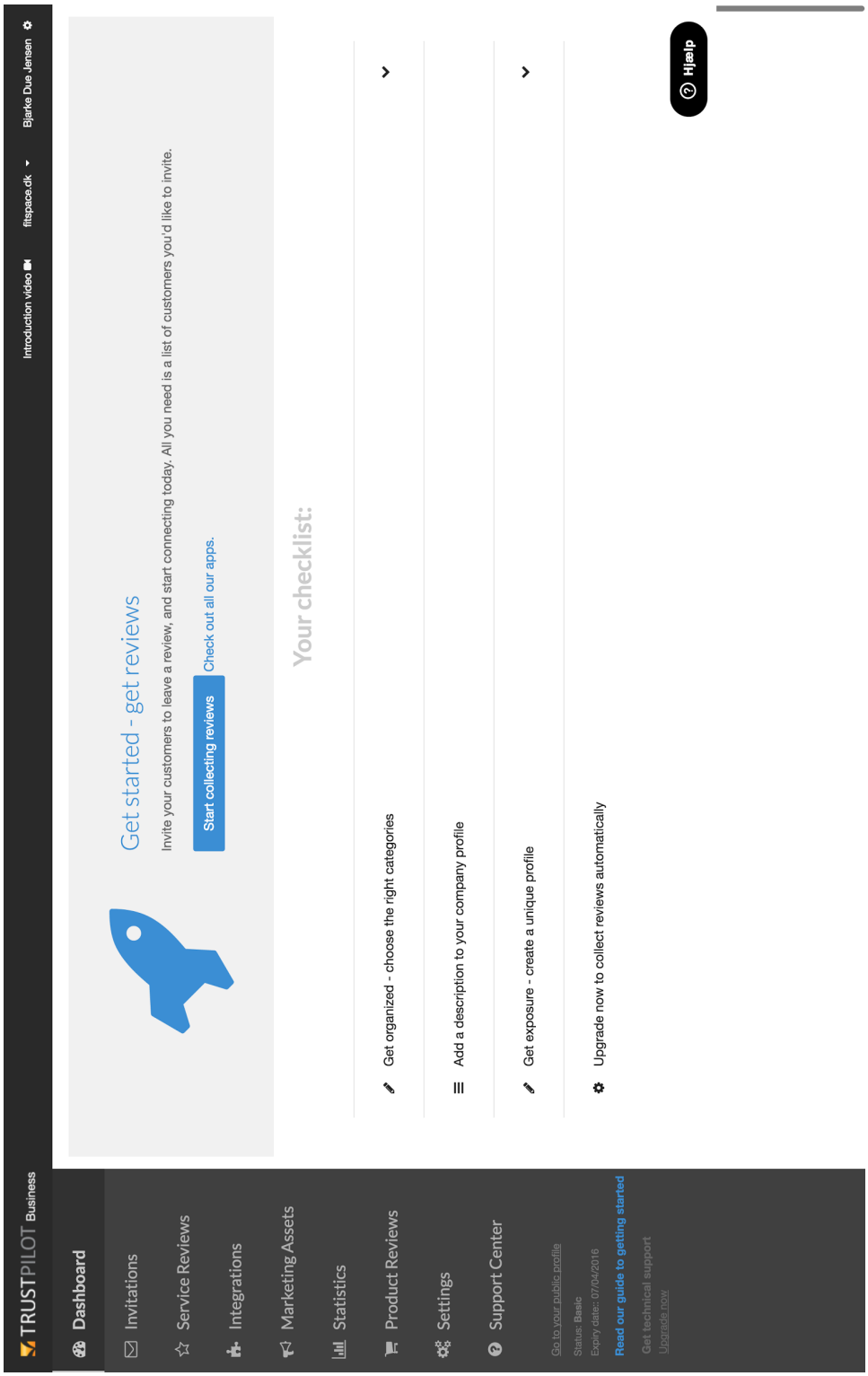
Appendix

List of Appendix

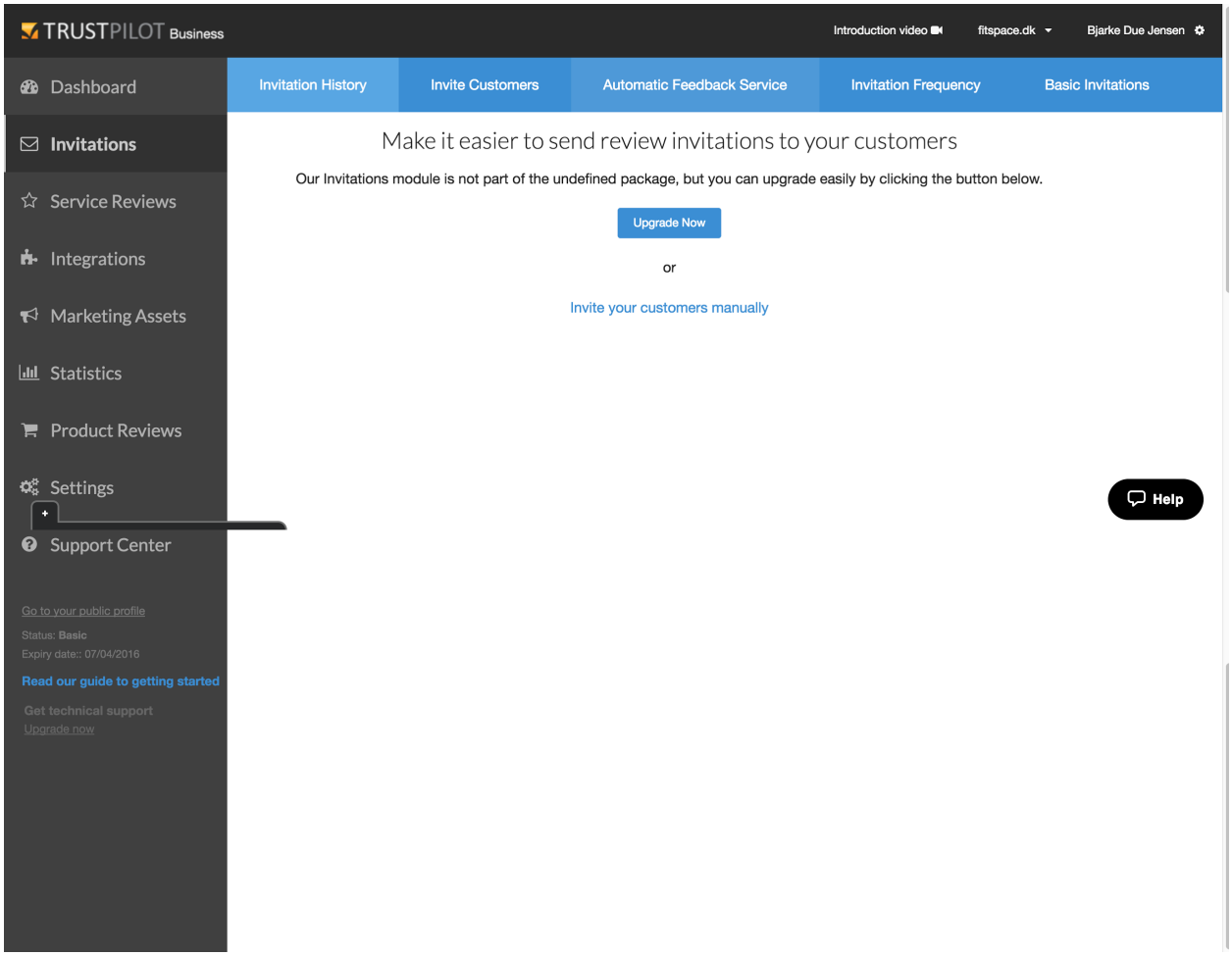
Appendix 1: LOD Cloud Diagram



Appendix 2: Trustpilot Business App Dashboard



Appendix 3: Trustpilot Business App Invitations



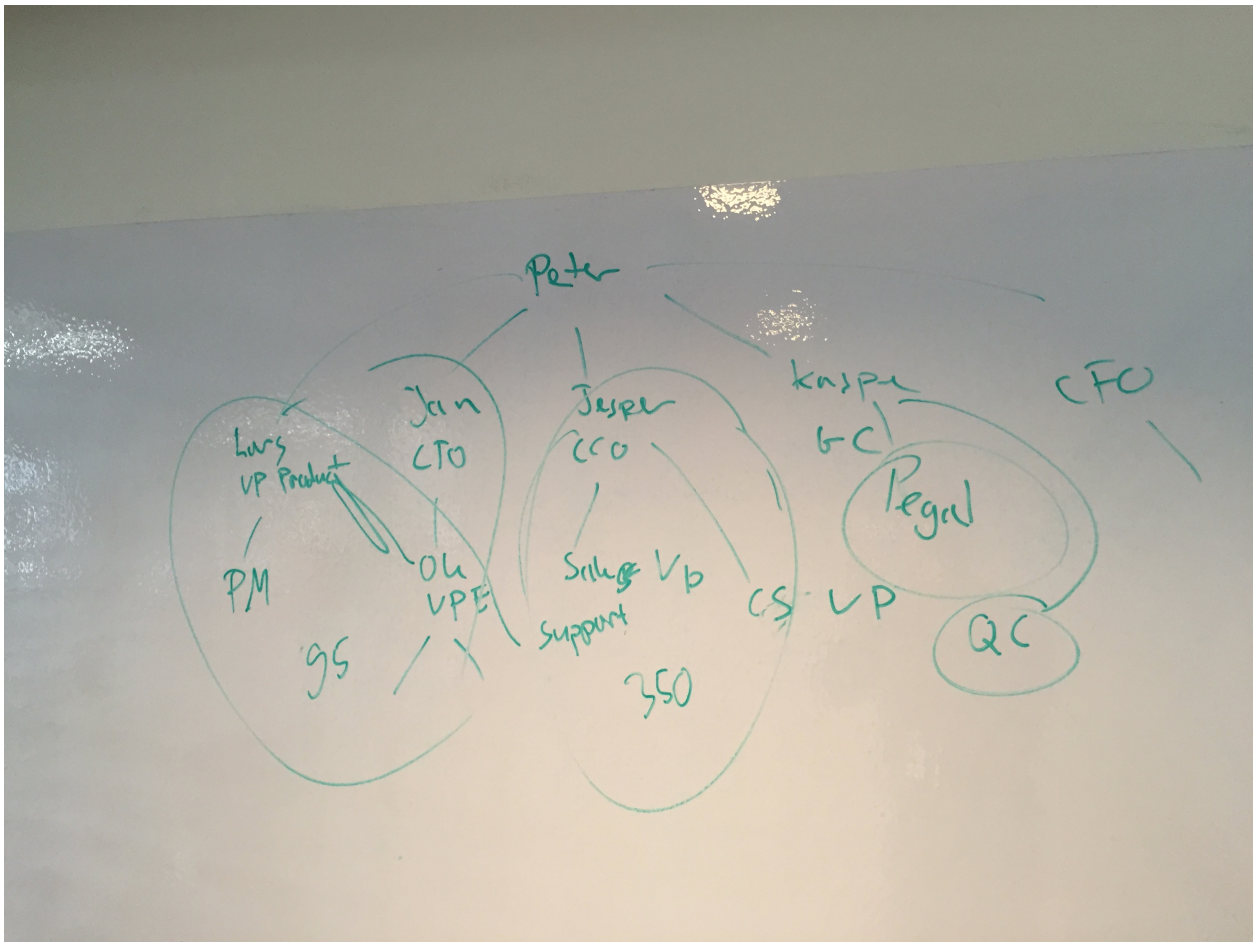
Appendix 4: Overview of Accessible Data in Trustpilot API

	"Public" Accessible Data	Customer Accessible Data
Business Unit	All Business Units including the attributes: - ID - Display Name - Website URL - Alternative Website URLs - TrustScore - Stars - Country - Number of reviews (count for each 1-5 star rating) - Reviews for the Business Unit - Profile URL - Evaluate URL - Embedded Evaluation URl - Categories the Business Unit is in.	Additional information about Business Unit: - Private Reviews - Private Tags
Categories	All Categories including the attributes: - ID - Display Name - Parent Category - Number of Business Units in Categories - Business Units in Category	
Consumer	All Consumers including the attributes: - ID - Profile URL - Display Name - City - Country - Gender - Profile Created At - About Text - Birth Year - Language - Profile Image - Number of reviews - Reviews written by the consumer	Additional Information About Private Consumers: - Email
Invitation		All invitation templates including the attributes: - ID - Name

Business Unit Reviews	Reviews - List of latest on Trustpilot Review - ID - Reviewing Consumer - Reviewed Business Unit - Stars - Title - Text - Language - Created at - Updated at - Company Reply (Text and date) - Tags (Group, Value) - Number of Likes - Likes by consumers - Is the review verified? - Web link	Additional Information About Reviews: - Referral email - Reference ID - Source - Status - Report Data
Product Review	Product reviews (and imported product reviews) - Number of Reviews (Count for each 1-5 star rating) - Stars average Product review (and imported product reviews) - Reviewing Consumer - Review Text - Stars - Language - Created At	Additional Information for Private Product Reviews: - SKU Additional Information for Private Product Review: - Product URL - Product Images - Product Name - Product SKU - Product GTIN - Product mpn - Product brand - Updated At
Resources	Links to Image Resources - Trustpilot Stars - Trustpilot Logo Supported Locales - Locale - Language - Region - Native Name - Translated Name - TLD - Link to Consumer Web App - Link to Business Web App Countries - Country Code - English Name - Translated Name - Locale Stars' verbal representations - Locale - Stars - String	

Appendix 5: Sketch of Trustpilot Organization

This Organization Diagram was drawn by Respondent 1 after the interview, February 2016.



Appendix 6: Trustpilot Business Model Canvas

Appendix 7: Guide to Interview guides, Audio and Transcripts

For the sake the sake of the trees we put our interview guides, Audio and Transcripts etc. online in this Dropbox folder:
https://www.dropbox.com/sh/9syf90wku9j8a60/AAAde2jEzGatPcQUqMnerx2ea

Also available through this short-link: https://goo.gl/8rI2zb

Overview of Interviews:

#	Referred to as	Role	Rank	Date	Duration (h:mm)
1	Respondent 1	Engineering	Very high	24th February, 2016	1:44
2	Respondent 2	Marketing	High	11th March, 2016	1:10
3	Respondent 3	Strategy	Very high	29th March, 2016	1:18
4	Respondent 4	Market and Sales	Medium	18th April, 2016	0:43
5	Respondent 5	Technical Specialist	Medium	18th April, 2016	0:31
6	Respondent 6	Integrations	Medium	20th April, 2016	0:39

Appendix 8: Codes

activedatasharing
adwords
apidocumentation
apimotivation
apisignup
appstore
authentication
bm/ch
bm/cost
bm/cr
bm/cs
bm/data
bm/ka
bm/kp
bm/kr
bm/rs
bm/vp
businesscase
businesses/amazon
businesses/apigee
businesses/bazaarvoice
businesses/bigcommerce
businesses/bing
businesses/bitcommerce
businesses/bizrate
businesses/boligportalen
businesses/comi
businesses/dandomain
businesses/falconsocial
businesses/fifo
businesses/gnatta
businesses/golfexperten
businesses/google
businesses/hootsuite
businesses/ikomi
businesses/linkedin
businesses/magento
businesses/mailchimp
businesses/momondo
businesses/navipartner
businesses/netflix
businesses/pricerunner
businesses/rejsegiganten
businesses/reviewuk
businesses/salesforce
businesses/saxo
businesses/sendgrid
businesses/shopify
businesses/slack
businesses/tripadvisor
businesses/twilio
businesses/twitter
businesses/woocommerce
businesses/yelp
businesses/yotpo
businesses/zapier
businesses/zendesk
businessprocess
casestudies
co-marketing
competitors
compliance
coopetition
copycat
cost-saving
customer-api
customer cross-over
customersuccess
datacollection
datawrites
dikw
economiesofscale
empowerment
fraughtdetection
ideas
innovation
intangible
integration
integration/customer
integration/pos
integrity
internalplatform
internalsystems
internalvsexternal
introduction
introduction/
employeebackground
introduction/thesispurpose
lead generation
lockin
marketing/brand
marketing/customerbehaviour
marketing/
customerexperience
marketing/integrations
media/dr
multi-sided market
mutualbenefit
networkeffects
nicheneeds
od
od/attribution
od/ch
od/dataaccess
od/de
od/motivation
od/mydata
od/op

od/tos
od/transparency
od/vp
organization
partnerprogram
partnerprogramstatus
partners
partnershipframework
partnersignup
partnerstrategy
passivedatasharing
pf
pf/2sided
pf/ec
pf/internalplatform
pf/ks
pf.ks
pfl
pf/l1/sf
pf/l2/pt
pf/l3/cr
pf/l4/io
pf/lf1/sf
pf/lf4/io
pf/ls
pf/ls1/pf
pf/ls1/sf
pf/ls3/cr
pf/ls4/io
pf/ni
pf/saa
pf/sv
pf.sv
privacy
product/features
product/feedbackloop
productreviewsfeature
product/seller
product/sellerratings
product/tags
qualityassurance
retention
reviewquality
sales/roleofintegration
sales/sellingpoints
security
servitization
specialization
strategicinvestment
strategy
trustpilotapp
trustpilotfuture
trustworthyness
userexperience
va/ca
va/co-creation
va/cr
va/value-add
winnertakesitall
wisdomofthecrowd
wordofmouth