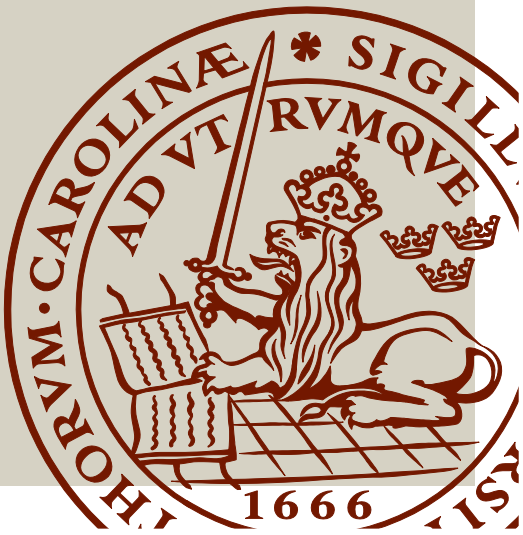
Open Data for Business Development in the Swedish ICT-sector

Hanna Bondesson and Axel Rahm

DIVISION OF INNOVATION ENGINEERING | DEPARTMENT OF DESIGN SCIENCES
FACULTY OF ENGINEERING LTH | LUND UNIVERSITY
2020

MASTER THESIS



Open Data for Business Development in the Swedish ICT-sector

Hanna Bondesson and Axel Rahm

LUND
UNIVERSITY

Open Data for Business Development in the Swedish ICT-sector

Published by
Department of Design Sciences
Faculty of Engineering LTH, Lund University
P.O. Box 118, SE-221 00 Lund, Sweden

Publicerad av
Institutionen för designvetenskaper
Lunds Tekniska Högskola, Lunds universitet
Box 118, 221 00 Lund

Subject: Innovation Engineering (INTM01)
Division: Innovation Engineering
Supervisor: Lars Bengtsson
Co-supervisor: Dennis Dreier, Combitech AB
Examiner: Jessica Wadin

Ämne: innovationsteknik (INTM01)
Avdelning: innovationsteknik
Huvudhandledare: Lars Bengtsson
Bitr. handledare: Dennis Dreier, Combitech AB
Examinator: Jessica Wadin

Abstract

Open data is broadly expected to generate innovation and economic value for both the public and private sector. Though the diffusion of open data has kept a fast pace, knowledge of visualisation and points of access of open data remains low. Moreover, literature of the private-sector application of open data is still scarce.

The first objective of this thesis is to increase the knowledge about the potential to use open data for business development for companies in the Swedish' information and communications technology sector. This is done through a literature study and a practical case study where open data is utilised. The case study showcases the potential to use open data for improved market insights and business intelligence. It also exemplifies the potential to create new products and services, through data-driven business model innovation, with open data.

The second objective of the thesis is to increase the knowledge of the technical barriers of open data from a user's perspective. Three use cases are presented, based on different Swedish open-data platforms to explore potential technical barriers. The limiting factors when utilising open data in Sweden today is found to be lack of primary, continuous data and the IT-knowledge needed to conduct a more fruitful data analysis.

The thesis shows that there is potential for ICT-companies to create or increase revenue streams through open data, but that lack of primary, continuous data could be hindering the adoption.

Keywords: open data, private sector, business development, technical barriers

Sammanfattning

Öppna data är allmänt förväntat att generera innovation och ekonomiskt värde för både offentlig och privat sektor. Även om spridningen av öppna data har ökat snabbt, är kännedom om åtkomstpunkter och visualisering av öppna data låg. Dessutom är litteraturen om den privata sektorns tillämpning av öppna data fortfarande knapp.

Det första syftet med uppsatsen är att öka kunskapen om potentialen att använda öppna data för affärsutveckling för företag i den svenska informations- och kommunikationsbranschen. Detta görs genom en litteraturstudie och en praktisk fallstudie där öppna data används. Fallstudien visar på möjligheten att använda öppna data för förbättrade marknadsinsikter. Det exemplifierar också potentialen att skapa nya produkter och tjänster, genom datadriven affärsmodellinnovation, med öppna data.

Det andra syftet med uppsatsen är att öka kunskapen om tekniska hinder för öppna data ur användarens perspektiv. Tre användarfall presenteras baserat på olika svenska öppna data-plattformar för att utforska potentiella hinder. Uppsatsen visar att den begränsande faktorn vid användning av öppna data i Sverige är brist på kontinuerliga primärdata och den IT-kunskap som fordras för att utföra en mer komplex och givande dataanalys.

Uppsatsen visar att det finns potential för ICT-företag att skapa eller öka intäktsströmmar genom öppna data, men att bristen på primär, kontinuerlig data kan hindra adoptionen.

Nyckelord: öppna data, privata sektorn, affärsutveckling, tekniska hinder

Acknowledgments

This thesis has been conducted in close collaboration with representatives at both Combitech Malmö and the Faculty of Innovation Engineering at LTH.

First, we would like to thank our supervisor at Combitech – Dennis Dreier, who has shown great passion and interest in our work as well as given much needed help and support along the process.

Second, we would like to thank our supervisor at the Faculty of Innovation Engineering – Lars Bengtsson, who has given us important feedback and helped us navigate throughout our work.

Lastly, we would like to thank our friends and families who remained encouraging during the course of this spring.

Lund, June 2020

Hanna Bondesson
Axel Rahm

Table of contents

1 Introduction

This chapter aims to introduce the topic of the master thesis by describing the background of open data and the current open-data landscape in Sweden. Furthermore, the research questions and delimitations are presented to define the scope of the study. Lastly, a disposition with summary outlines of each chapter is given.

1.1 Background to Open Data

Open data has increasingly gained attention in Europe and in 2015, the European Commission released a report together with Capgemini called "*Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources*" (Capgemini & European Commission, 2015). They estimate that the direct market size of open data in Sweden will amount to 1.713 billion euros in 2020. In Figure 1, the estimated total market size of open data in EU28+ countries can be found. They also estimated an annual growth rate of 8.2% over the period 2016–2020. The direct market size is based on revenues and gross value added that are direct results from benefits of open data, while indirect market size is based on indirect economic benefits, such as: new job potential, new goods and services, and time savings (Capgemini & European Commission, 2015).

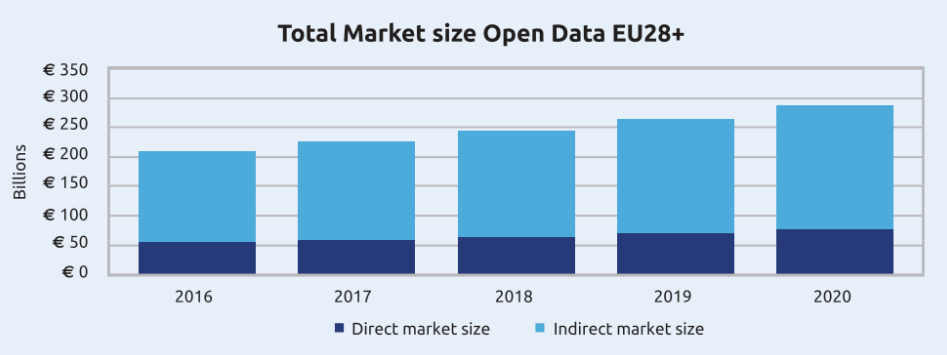


Figure 1: Total market size (high bound), split in direct and indirect size for EU28+in billions, 2016–2020 (Capgemini & European Commission, 2015). Licensee: CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

In the report, the European Commission and Capgemini write *that "public administration is by far the sector that will gain the most from opening data, with a value of 22 bn EUR in 2020. This confirms that the public sector is the first re-user of its own data"* (Capgemini & European Commission, 2015, p. 10). Janssen, Charalabidis and Zuiderwijk (2012), who researched the benefits and barriers of open data already in 2012, found that much of the research is focused on the suppliers of data. While the main focus in literature was on benefits and positive implications of open data, many barriers in terms of complexity, use and information quality have remained unexplored. They concluded that more research was needed from the user's perspective before open data systems would be freely adopted. Though several years have passed since then, a more recent article called "*Metamodel to support decision-making from open government data"* (Rojas et al., 2018) showed similar findings. The authors concluded that research into mechanisms to facilitate search, extract, analysis and use is necessary in order to take full advantage of the potential of open data (Rojas et al., 2018).

1.1.1 Definition of Open Data

The European Commission and Capgemini defines open data as *"... information that can be freely used, modified, and shared by anyone for any purpose. It must be available under an open license and provided in a convenient and modifiable form that is machine readable"* (Capgemini & European Commission, 2015, p. 7).

It is important to note that *open* refers to free or liberal and does not require the data to be free of charge. The Open Knowledge Foundation says that the data must be provided at no more than a reasonable one-time reproduction cost and that it should be downloadable via the Internet free of charge (Open Knowledge Foundation, n.d.). Moreover, only because data is visible on a website, it does not imply unrestricted re-use or distribution. Open data must be accompanied by an open license that specifies the terms of use (Ibid.).

Scientific studies on the open-data ecosystem argue that data providers can be divided into two categories: (1) organisations that provide data *for free*, with no or few limitations to the use of data, and (2) organisations that do business from selling access to the (Kitsios et al., 2017). Most organisations that distribute data for free are public administrations. This is because there are other benefits for governments when making government data public, such as those stated in the book "*The World of Open Data"* by (Charalabidis et al., 2018):

i) to increase transparency and accountability,
ii) to stimulate engagement and participation,
iii) to encourage innovation and improved efficiency for companies and society, as illustrated in Figure 2.

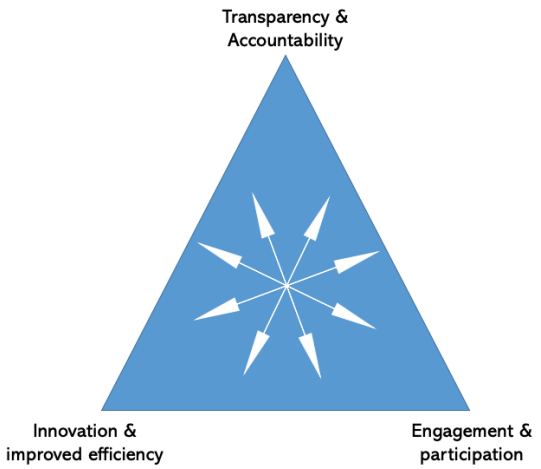


Figure 2: Objectives of open government data (adapted from Charalabidis et al., 2018).

1.1.2 Open data, Big Data, and Public Sector Information

Open government data is a subset of open data and public sector information (PSI) and open PSI is information that is collected, processed, preserved, and maintained by the government (Capgemini & European Commission, 2015). Open data and open government data can be a sub-set of big data. Big data is in turn a collection of large information sets, or "high-volume, high-velocity, and high-variety information" (Ibid.). The relationships and sub-sets of open data, big data, and PSI can be seen in Figure 3 below.

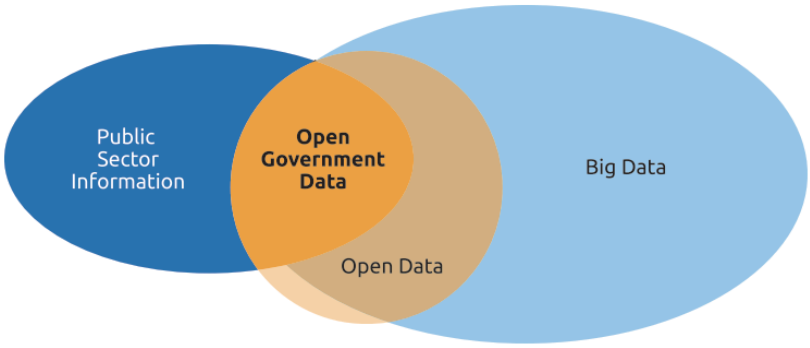


Figure 3: Relationship between PSI, open government data, open data, and big data (Capgemini & European Commission, 2015). Licensee: CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

Examples of open government data can be statistics, such as metric tons of ecological meat produced per year (Jordbruksverkets statistikdatabas, 2009), the population's level of education per year (Statistics Sweden Database, 1997), and retail sales growth by industry per month (Ibid.). An overlap of open government data and big data is, for example, weather or health data such as meteorological forecast data e.g. SMHI (2014). Where open data does not overlap with PSI or big data is usually business data which has been made public and open, an example could be customer complaints (Capgemini & European Commission, 2015).

According to DIGG, the Agency for Digital Government in Sweden, the most requested datasets are information about traffic, geographic descriptions, socioeconomics and real estate (DIGG Agency for Digital Government, n.d.). Requests for new datasets often come from private citizens, small to mid-sized companies or new businesses; and they most often relate to public

administration, geo-related data, environmental data and transport data (Ibid.).

1.1.3 Application Examples of Open Data in Scandinavia

Table 1 provides an overview on use cases for open data. The companies in the list use open data either to extend their service offerings, such as for Spotify; to tailor products for markets, such as for Volvo and IKEA; or to become more efficient in their logistics, such as for Carlsberg. There is also an example of a public-sector initiative based on open data: the online platform openaid.se, which uses open data to increase transparency of government aid funding.

Table 1: Examples of Use of Open Data in Scandinavia, Found by Authors

Company	Sector/Business area	Use of open data
Spotify	Music-streaming service	*"Spotify has taken an open approach to the data that underpins its music streaming service. When it comes to finding out about artists and music, Spotify uses the fan-created and openly licensed MusicBrainz."* (European Data Portal, 2017, p. 1).
Volvo	Manufacturing	Volvo uses temperature statistics to dimension cooling systems internationally (Forgard et al., 2020). They want to see where in the world the systems fulfil certain temperature criteria (Ibid.).
IKEA	Producer of furniture	To IKEA, a comprehensive report of temperature statistics was delivered that they use to produce products that fit all the different climates of the world (Forgard et al., 2020). They were able to divide the world market into different regions with a similar climate (Ibid.)
Carlsberg	Producer of alcoholic beverages	Climate statistics that cover the whole world (from SMHI) has been used to give the beer producer an important basis for decision-making when planning their transports on the international market (Forgard et al., 2020).
Openaid.se	Public sector initiative	*" An example of how open data is re-used in Sweden, is Openaid.se. The platform is built on open government data and visualises when, to whom, and why aid funding was paid out and what the results were. The platform aims to increase the transparency of aid spending and performance.*" (Capgemini & European Commission, 2019b, p. 3).

1.2 Background to Thesis

In interviews from 2017 with actors from the open-data ecosystem, both utilising data and providers of data, it was found that the knowledge of open data from the business perspective was limited (Kitsios et al., 2017). One service provider of open data said: *"It is important to inform and educate users about the significance and value of utilizing open data. People and enterprises have already understood their value but there is still a need for further education in the use of data"* (as quoted in Kitsios et al., 2017, p. 404). Based on discussions with supervisor Dennis Dreier from Combitech, it was decided that a case study based on utilisation of open data sources would be a good complement to existing studies on open data. Furthermore, research on the innovation adoption of open data in Sweden showed that knowledge of visualisation, points of access, data depositories and even basic knowledge of what open data is, remains low in Sweden (Temiz, 2018); a practical case study of utilising open data sources contributes to increasing this knowledge.

When talking to supervisor Lars Bengtsson from LTH it became clear that limiting the scope to one area of application, and to one market sector, would be necessary to maintain focus throughout the project. In a survey study from 2013 with 138 Swedish ICT-entrepreneurs, 43% of respondents said that open government data was essential for their business model (Lakomaa & Kallberg, 2013). ICT refers to information and communications technology, and companies in the ICT-sector are made up of manufacturing and service companies whose main activity is linked to development, production, commercialisation, and intensive use of new technology. The research from Lakomaa & Kallberg (2013) implies that start-up companies in the ICT-sector are early adopters of open data in Sweden. It was therefore considered that this market sector can provide information about the potential application areas for other sectors as well.

Consequently, investigating how open data can be used for business development for companies in the Swedish ICT-sector, was considered an appropriate scope for this thesis. In this thesis, business development is defined by creating long-term value for businesses, and involves developing processes, products, and markets, to create new opportunities and value (Pollack, 2012).

1.3 Objective

The objective of this master thesis is to increase the knowledge of the potential to use open data for business development in the Swedish ICT-sector and of the technical barriers when utilising the resource. The following two research questions are answered in this thesis:

1.3.1 Research Questions

1. How can open data be used for business development for companies in the Swedish ICT-sector (RQ1)?
2. What technical barriers exist from a user's perspective when trying to identify, access and process open data (RQ2)?

1.4 Delimitations

i) RQ1: the analysis will be limited to the ICT-sector in Sweden. The ICT-sector refers to companies selling products and services connected to information and communications technology. It is made up of manufacturing and service companies whose main activity is linked to development, production, commercialisation and intensive use of new technology. Companies in this sector are considered to be early adopters of new technology and data, such as open data.

ii) RQ2: the analysis will be limited to the data sources: Statistics Sweden, Oppnadata and Swedish Public Employment Service. Statistics Sweden is chosen, because they coordinate the system for the official statistics in Sweden (scb.se); Oppnadata, because they have been given the mission to facilitate to public sector data (oppnadata.se); and Swedish Public Employment Service, because they are the main portal for job advertisements, which is relevant when conducting market analyses.

1.5 Outline

The outline of the remaining thesis is as follows:

Chapter 2: Methodology
This chapter describes the methodology used in the thesis. The research strategy is discussed, and the research design is described. The data collection section specifies what open data sources were used as well as how. Lastly, the credibility of research is discussed.

Chapter 3: Frame of Reference
This chapter describes the theoretical frame of reference for the project. It presents methods for data-driven business model innovation and existing application areas of open data in the ICT-sector. Furthermore, adoption barriers of open data and literature findings from Sweden are presented.

Chapter 4: Case Study Analysis
In this chapter the case study results are presented and analysed. The case study aims to identify business development opportunities for a company in the Swedish ICT-sector, through utilisation of open data.

Chapter 5: Use Case Analysis
Provides insight into the technical process involved when using open data. This chapter also intends to explain the technical challenges that users face when trying to utilise open data in Sweden.

Chapter 6: Discussion
This chapter discusses the results of the case study, use cases and literature study and presents a summary of the potential application areas and technical barriers.

Chapter 7: Conclusion
This section presents the conclusions to answer the two research questions. Contributions, limitations and suggestions for future work are also presented.

2 Methodology

This chapter describes the methodology used in the thesis. The research strategy is discussed, and the research design is stated. The data collection section specifies the open data sources used. The last section discusses the credibility of research.

2.1 Work Process

The work process has been iterative and is illustrated in Figure 4 below. The first part of the work process was the design phase, which involved defining the research questions and deciding on an appropriate method of study. Simultaneously, the available academic literature on open data was reviewed, to identify research gaps in the field of open data, to which the thesis can contribute.

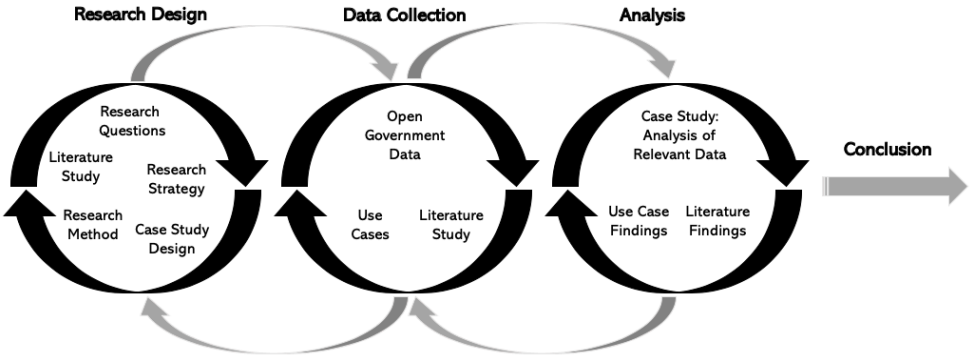


Figure 4: The work process (created by the authors)

Second, the data collection was done. This involved scanning open-data sources in Sweden and mapping out datasets that could be relevant to the case study. Moreover, the interaction between user and open-data sources were recorded to take note of potential obstacles of interoperability and begin collecting information for the use cases.

Lastly, the analysis phase was initiated. This involved visualising and modelling open data in Qlik Sense Desktop (Qlik, 2019), a free-to-use analysis tool for visualising structured data. Data-driven business model innovation theory (Sorescu, 2017), the Ansoff-matrix, sometimes called the product/market expansion grid (Meldrum & McDonald, 1995), and theory found in "*Fundamentals of Strategy"* (Johnson et al., 2012), were the foundation for the analysis and conclusions. From the use cases, technical barriers connected to identifying and accessing data were identified and compared to scientific literature on open data barriers for validation and/or critical discussion.

2.2 Research Design

Since the use of open data is relatively new and only to a limited extent explored, many studies have focused on the distributing part and the importance of common standards and policies when publishing data (Kalampokis et al., 2019). For this reason, the research in this thesis is mainly exploratory, with the goal to understand the application areas of open data outside the public sector as well as the technical issues that users' face when trying to utilise the resource.

The two research questions require different types of data collection and analysis processes and therefore, they are mainly independently treated, as shown in Figure 5. However, information regarding the technical barriers connected to *processing* open data is gathered through the case study application of open data: represented in Figure 5 by the horizontal arrow.

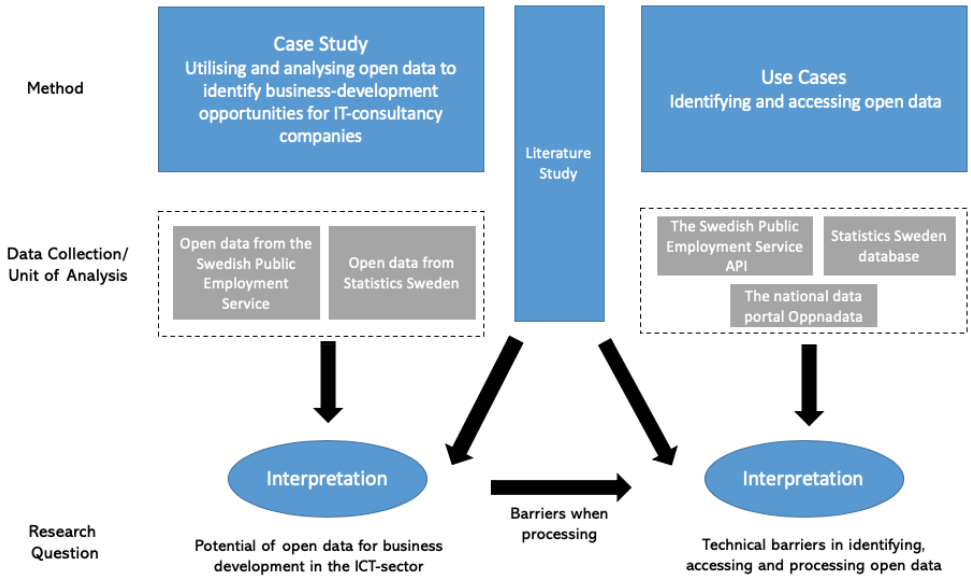


Figure 5: Research design (created by authors)

2.2.1 Literature Study

A literature study has been conducted to give theoretical background to the subject of open data, particularly regarding its application to business development, the private sector, and the current technical challenges. It brings together a comprehensive view of what existing literature has found on the subject and is, therefore, a good complement to the practical nature of the case studies. The study has been done through different sources and databases:

i) "*Google Scholar*" – a freely accessible web search engine that provides scholarly literature across several publishing formats and disciplines.
ii) *"LUBsearch"* – Lund University's common search engine for research articles, disputations, journals etc.
iii) *Others* – articles and books recommended from supervisors (Bengtsson, 2020) (Dreier, 2020).

To be considered relevant, articles either had to be related to application of open data in the private sector, or the technical barriers and quality measurement of open data. Keywords used include: "open data", "public data", "open statistical data", "value", "innovation", "decision-making", "private sector", "private businesses", "barriers" and "quality".

2.2.2 Case Study: Improved Market Insights with Open Data

In *"Fundamentals of Strategy",* Johnson, Whittington & Scholes (2012) argue that an organisation must have a good understanding of their current business situation in order to successfully develop their business. They divide such knowledge into three different categories or scopes which are shown in Figure 6: *macro-environment*, *industry or sector* and *competitors and markets. The macro-environment* is the "broad environmental factors that impact to a greater or lesser extent almost all organizations". Such factors are for instance political, technological, social, and environmental trends. *Industry* is the scope that contains all actors producing the same goods or services, as well as industry specific trends. *Competitors and markets* is the scope defined by the customers as well as the immediate competitors (Johnson et al., 2012). In line with this framework, the aim of the case study is to use open data to gain market insights for companies in the Swedish ICT-sector, which means increasing the knowledge of the business environment connected to *competitors and markets* and *industry or sector*.

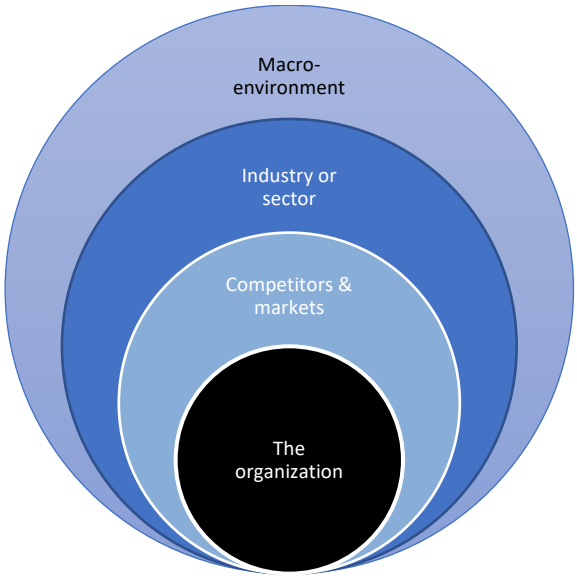


Figure 6: Layers of the business environment (adapted by Johnson et al., 2012).

The case study is based on visualising and analysing data from open-data sources in Sweden, to gain insights about the demand for data and IT competence in Sweden. These insights are then analysed in the context of what it means for a fictive IT-consultancy company, looking for business-

development opportunities and wanting to make data-driven decisions. The data that is used comes from real open-data sources, described in detail under section *2.3 Data Collection.* However, the company is fictive, and used to showcase how the improved market insights can be used to make business-development decisions. Therefore, the company is not described in detail and the choice of attributes (such as company location, competence area and size) was randomly chosen by the authors of the thesis. Information about the fictive company is provided under section *4.1 Description of Fictive Company.*

The case study is used to give practical insight into open data sources and the potential use of the resource. This is the goal of a case study approach, to conduct an in-depth analysis of a phenomenon within a real-world context (Guetterman & Fetters, 2018). A single case study approach was used in this research. The reason for choosing a single case study is that it represents the phenomenon of interest, i.e. utilisation of open data, and allows for a more concrete analysis of an otherwise rather intangible subject.

2.2.3 Use Cases on Swedish Open-Data Sources

A use case is a method for evaluating some functionality of a system. In the book "*Object Oriented Software Engineering*: *A Use Case Driven Approach"* by Jacobson (1992), the methodology and purpose of use cases is described. The goal of doing a use-case analysis is to examine a specific part of a system in great detail and draw conclusion regarding its drawbacks. By doing a use case from an external user's perspective, i.e. not as an internal developer or maintainer, insights can be generated that otherwise would have been missed otherwise (Jacobson, 1992). Some essential definitions are given next that are used throughout the thesis.

2.2.3.1 User

The *user* is a person who will interact with a system in a certain way (Jacobson, 1992). Jacobson uses a recycle machine to exemplify a use case. In that case, one user would be a customer who wants to recycle her cans in the machine, and a second user would be the machine operator who wants printed status reports from the machine regarding the number of recycled items. In each use case, focus is only on one type of user. To get a broader view of the interaction between system and users, multiple use cases for several users can be conducted (Jacobson, 1992).

2.2.3.2 Problem domain

The *problem domain* is the part of the system that the defined user interacts with (Jacobson, 1992). The goal is seldom to evaluate the entire system in a single use case, instead, one breaks down the system in smaller sub-system domains and describes the user's interaction in detail. In the recycling machine example, one problem domain could be the screen and buttons, with the instructions for recycling, the hole for items to be deposited and the final receipt. These parts of the system define the interaction between the machine and the (user) customer (Jacobson, 1992).

2.2.3.3 Initialisation

Once the user and problem domain are defined, one can begin to initialise the use case (Jacobson, 1992). The goal of the use case is to identify all possible turn of events in the interaction between the user and the system in order to evaluate both functionality and user interface. If one wants to do such an evaluation for the entire system, one could do several use cases, one for each type of user (Jacobson, 1992).

2.2.3.4 Minimal and success guarantee

The use case begins with a definition of the *success guarantee* (Jacobson, 1992). That is, what is the desired outcome when the user interacts with the system? In the case with the recycling machine, the success guarantee would be that the customer successfully deposits all her items and the machine prints a receipt of her total amount. The *minimal guarantee* is the worst acceptable outcome from the interaction. An example of minimal guarantee could be that the machine does not recognise any of the items that the customer wishes to deposit, and informs her via the user interface (Jacobson, 1992)

2.2.3.5 Main success scenario

The main success scenario is a step-by-step description in chronological order of what happens in the interaction between the user and the system before the success guarantee is reached (Jacobson, 1992). Through this process, the system is studied in detail and possible inadequacies could be detected (Jacobson, 1992).

2.2.3.6 Extends

"A powerful concept that is used to structure and relate use case descriptions is the extend association" (Jacobson, 1992). Extends are possible deviations

from the main success scenario. These are also listed in chronological order, like the main success scenario, but describe all alternate outcomes other than the main success scenario. An example could be that an item gets stuck in the recycle machine (Jacobson, 1992)

2.3 Data Collection

The thesis is based on three components of data collection: open datasets relevant to answer RQ1, qualitative information gathered when writing the use cases for RQ2, and the literature study to compare and extend the empirical results with existing literature.

2.3.1 Open Data (RQ1)

First, open data relevant to the ICT-sector was identified using the Swedish national open data portal (Öppnadata och PSI). The data originates from the Swedish Public Employment Service (Arbetsförmedlingen) and Statistics Sweden (SCB). Statistics Sweden is responsible for coordinating the system of official statistics in Sweden, and authors believed that it was a suitable source of nation-wide statistics concerning the macro-scope. Swedish Public Employment Service on the other hand, was chosen to provide ICT specific information regarding demand for IT-services. Other sources of open data were available, but they were left unexplored due to an insufficient time frame.

2.3.1.1 The National Data Portal, Oppnadata

One main component for identifying open data has been through the national open-data portal (Öppnadata, n.d.). They play an important role in identifying open data, since they allow users to search for data from many different sources. The basic functionality of open-data portals include metadata descriptions, which allow the user to search for datasets via keywords, resource format, publisher, topic categories and location (Charalabidis et al., 2018). However, Oppnadata does not publish the data itself, instead the website provides references to the original sources. Oppnadata was used to identify relevant datasets to the case study by filtering on topic category.

2.3.1.2 Data from the Swedish Public Employment Service

From the data portal and through discussions with supervisors, data from the Swedish Public Employment Service was identified. Their open data consists of job-advertisement information, published on their website. Two types of primary data were available and collected. The first type were all employment ads for ICT-positions published at their website between 2010 and 2017 (188,434 ads), which were available for download. This data was chosen to investigate any long-term trends in the ICT job market in Sweden. The second type were all ads that were active on their website at the time of the data collection process for the analyses presented in this thesis (3,445 ads). The authors collected these ads through an API (Application Programming Interface) provided by Swedish Public Employment Service. The API provided the authors with a set of rules on how to interact with the database containing the job ads.

Complementary data regarding geographical areas was gathered to link the data points with datasets from Statistics Sweden, described below. All ads were downloaded in JSON (JavaScript Object Notation)-format, concatenated with the complementary data, cleaned, and converted into CSV (Comma Separated Values)-format using self-developed Python scripts.

2.3.1.3 Data from Statistics Sweden

Through Oppnadata, the authors found that several datasets regarding information on the ICT sector originated from Statistics Sweden. Statistics Sweden collect various datasets from Swedish authorities and via surveys and makes those results accessible as secondary data via statistikdatabasen.scb.se. If one wanted the primary data, Statistics Sweden charges 30–60 thousand SEK with a 1 – 6 months delivery time depending on the usage purpose (Statistics Sweden, n.d.). Only free of charge data was used in this thesis to increase transparency and reproducibility at low cost.

The authors searched the statistics database for those datasets containing useful information about the ICT sector. Such datasets were found under categories such as *Enterprises' IT expenditure* and *ICT usage in enterprises*, etc. The datasets could be filtered using pre-determined filters, e.g. company size or region, and downloaded as CSV-files. Since the data mainly consisted of survey results and official reports, the available data from Statistics Sweden was updated more rarely than their counterpart at the Swedish Public Employment Service. The data in the datasets was aggregated to annual values.

2.3.2 Information for Use Cases (RQ2)

The data collection for the use cases involved interacting with the open-data systems and exploring how the system responded based on different requests, such as requesting all ICT ads from 2015 or requesting the IT-expenditure from different business sectors 2018. Note that throughout the thesis, the authors take the role as *users* and that the results of the use cases are based on the authors experience when interacting with the platforms. The use cases are based on the same open-data platforms as used in the data-collection for RQ1, namely the national data portal, the Swedish Public Employment Service and Statistics Sweden. The information was noted in the form of use cases according to the description in section 2.2.3.

2.4 Data Analysis

2.4.1 Visualising Data in QlikSense (RQ1)

For the case study, the data was cleaned, merged and visualised through QlikSense Desktop (Qlik, 2019). After the data preparation process, a data model was developed in QlikSense that links the different data sources into a coherent schema as shown in Figure 7. In the figure, all the attributes from the data sources can be seen. The analysis of open data is conducted so that it relates to theory on business development.

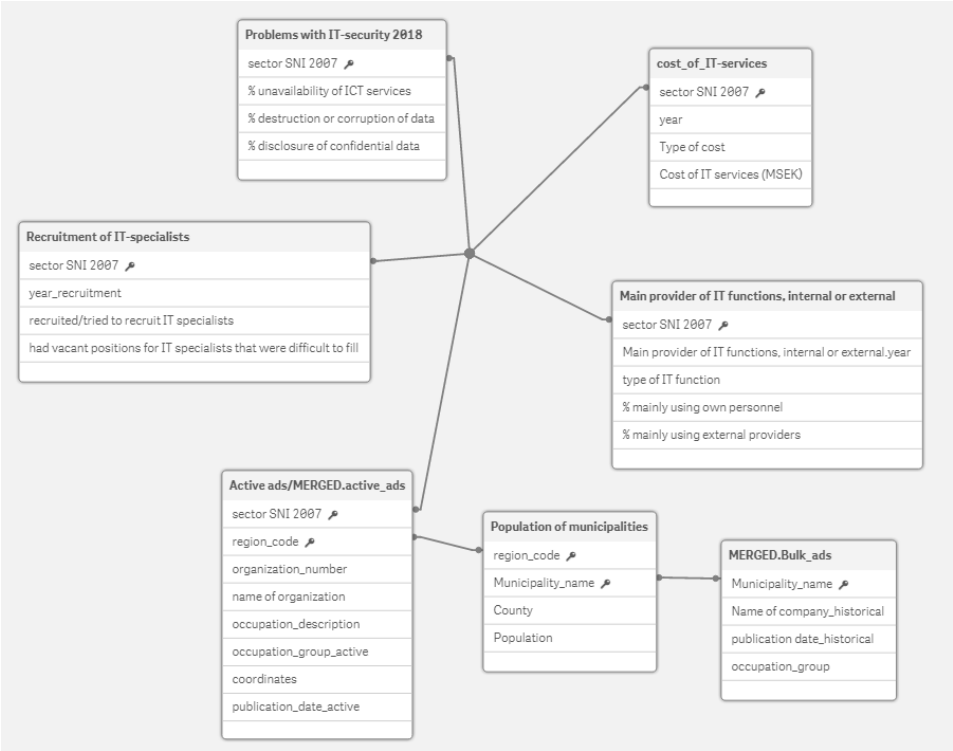


Figure 7: Visualisation of the data model in form of a Snowflake schema in Qlik Sense Desktop. The following datasets come from Statistics Sweden: problems with IT-security 2018, cost of IT-services, recruitment of IT-specialists, main provider of IT-functions (internal or external) and population of municipalities. Active ads and Bulk ads come from the Swedish Public Employment Service.

2.4.1.1 Business Development Theory: the Ansoff Matrix

As previously mentioned, the case study uses a fictive case company to showcase how the insights can be used on a strategic level. Therefore, a framework to show the strategic directions that companies can take is needed. The Ansoff-matrix (Meldrum & McDonald, 1995) is a simple depiction describing the options that organisations have for increasing revenue or profitability, see Figure 8. It encapsulates all of the strategic directions that an organisation can adopt and focuses on the relationship between new and existing products and customers (Meldrum & McDonald, 1995).

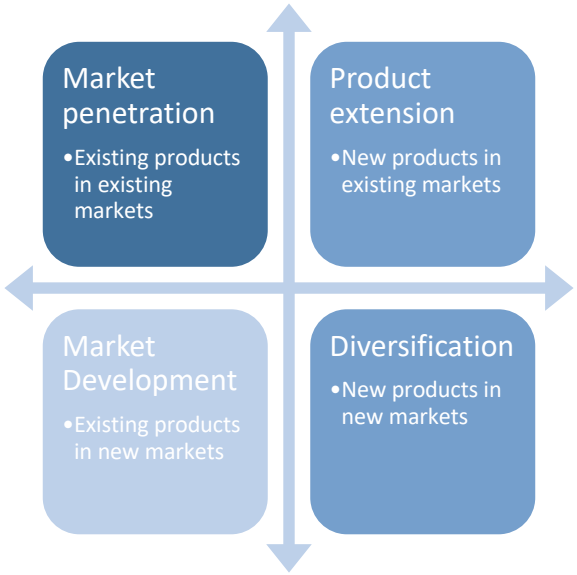


Figure 8: The Ansoff matrix (Meldrum & McDonald, 1995).

Consequently, companies have four strategies to choose from: concentrating on improving existing products to existing markets, developing new products for existing markets, entering new markets with existing products or diversifying with new products for new markets (Meldrum & McDonald, 1995). The framework is used to provide the fictive company with recommendations for developing their business.

2.4.2 Analysing the Use Cases (RQ2)

The use cases were analysed based on a methodology for information quality assessment presented by Lee, Strong, Kahn & Wang (2002). In Table 2, the statements used to evaluate different quality measures are described. These metrics are referred to in articles and books when discussing the quality of open data, e.g. by Vetrò et al. (2016) and Charalabidis et al. (2018). The statements under "*Description"* in the right column were used to identify where barriers exist. If the statements attune with the user's opinion, the quality is considered high. If the user believes a statement followed by an R (reverse) to be accurate, it means that the quality connected to this measure is considered low. Note that the users evaluating the platforms are the authors of this thesis.

Table 2: Information quality measurements (Lee et al., 2002)

Measure	Description
Accessibility	This information is easily retrievable. This information is easily accessible. This information is easily obtainable. This information is quickly accessible when needed.
Appropriate amount	This information is of sufficient volume for our needs. The amount of information does not match our needs. (R) The amount of information is not sufficient for our needs. (R) The amount of information is neither too much nor too little
Completeness	This information includes all necessary values. This information is incomplete. (R) This information is complete. This information is sufficiently complete for our needs. This information covers the needs of our tasks. This information has sufficient breadth and depth for our task
Concise representation	This information is formatted compactly. This information is presented concisely. This information is presented in a compact form. The representation of this information is compact and concise.
Consistent representation	This information is consistently presented in the same format. This information is not presented consistently. (R) This information is presented consistently. This information is represented in a consistent format.
Ease of operation	This information is easy to manipulate to meet our needs. This information is easy to aggregate. This information is difficult to manipulate to meet our needs. (R) This information is difficult to aggregate. (R) This information is easy to combine with other information.
Accurate and objective	This information is correct. This information is incorrect. (R) This information is accurate. This information is reliable.
Interpretability	It is easy to interpret what this information means. This information is difficult to interpret. (R) It is difficult to interpret the coded information. (R) This information is easily interpretable. The measurement units for this information are clear.
Timeliness	This information is sufficiently current for our work. This information is not sufficiently timely. (R) This information is not sufficiently current for our work. (R) This information is sufficiently timely. This information is sufficiently up to date for our work.

2.5 Credibility of Research

The "*Criteria for Assessing the Trustworthiness of Naturalistic Inquires"* according to (Guba, 1982), state that the credibility of a qualitative study should be evaluated based on the four aspects of trustworthiness:

i) Credibility
ii) Transferability
iii) Dependability
iv) Confirmability

The four aspects are described in the following.

2.5.1 Credibility

Credibility refers to the truth value of the collected data as well as the interpretation of ditto (Guba, 1982). To test the credibility, one should collect data from multiple sources and perform so-called *member checks*, i.e. letting other people interpret the collected data (Guba, 1982).

The authors collected data from multiple reliable sources which were distributed by government agencies. Furthermore, the authors had continuous discussions with their academic supervisor as well as co-supervisor from Combitech concerning the choice of datasets and interpretation of results. The results and analysis of the use cases have been validated with supervisor Dennis Dreier, Systems Designer/Architect at Combitech.

2.5.2 Transferability

Transferability refers to the applicability of the study, i.e. to which extent the result would hold in alternative setting (Guba, 1982). To test the transferability of a study the authors should give a detailed description of the data and the data collecting process. Such description should provide the reader with an adequate possibility to determine the generalisability of the results (Guba, 1982). To increase transferability a detailed description of the data collection process is provided as well as references to the sources: (JobTech Development & Swedish Public Employment Service, n.d.) (Statistics Sweden Database, 1997).

2.5.3 Dependability

Dependability refers to the ability to replicate a study (Guba, 1982). Guba argues that no study could be considered trustworthy if it is not replicable. The authors have given a thorough description of the method, data collection and analysis method to facilitate replication of the study. Notably is that all datasets used in this thesis are available for free and can be downloaded, which further eases the process of replication.

2.5.4 Confirmability

Confirmability refers to the objectivity and neutrality of a study, i.e. to which extent the result is data driven, not a result of the author's predispositions (Guba, 1982). Both authors have been involved in every part of the data collection and the process of conclusion. Guba argues that by including more researchers' points of view in the research process, one limits individual predispositions, in favour for more data driven results (Guba, 1982).

3 Frame of Reference

This chapter describes the theoretical frame of reference for the project. It presents methods for data-driven business model innovation and existing application areas of open data in the ICT-sector. Furthermore, adoption barriers of open data and literature findings from Sweden are presented.

3.1 Data-Driven Business Model Innovation

The article "*Data-Driven Business Model Innovation"* (Sorescu, 2017) lays the foundation for the theory needed to investigate the business development potential of open data. The article covers how organisations, mainly ICT-companies, can leverage big data and information networks to innovate their current business models, and develop new ones.

Sorescu (2017) present three methods to build or improve a business model around big data: *leveraging external data, leveraging internal data* and *microanalytics to macroanalytics.* These are described in Table 3 below.

Table 3: Data-Driven Business Model Innovation Methods (based on Sorescu, 2017)

Method	Description	Examples
Leveraging External Data	Involves retrieving, repackaging, and selling relevant data from external sources.	*"Gnip, a social media API aggregation company that gathered data from specific providers (e.g., Twitter, Digg, or Vimeo) and instantly sent it to consumers"* (Sorescu, 2017, p. 694)
Leveraging Internal Data	Involves either: (i) analysing internal customer data to serve them better and create revenues, or (ii) accumulating data about consumers and selling the relevant information to other firms	*"product recommendations on Amazon.com are based on a sophisticated analysis of the browsing and purchase data of its customers."* (Sorescu, 2017, p. 694) *"The challenges of harnessing insights from big data are turning former hardware and software manufacturers like Microsoft and IBM into big data analytics service providers."* (Sorescu, 2017, p. 695)
Microanalytics to Macroanalytics	Involves transforming microdata, for instance consumer-level data by aggregating it with other such data to create macroanalytics that can assist policy makers.	*"Assume a health organization needs to make available a vaccine that is in short supply: what is the best way to deploy it? Individual patient data could be aggregated to the general population level in order to help the health organization make a more informed decision (Schmarzo, 2015)"* (Sorescu, 2017, p. 695)

3.2 Application Areas of Open Data in the ICT-sector

There is an opportunity for ICT-companies, with knowledge and capacity to handle the velocity and variety of big data, to turn big data into business intelligence for themselves and other organisations. Potentially, open data can also support the data-driven business model innovations presented by Sorescu (2017). To begin to understand how open data is used in the ICT-sector today, two research articles are used:

1. *"Open government data and the private sector: An empirical view on business models and value creation"* (2017), Magalhaes and Roseira analyse 178 firms that use open government data across various industries in the U.S.. They argue that there are two ways that open data is used in the private sector to build competitive advantage: (1) through the creation of new products and services, and (2) through improved market insigths, products, services and processes (Magalhaes & Roseira, 2017).

2. "*Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs"* (2013), Lakomaa and Kallberg describe how open data influences the ability for Swedish ICT-entrepreneurs to execute their business plan. The findings are based on survey data with 138 Swedish ICT-entrepreneurs; 43% of respondents said that open data was essential for their business model, and 82% that access to open data would support and strengthen their business plan (Lakomaa & Kallberg, 2013).

Magalhaes and Roseira (2017) found several examples of products and services that were created through open data in the private sector. The main themes were those connected to: *advocacy, consultancy, data refining and structuring, single purpose apps, interactive apps, data platforms and open data portals.* The study also showed that all of these uses were prevalent in the ICT-sector. Morover, connected to improved market insights, products, services, and processes, Magalhaes & Roseira (2017) found that companies in the ICT-sector utilise open data for *business intelligence*, *process optimisation*, *product/service improvement* and *research and development.*

Lakomaa and Kallberg (2013) found that the primary use of open data among Swedish ICT-entrepreneurs was to *resell public sector information* after the data had been processed and merged with other data. In other words, using open data to create new information products. Furthermore, Lakomaa and Kallberg (2013) found another application area of open data for Swedish ICT-entrepreneurs: to get *information about potential markets*. For start-ups, internal data can be limited or non-existent. However, the information gathered by government agencies is enormous compared to any private data repositories and since each sector of society has a government agency or department collecting information to manage public programs, and monitor developments etc., this data can be used instead (Ibid.).

3.3 Technical Barriers of Open Data

As discussed, much of the research into the technical barriers of open data focuses on the supply side of open data (see e.g. Zuiderwijk & Janssen, 2014; Martin, 2014). However, two projects researching the adoption barriers of open data include the user's perspective:

1. In *"Benefits, Adoption Barriers and Myths of Open Data and Open Government",* Janssen, Charalabidis, and Zuiderwijk (2012) present issues with open data that are hindering the adoption. The findings are based on a workshop and interviews with different actors in the open data landscape, including users.
2. In Temiz's doctoral thesis "*OPEN DATA AND INNOVATION ADOPTION: Lessons from Sweden"* (2018), he investigates the potential determinants of open data adoption for open, data-driven innovation. The findings are based on a survey study with individuals active in the open data ecosystem in Sweden.

Mainly, Janssen et al. (2012) found barriers connected to themes they call *task complexity* and *use and participation* from users. Use and participation has more to do with not understanding how to make use of the data or lack of statistical knowledge. However, issues related to task complexity gives insight into technical barriers encountered by users. For instance, Janssen et al. found that users found it difficult to discover appropriate data, as well as searching and browsing for data. Another issue was not having access to original data, i.e. only processed data. Users also found it difficult that data could be provided without explanations making it difficult to interpret (Janssen et al., 2012).

One of the first barriers that Janssen et al. mention, is having to registrate and pay a fee to get access to open data. While this is not merely a technical issue it hinders the use. Lakomaa & Kallberg (2013), for instance, found that the willingness to pay for data was far lower than the cost for data made available by government in their study of Swedish ICT-entrepreneurs. One of the government agencies that they mention, the Swedish cadastral and land survey authority (Lantmäteriet), has now made their open geo data available through an API without charge (Lantmäteriet, n.d.), but the Swedish Incorporation Authority (Bolagsverket), for instance, still denies access to free public data (Bolagsverket, 2016).

Furthermore, Temiz (2018) writes that open data often is described as resembling open source software, where the user has access to the source code and can modify it freely, which is not in line with what his research found. Instead, the current state of open data in Sweden more resembles closed applications that provide open APIs to third-party developers. Third-party developers do not have any rights or tools to access the source code and can therefore not modify the code or data to fit their needs (Temiz, 2018).

3.4 Open Data Maturity in Sweden

In a larger context, there are differences between how far countries in the EU have come in policy making, publishing, and re-use of open government data, which influences the user experience. Capgemini and the European Commission release a report describing the open-data maturity level of EU countries every year (Capgemini and European Commission, 2019a). In it, countries are compared in terms of policies, portals, impact, and quality. In *"Open Data Maturity Report 2019"* the overall maturity level rating in Sweden is 55%, while EU-average is 66% (Ibid.).

Looking solely at the quality measurement, Sweden ranks lower than average, with a score of 58% while the EU-average is 65% (Capgemini & European Commission, 2019b). In Figure 9 the individual quality dimensions and the Swedish' scores are shown. It is important to note that the evaluation is based on the metadata quality, and not the quality of single datasets (Capgemini & European Commission, 2019a, p. 51-59). A high quality-score means that more data is accessible through data portals (Ibid.), which is important for identification of data.

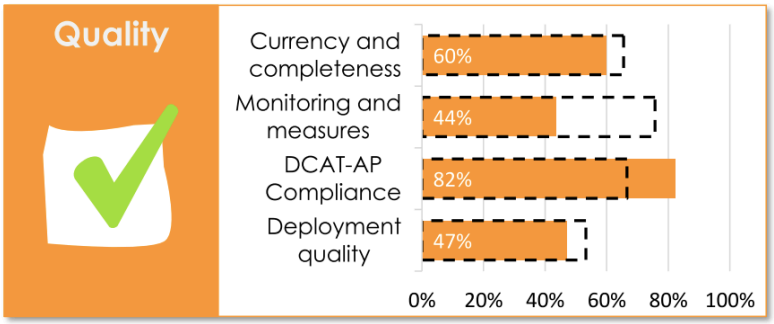


Figure 9: Quality Dimension Performance of Open Government Data in Sweden (orange), compared to the EU28 average (dotted) (Capgemini & European Commission, 2019b). Licensee: CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

4 Case Study Analysis

In this chapter the case study results are presented and analysed. The case study aims to identify business development opportunities for a company in the Swedish ICT-sector, through utilisation of open data.

4.1 Description of Fictive Company

The analysis is based on the perspective of a fictive start-up company located in Lund, Sweden. The company consists of five consultants, specialising in web development, both front-end and back-end, and their current customers are primarily larger manufacturing companies in Skåne. With a couple of profitable years behind them, the company wishes to develop their business model. The amount of internal data is limited, and the company seeks to utilise open data to make an informed, data-driven decision.

The company wishes to *gain a thorough understanding of its market environment*. In the analysis below, geographical and competence insights are generated by analysing the employment ads for data and IT positions over the period 2010-2017 (JobTech Development & Swedish Public Employment Service, n.d.). Additionally, insights regarding consumption of IT-services for potential markets is gained by combining datasets from Statistics Sweden (Statistics Sweden Database, 1997).

4.2 Demand for Data and IT Competence

Over the period 2010–2017, 188 000 employment ads within data and IT have been published at the Swedish Public Employment Service. Half of the sought-after competences have been within software and systems development, as shown Figure 10. In March of 2020, a sample dataset of 3,445 active ads was collected, and showed the same tendency; 56.6% of the active ads were for software and system development positions.

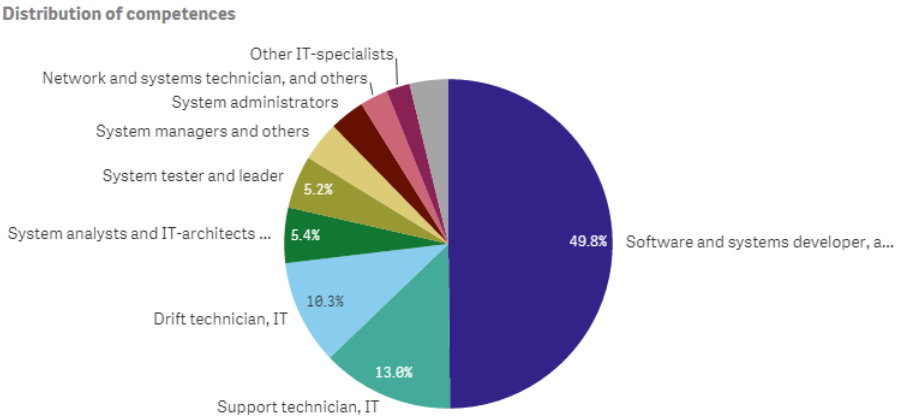


Figure 10: Distribution of Data and IT positions from The Swedish Public Employment Service, 2010-2017 (188,000 employment ads)

Looking historically, the number of ads for software and systems developers has been consistently increasing since 2013, (see Figure 11). From then, the ads have been growing, on average, by 23% per year. While other occupation groups also have been growing in demand, such as system analysts, IT-architects, and drift technicians, none of them have increased in the same rate as software developers and system developers.

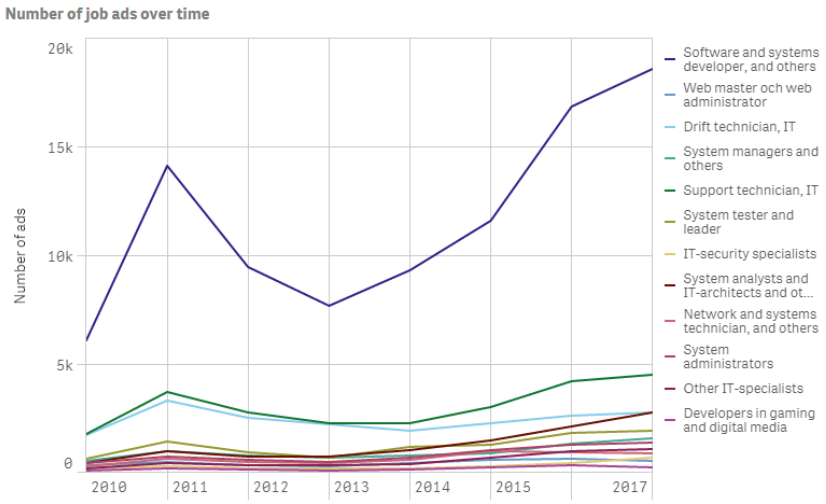


Figure 11: Number of job ads for different occupation groups, 2010-2017. (188,000 employment ads)

The demand for software and system developers has been growing fastest, and already represent more than 50% of the IT-demand. Web development falls within this category which bodes well for the company. Looking within this group, the most prominent demand is for programmers, as seen in Figure 12. Frontend and backend-developers account for roughly 14% of the demand.

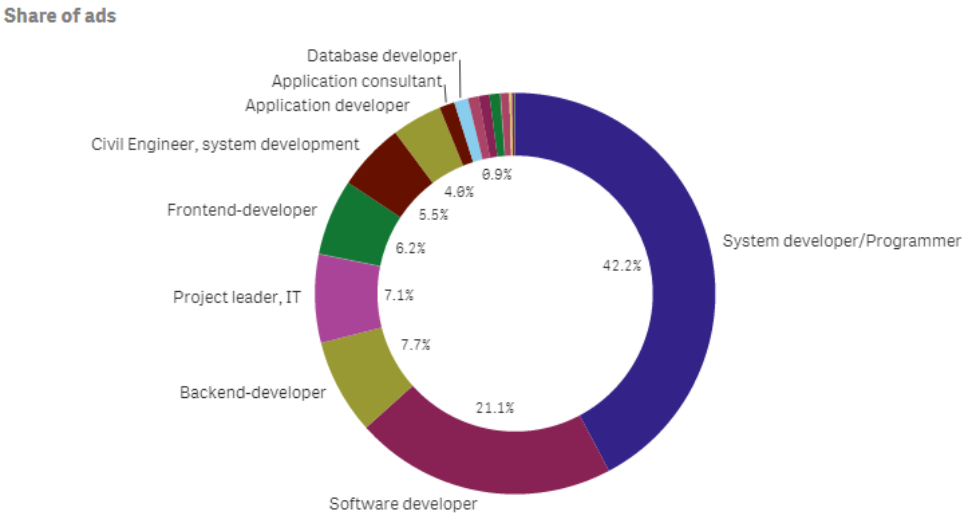


Figure 12: Employment ads for software and system developers split into competence area, 2010-2017. (94,000 employment ads)

4.3 Demand for Data and IT Competence across Sweden

Figure 13 shows the ad density, i.e. number of employment ads per capita, of data- and IT-related positions across Sweden. The data is presented on municipality level, and the values are adjusted according to the population in each area. The figure show that the number of job ads per capita is higher in the county of Stockholm, Östergötland and Västergötland. Furthermore, the ad density is particularly distinct for the cities of Stockholm, Linköping and Gothenburg.

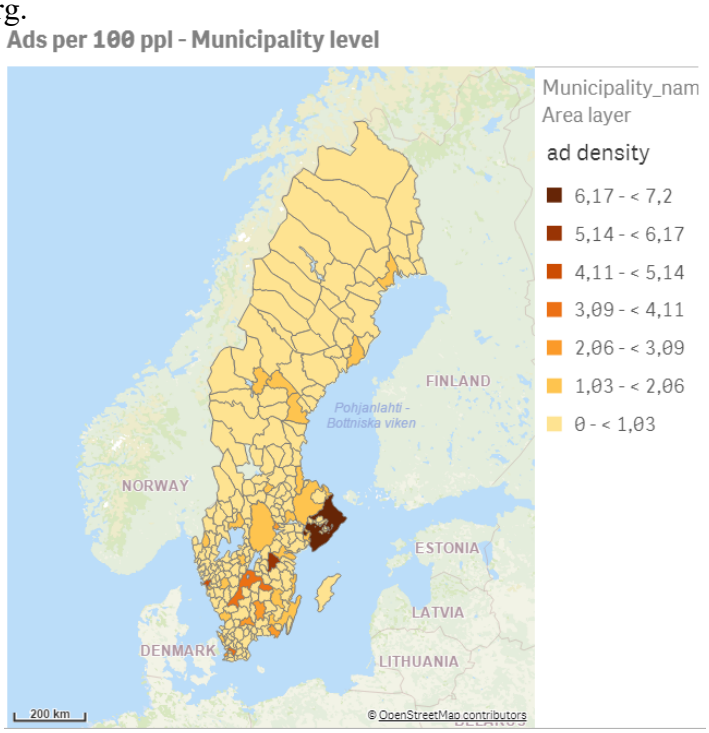


Figure 13: Number of job ads for Data & IT positions per 100 people, Municipality level, 2010-2017. (188,000 employment ads)

When analysing this on a yearly basis, the data shows that the above-mentioned municipalities have remained the most advertised regions during the time period 2010–2017, as shown in Figure 14. At the same time, other locations have emerged and gained in importance such as Jönköping, Lund and Luleå.

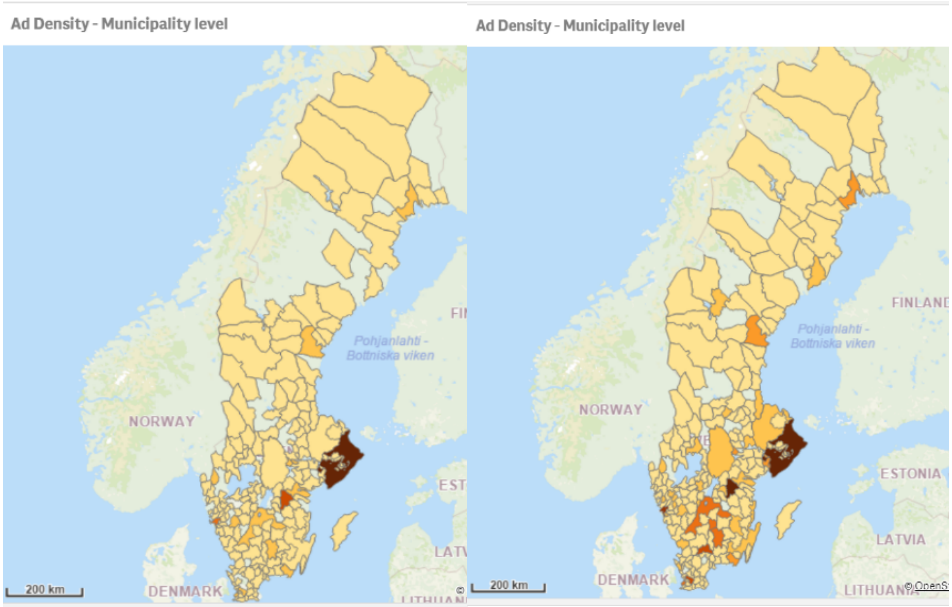


Figure 14: Number of job ads for ICT positions per 100 people, Municipality level, 2010 and 2017. (188,000 employment ads)

The job ad density for data and IT positions could be interpreted as a measurement for a region's demand for IT knowledge. The analysis shows that population size is not necessarily decisive; Linköping had a significantly larger demand than more populated cities such as Malmö, Uppsala and Örebro. In areas with higher demand, it could be expected that many companies request ICT personnel and that the market for IT-solutions is generally larger.

In order to test if these historical observations are still relevant, Figure 15, shows the same mapping but for the sample of active ads at The Swedish Public Employment Service in March 2020. Although the dataset is not as extensive as the previous one, it reveals the same tendencies, i.e. Stockholm, Linköping and Gothenburg and shows the highest demand of ICT competence per capita.

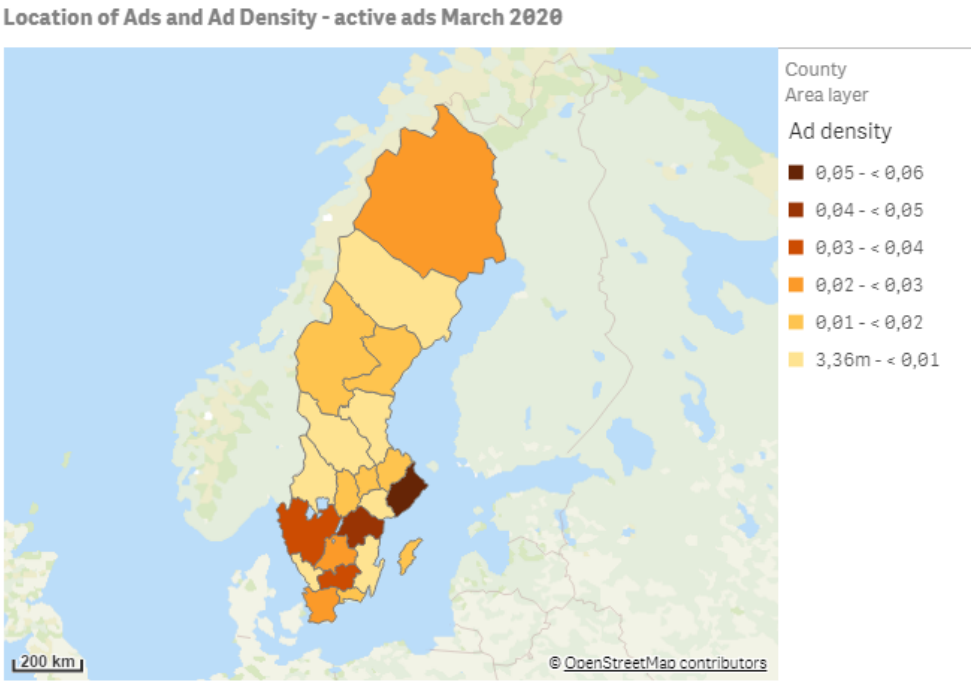


Figure 15: Number of job ads for ICT positions per 100 people, County level. Active ads, March 2020. (3,445 employment ads)

4.4 Demand for Data and IT Competence across Sectors

Data from Statistics Sweden shows that there are some differences between the business sectors concerning which IT-functions require external help, see Figure 16. Out of the sectors*, real-estate companies and managers, energy and recycling* as well as *manufacturing*, are the sectors in which enterprises have the highest demand for external providers of IT-services.

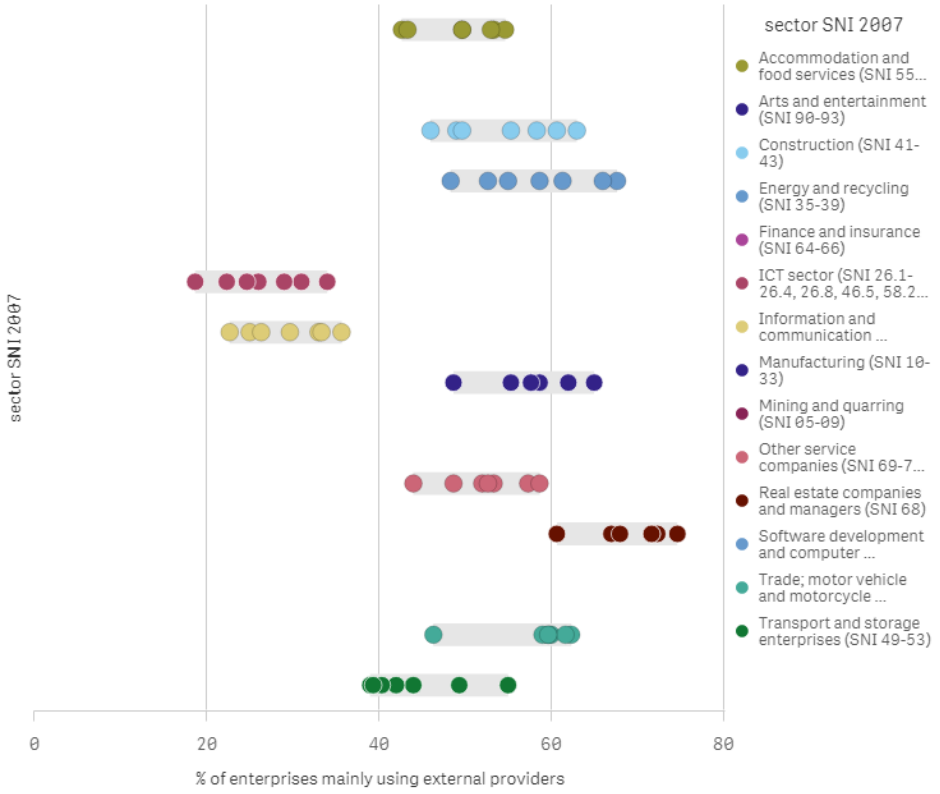


Figure 16: percentage of enterprises mainly using external providers for IT-services, across different sectors. The data points represent different IT-service types. (source: Statistics Sweden).

In Figure 17 below, each sector is represented by a colour, and within the sector block, the partition shows the distribution of ads for different occupation groups. For example, in the sector for *wholesale and retail trade*, more than 50% of the IT-oriented job advertisements are for software and system developers. The same goes for all other sectors, where almost half of the IT-employment ads are related to software and system developers.

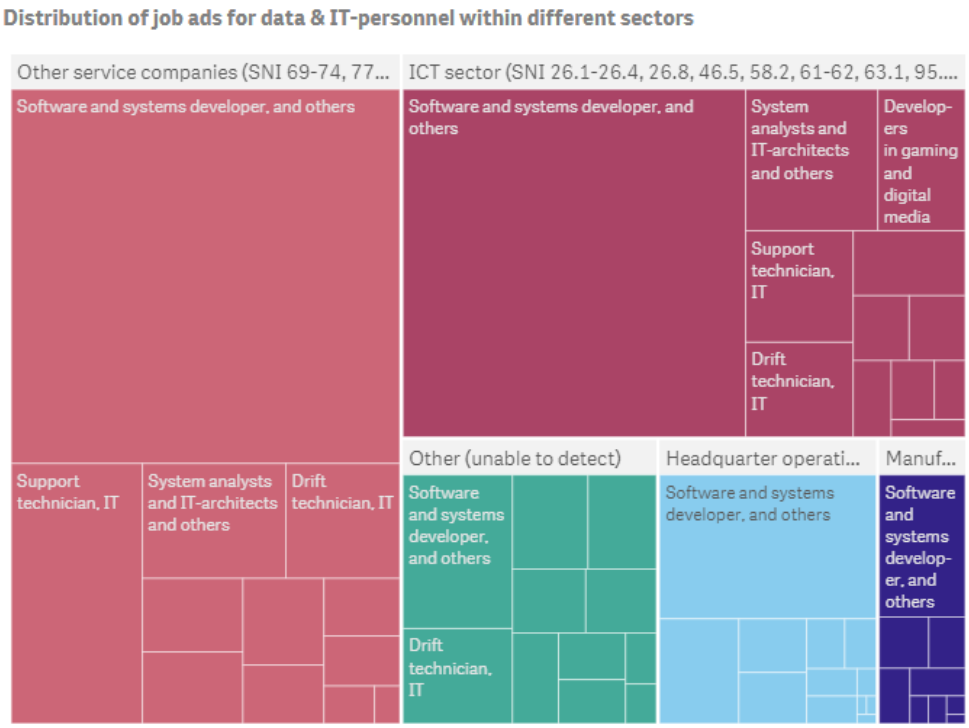


Figure 17: Distribution of job ads for data and IT-personnel in different sectors (source: 3,445 active ads on The Swedish Public Employment Service in March 2020 & complementary industry data).

4.5 IT-Service Expenditure across Sectors

The market size IT-services in each of these sectors is also of interest. Based on statistics on IT-expenditure in different sectors, it was found that the retail industry spent 15 billion SEK on external IT-services in 2018, see Figure 18. Furthermore, the expenditures have been growing by an average of 14% per year since 2014. Besides, *manufacturing* and *finance and insurance* companies spent more than 6 billion SEK each. Based on this, the company already serves an attractive sector, i.e. the manufacturing sector, but could want to explore opportunities in the retail industry.

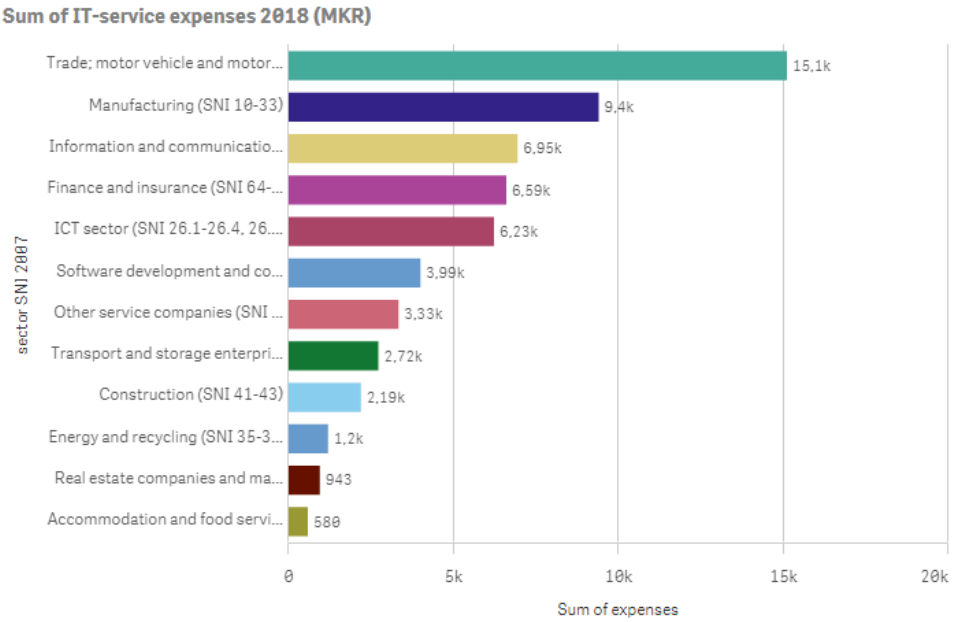


Figure 18: IT-service expenditure in 2018 for different sectors (source: Statistics Sweden)

4.6 Business-Development Opportunities for Fictive Company

While the visualisations and market insights are based on real open data, the coming recommendations are based on the fictive case company setting. Based on the insights obtained by the analysis, the case company can find business development opportunities by using the Ansoff matrix, introduced in Section 2.4.1.1, either by developing their services or markets.

4.6.1 Service Development

The analysis showed that the company is providing a service with high demand: software and system development. Within that category, however, backend and frontend development does not seem to be as desired as programmer competences. It is possible that the tasks performed by web, backend and frontend developers also falls into the category *programmers*, and that the companies seeking such competence are using a more generic title. It is also plausible that the competence needed for *programmer* positions resembles those possessed by the current employees. The company should not be discouraged by the relatively small demand for frontend and backend developers but rather seek the opportunity to broaden its supply to similar services.

4.6.2 Market Development

The analysis showed that the company's current sector – manufacturing sector, is an attractive market for IT-service providers. The retail industry has, however, had even higher increased demand for IT-services, implying an opportunity to take market shares in this sector. The company can explore the opportunity to serve the retail sector, perhaps through their current business contacts or by marketing campaigns. Additionally, the geographical analysis shows that the demand for ICT services is larger in Stockholm, Göteborg and Linköping. The higher density of employment ads from these cities might indicate that there already exist many ICT companies there that recruit. If the company were to locate an office in Linköping for example, they can harvest the fruit from such cluster, e.g. in business partners, access to competence, etc. By opening a new office, instead of a complete relocation, the company can maintain its lucrative business in Skåne.

4.6.3 Recommendation for Fictive Case Company

Based on the open data the company should: (1) maintain their service within web development segment and explore the opportunity to use current competences elsewhere, (2) try to serve the retail business in order to gain market shares and (3) consider opening a new office in a city with a high density of ICT demand to improve business opportunities and benefit from similar companies.

5 Use Case Analysis

This chapter provides insight into the technical process involved when using open data. This chapter also intends to explain the technical challenges that users face when trying to utilise data. This can be relevant to the publishers of open data to improve their services. Moreover, it is relevant to actors who plan to use open data, since the chapter describes the steps involved in identifying and accessing open data.

5.1 Overview of Use Cases

A use case describes the behavioural contract between stakeholders with interests (Jacobson et al., 2011). Since companies can use many different analysis programs (e.g. Excel, Qlik Sense, Power BI, Tableau or other business intelligence tools) to handle and analyse data, the focus is here on the interaction with open data platforms for retrieving data as well as data transformation processes. The use cases are described in Table 4.

Table 4: Use Case Briefs

Actor	Goal	Brief
Data Analyst	Extract filtered open government data through API	Data Analyst adds metadata (dates, companyID, industryID, coordinates, etc.) to API-request and data is extracted.
Data Analyst	Downloading datasets from SCB	Data analyst finds relevant data sets from Statistikdatabasen and downloads them using desirable filters.
Data Analyst	Finding open data sources from oppnadata.se	Data analyst finds and gets redirected to sources for relevant open data.

5.2 Downloading Filtered Open Data from the Swedish Public Employment Service

The Swedish Public Employment Service gives users access to primary (raw) data, i.e. data which has not been altered or processed by a secondary actor. At jobtechdev.se, one can find several APIs for fetching different types of data. All historical job ads, that were posted on their employment portal from 2010 to 2017, can be directly downloaded via this website.

5.2.1 Job Search API

The use case below evaluates the *Job search API*, that is used to fetch active ads filtered by selected criteria. The initial menu for the API is presented in Figure 19. At the very top of the page, the API key must be submitted, i.e. an identification key for authorisation of the open data user that can be applied for via an online form.

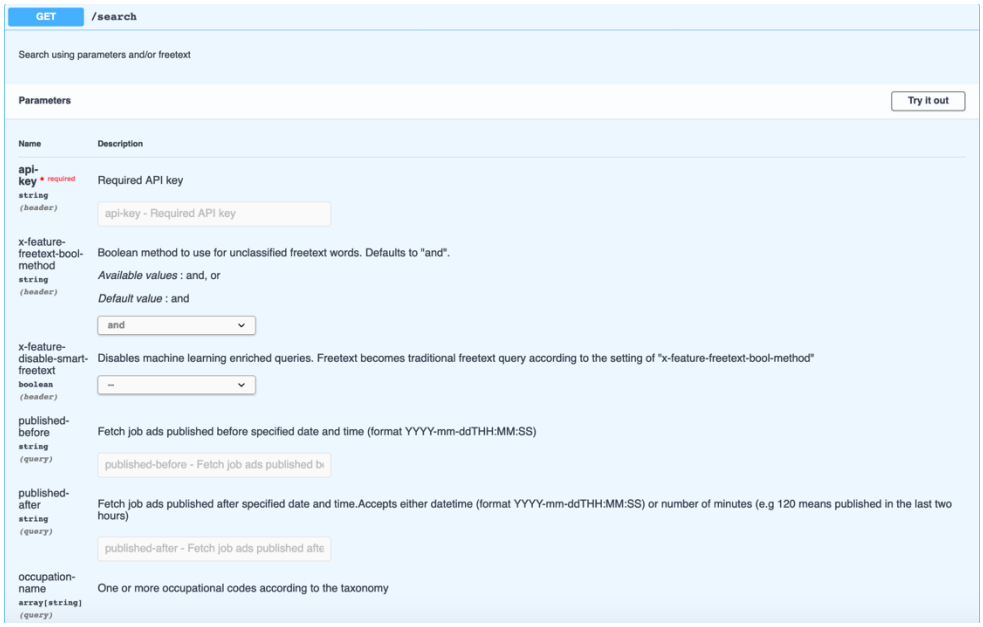


Figure 19: User interface from the Swedish Public Employment Service's open data API (screenshot taken 2020-05-13)

Next follows several different categories which can be used to filter the results. The input for these categories must follow the official taxonomy

provided by another separate API. When a successful request is made, i.e. the user has provided a valid API key and filters according to the taxonomy, the result can be downloaded as a JSON-file.

5.2.2 Use Case – Job Search API

Primary actor: Data Analyst
Scope: The Swedish Public Employment Service's Job search API for open data (subsystem scope)
Level: User goal
Stakeholders and Interests:
Data Analyst – wants to download data from the the Swedish Public Employment Service API only from certain years, and certain entry-categories.
The Swedish Public Employment Service – wants to provide open data in a user-friendly manner
Precondition: User already has an API key from The Swedish Public Employment Service and has insight into how API-requests are made.
Minimal Guarantee: System is able to detect that something went wrong and provide the user with information on how to proceed and/or ask the user to provide details.
Success Guarantee: System is able to filter the data according to the user's request and return it in JSON file.
Main Success Scenario:

1. User enters API key and request data filtered on industry
2. System accepts key and recognises the filter
3. System selects the desired data from the data base.
4. User receives the response as a JSON file.

Extensions
2a) The system does not recognise the key as a valid API key
2b) The system does not recognise the filter as valid taxonomy
3. The system fails to collect the desired data from the database
4. The response is insufficient and/or not presented in a JSON format.

5.2.3 Key Takeaways

The job search portal requires the user to gather rather deep knowledge about the taxonomy to use it correctly. The API offers many ways of sorting the datasets, but specific terminology must be followed, e.g. to fetch data from

Data and IT the parameter *apaJ_2ja_LuF* must be provided in the request url. This seemingly random value is only obtainable by doing a specific request to another API.

The data received by the portal can be downloaded as a JSON file. Although this file contained all raw data, the user base is limited to those know how to use APIs.

5.3 Downloading Datasets from Statistics Sweden

When downloading data from Statistics Sweden, their data portal *Statistikdatabasen* is used.

5.3.1 The Statistics Database – Statistikdatabasen

Data available through *Statistikdatabasen* is secondary, i.e. the data has been collected, processed, and presented by a secondary actor, not the authors of this thesis. The data is categorised in a tree-like structure, where each branch has sub-categories, as presented in Figure 20. This structure can be used to identify a suitable dataset.

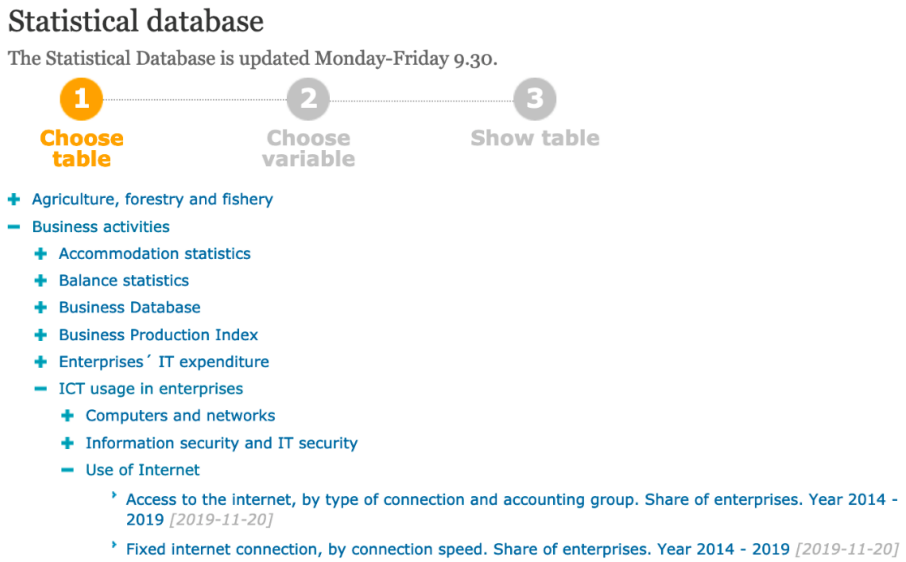


Figure 20: User interface from the Statistics Sweden Database, showing the tree-structure for identifying data (screenshot taken 2020-05-13)

The data can be sorted according to pre-determined filters given by Statistics Sweden, as shown in Figure 21. The user cannot sort the data in any other way, nor can the user combine multiple study domains for a more delicate result. Once the selection of filters is done, the result can be downloaded in several different formats.

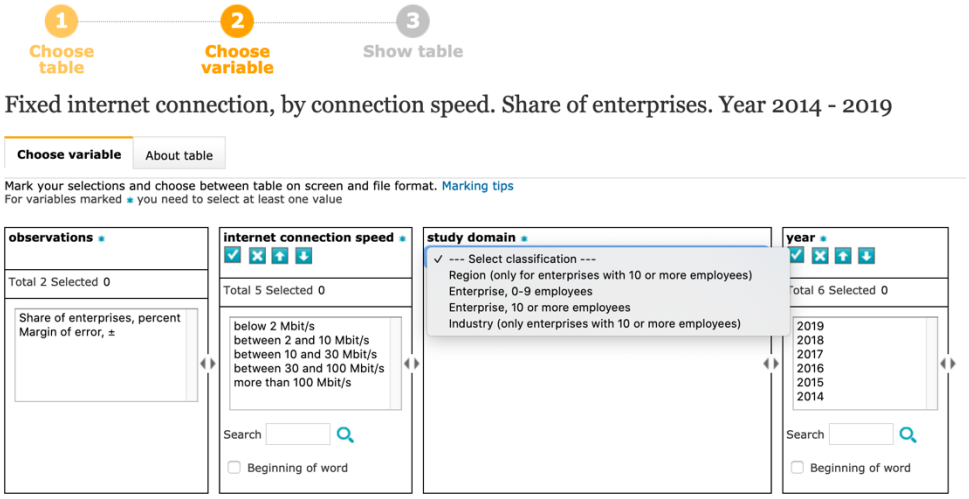


Figure 21: User interface from the Statistics Sweden Database (2), showing the pre-determined filters or study domains (screenshot taken 2020-05-13)

5.3.2 Use case – The Statistics Database

Primary actor: Data Analyst
Scope: The Statistics Database, *Statistikdatabasen* from Statistics Sweden
Level: User goal
Stakeholders and Interests:
Data Analyst – wants to download data regarding desired domain, filtered by certain critera.
Statistic Sweden – wants to provide open data in a user-friendly manner
Precondition: User must know what type of data he/she is interested in.
Minimal Guarantee: System present the datasets in an organised matter such that the user can determine whether the desired data is available.
Success Guarantee: The user can find the desired data and can filter it in according to his/her criteria.
Main Success Scenario:

1. User enters the website and is presented with categories for dataset
2. User finds a relevant category and can navigate through sub-categories until desired data table is found
3. User can filter data table according to his/her preferences and request the final data table
4. System successfully returns the data table in the file format requested by the user

Extensions

2a) User does not find a suitable category for their need.
2b) System does not provide a data table containing information needed by the user
3a) System lacks desired filters
3b) System does not support cross-filtering with desired filters
4. The response is insufficient and/or not presented in the desired format.

5.3.3 Key Takeaways

Statistikdatabasen contains a wide variety of data sets. These sets are accessible once one knows where to find them. However, if that is not the case one might have to navigate back and forth through the tree-like structure.

One can easily choose which of the pre-determined parameters to include or exclude in the result. The secondary data limits the possible modifications and combinations with other datasets.

5.4 Downloading Datasets from Oppnadata

Oppnadata arrange sources for open data and public sector information (PSI) in a structured way on their website.

5.4.1 The National Data Portal - Oppnadata

When searching for a data source on Oppnadata, one is presented with a search box as well as some categories (see Figure 22).

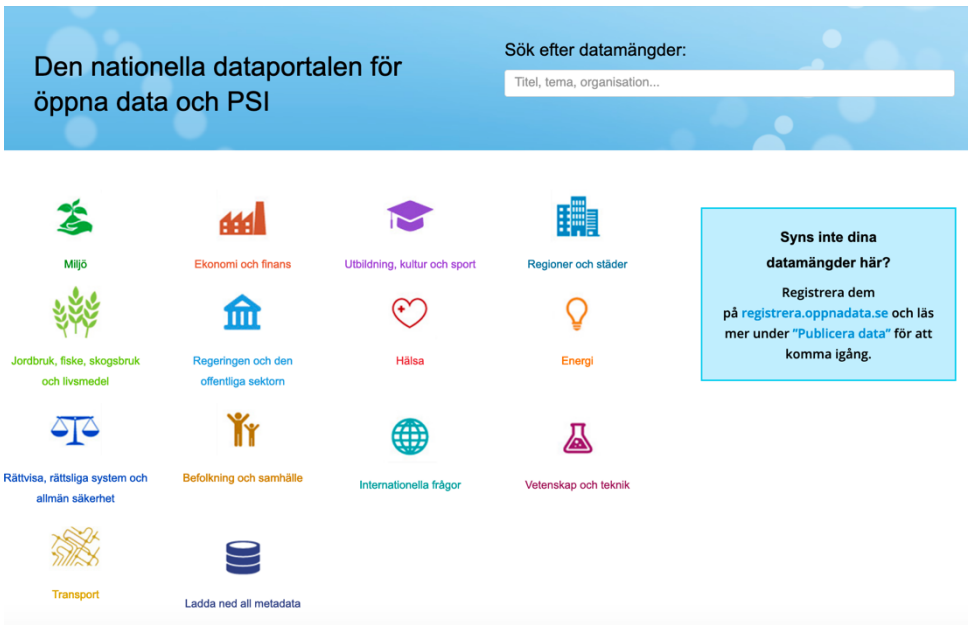


Figure 22: User interface from the Swedish national data portal, Öppnadata (screenshot taken 2020-05-13)

After choosing a category or by using the search box, one is presented with all related sources. To navigate further, one can add one of more search terms or filter the result by keywords, sources, or licenses, all provided by the website. Once a suitable dataset is found, the website provides a link to the source.

5.4.2 Use case – Oppnadata.se

Primary actor: Data Analyst
Scope: Oppnadata's data portal (oppnadata.se)
Level: User goal
Stakeholders and Interests:
Data Analyst – wants to download data regarding desired domain, filtered by certain criteria.
Oppnadata – wants to provide open data in a user-friendly manner
Precondition: User must know what type of data he/she is interested in
Minimal Guarantee: System can handle users searches and inform user if no dataset matches the search phrase.
Success Guarantee: System successfully presents user with a source containing relevant data
Main Success Scenario:

1. User enters website and is presented with data categories and a search box.
2. User searches for relevant data field and is presented with one or more data sources
3. System presents keywords which enable the user to narrow down the sources.
4. User finds a source to a desired data source.

Extensions
2a) System does not find any matching dataset for the search term
3a) System does not provide any keywords which narrow down the search result
3b) System return too many results in an unorganised matter
4. User cannot find any source for the desired data

5.4.3 Key Takeaways

Much like Statistikdatabasen, Oppnadata presents a variety of different data sets (or sources for ditto). Specific sources can be tedious to find, among the large supply. There is a possibility that a user might find themselves going back and forth from the possible categories.

6 Discussion

This chapter discusses the results of the case study, use cases and literature study and presents a summary of the potential application areas and technical barriers. Also, ethical aspects of using open data in the private sector are considered.

6.1 Open Data for Business Development in the Swedish ICT-sector (RQ1)

6.1.1 Open Data for Improved Market Insights

The case study is an example of how open data can be used for improved market insights in the ICT-sector. It thereby showcases what Magalhaes & Roseira (2017) wrote about open data providing companies with a competitive advantage through improved market insigths, as well as what Lakomaa & Kallberg (2013) found about start-ups using open data for information about potential markets. It also acts as a practical example of how open data can *complement* Sorescu's (2017) business-model innovation method of leveraging internal data to improve the customer offering and create new revenues.

6.1.1.1 Complement Internal Data with Open Data

Sorescu (2017) writes that companies can leverage internal data to improve the customer offering and create revenues. In this instance, data is used for improved insights about customer behaviour and preferences, and used either internally or sold to external parties. Since information gathered by government agencies and departments is enormous compared to any private data repositories, open data can complement internal data to make the information more rigorous.

In order to use open data for improved market insights, one needs to know where to find relevant data. Since it is mainly government agencies that provide open data, companies need to identify what government agencies possess relevant information about the market. Dataportals can be used for navigating and identifying sources of data and Statistics Sweden offers information about sector trends through the statistics published. However, as seen from the case study, agencies such as the Swedish Employment Services, that might not be top-of-mind for market analysis, can also provide informative data.

From the case study we could see that the information gained through the employment ads was more extensive then that from Statistics Sweden; the data cleaning and handling was also more complex. After data cleaning and visualisation, the insights from the employment ads turned out to be, indeed, relevant and informative, but it is important to note that this can not be guaranteed beforehand. Utilising open data is an exploratory matter. It is not until users try to extract, clean and visualise the data that potential barriers, but also the value of insights, make themselves visible.

It is also necessary to discuss whether the information about the market could have been gained through other sources than open data. The purpose of open data is to some extent that it can be *re-used*. If data is used for internal decision making, paying for survey studies or consultancy services could be a valid alternative. The time needed to identify relevant sources of data and clean and handle it, should be weighed against the cost of external consultancy services. Therefore, there might be more potential for some ICT-companies to become conversant at utilising open data and selling the products or services to other companies, rather than merely using the competence internally.

6.1.2 Open Data for New Products and Services

The case study also acts as an example of how open data can be used to create new products or services. For instance, ICT-companies could consider doing similar market analyses for other sectors and selling the information to organisations. Open data can then be used for consultancy services, which is an existing use of open data in the ICT-sector (Magalhaes & Roseira, 2017). The case study thereby also shows how open data can *complement* or *be the*

starting point for Sorescu's (2017) business model innovation of leveraging external data.

6.1.2.1 Business Model Based on Leveraging External Open Data

Sorescu (2017) writes that ICT-companies can build their business model on external data, by repackaging it and selling the relevant data to other organisations. Open data, which is free to modify and use for the development of new products and services, can therefore complement this innovation strategy. Since the market size of open data is increasing, and more and more data from the government is published as open, the products and services based on open data can be expected to improve in the coming years.

Companies in the ICT-sector are also leading-edge when it comes to handling large velocities of data. However, in some of the product and service categories presented by Magalhaes & Roseira, there were more companies from other sectors utilising open data than the ICT-sector (for instance companies in business and legal services). Though the sample size of companies from each sector in the study differs, it shows that the ICT-sector is not the only sector with competence or desire to tap into the market of new products or services created by open data. This being said, handling open data is time consuming. Companies in the ICT-sector have an opportunity to become conversant at utilising open data, which could make it difficult for companies in other sectors to compete with the speed and capability of these specialised technology companies.

6.1.2.2 Macrotrends from Micro Open Data

The case study is also an example of how microdata is used to inform about macrotrends. In the case study, individual employment ads are aggregated to give information about the macroenvironment of IT-competence demand in Sweden. Again, the information could then be resold to other organisations, or used for internal decision-making. However, though the method can be used for internal decision-making, it is a more time-consuming process than utilising data that has been pre-processed. Again, the trade-off between handling data in-house or hiring external consultancy services is relevant.

This method was discussed by Sorescu (2017) as a data-driven business model innovation, who suggests that companies tap into the opportunity use data which has been gathered to inform about a specific matter and combine it with other such data to create a dataset concerning the macro trends of the

entire data population. This method can in general be used both with internal company data and external data, and the finished product can also be used internally and sold to other organisations. Therefore, the actual application areas of open data remain the same when using this method: to create new products or services and for improved market insights, products, services or processes.

Moreover, the point of the strategy is to identify macro trends and a limitation of using open data lies in finding enough primary and big open data. Such data is for instance weather data, geodata and monetary data (or as the case study showed, it could also be employment ads). This being said, the thesis and case study were limited to Swedish sources of open data. Open data knows no borders and the maturity of open data is dispersed throughout Europe. As presented in the *"Open Data Maturity Report 2019"* released by Capgemini and the European Commission (2019a) Sweden's maturity level in publishing open data is considered lower than average in the EU. Perhaps it would be easier to find appropriate data through international data portals. Some international data may be applicable to Swedish' interests as well, if not for use in Sweden, for export.

6.1.3 Open Data for Improved Products, Services and Processes

Though this thesis has not explored the full potential of using open data to enrichen *current* products, a few example of such success appeared the when researching the background to open data. Spotify incorporated open data from the database MusicBrainz in its music streaming service and by providing the customer with relevant information related to the main service, the company improved the user experience. Moreover, Magalhaes & Roseira found that companies in the ICT-sector use open data to complement many other products, services and processes (as shown in Section 3.2).

6.1.4 Summary of Open Data for Business Development

A summary of the application areas of open data, as well as methods for utilising the resource, in the Swedish ICT-sector can be seen in Table 5 below.

Table 5: Summary of Application Areas of Open Data for Business Development in the Swedish ICT-sector

Competitive Advantage	**Improved Market Insight**	**Creation of New Products and Services**	**Improved Products, Services or Processes**
Examples	Business Intelligence, Research & Development	Advocacy, Consultancy, Apps, Data Refining & Structuring, Data Platforms, Open Data Portals, New Information Products	Process Optimisation, Product & Service Improvement
Method	Complement internal data with external open data	Leverage external open data	Extend product and service offering with external open data
Description	Retrieve and repackage open data from one or many sources for market insights Transform micro data to macro trends Utilise already processed open data	Retrieve, repackage and sell open data from one or many sources Transform micro data to macro trends Building portals for the identification of open data	Integrate open data in products and services Providing users with relevant information regarding the main service for enhanced customer experience.

6.2 Technical Barriers for Utilisation of Open Data (RQ2)

This chapter will evaluate the technical barriers encountered concerning the use of open data. The barriers can be divided into two categories:

i. Barriers encountered when searching for and extracting data (Chapter 6.2.1). These are related to the system used for such action and are covered in the use cases.
ii. Barriers encountered when processing the extracted data before the visualisation for the case analysis (Chapter 0).

6.2.1 Barriers when Searching for and Accessing Data

6.2.1.1 Access Requiring Knowledge of APIs

The accessibility varied across the encountered platforms. The tree-like structure such as Oppnadata and Statistikdatabasen provide a user-friendly and intuitive interface suitable for inexperienced users of open data. Pre-defined categories facilitate the searching process for specific information. In turn, this requires knowledge of what type of data is required to do an analysis. Especially, since the data is not necessarily gathered for market-analyses purposes and rather provided and structured in a more general-purpose format.

The API solution found at The Swedish Public Employment Service requires more technical skills for complex data handling, albeit it gives experienced users more control over the output data. The user is in control of the entire request and can specify it according to their needs, whereas the tree-like structure is limited by the systems design. Thus, it can be a great advantage of having full control on the output data concerning format and information included when developing a complex data model comprising of several datasets.

This is in line with the research by Janssen et al. (2012) about users sometimes lacking the ability to discover appropriate data. They also argue that too much focus is on distributing single datasets when the real value might come from combining different datasets. A cited interviewee said: *"Use is limited to the happy few, those who are educated and have time to explore new business opportunities."* (Janssen et al., 2012, p. 263).

Developing user-friendly platforms for retrieving large quantities of open data can be a solution for bridging that gap. Until then, the necessity of deeper computer science knowledge to extract and process open data could hinder the adoption of open data outside the ICT-sector.

6.2.2 Barriers when Processing the Data

The barriers found when processing the data mainly have to do with lack of primary data, interpretability, and consistency issues.

6.2.2.1 Lack of Primary Data

One of the main issues found when trying to merge data for the case study, was the lack of control because of only having access to processed data from Statistics Sweden. Because the data was secondary, information could only be sorted by pre-determined parameters which limited the analysis. Additionally, it was not always possible to sort the data on multiple parameters at the same time, e.g. the data could not be sorted based on geography *and* business sector. This is a practical example showcasing how limited access to original data could hinder the adoption of open data, as found in the research by Janssen et al. (2012). While Statistics Sweden does offer the primary data from their survey studies in an anonymised format, this is only available upon request for research purposes (and in limited cases for businesses). Such data costs 30-60 thousand SEK and the delivery time amounts to 1–6 months according to their website (Statistics Sweden, n.d.).

Additionally, datasets published at Statistics Sweden often contain aggregated annual data, which limits the detail of analysis to that aggregation level. This is of course useful for spotting long-term trends, but such discontinuous data points make it significantly harder to perform any more advanced statistical analyses such as predictions on a daily, weekly, or monthly scope.

Similar observations have been published in the article "*OSeMOSYS-PuLP: A StochasticModeling Framework for Long-Term Energy SystemsModeling"* (Dreier & Howells, 2019). The authors discuss benefits and drawbacks from using aggregated (secondary) data for data analysis: *"Advantages are time-efficiency for both data collection and preparation."* (Dreier & Howells, 2019, p. 2) *and "(...) drawbacks also exist such as: the dimensionality of the analysis is restricted to the aggregation level of the aggregated data; it is*

potentially impossible to trace back the aggregation to the original raw data" (Ibid.).

The API from The Swedish Public Employment Service, however, provide to the users full control over the (anonymised) raw data and allows free modification according to the users' needs. This enables opportunities of utilising the data, while it still requires work to clean and merge the data with other datasets. Interviews with open-data actors also confirms that almost real-time data, i.e. frequently updated data with minimal time-delay, provides a competitive edge (Kitsios et al., 2017)

6.2.2.2 Difficult to Interpret Data

A second issue which occurred when working with data from Statistics Sweden had to do with interpretability. Information describing the datasets was sometimes lacking, e.g. the definitions of variables or values were dubious or non-existent. Such cases undermine the quality of the conclusions that the user means to draw. This is problematic since the absence of explanations to open data was found to be an adoption barrier in the research by Janssen et al. (2012).

6.2.2.3 Inconsistent Representation and Incomplete Datasets

Lastly, fields in the datasets from Statistics Sweden were sometimes not titled or named the same way in different datasets, i.e. inconsistency. For instance, industry categorisation was inconsistent across datasets, which complicated the process of merging datasets. Additionally, in some datasets, random years of data were missing, which hindered any historical analyses from being done on the data. Again, this inconsistency has the consequence of hindering open data adoption.

This could be the result of several available portal-software frameworks, which leads to a non-uniform publication of open datasets (Kubler et al., 2018). However, as open-data portals continue to be developed, more advanced functionality are expected to be added, which puts less requirements on the user's ability to handle the inconsistencies.

6.2.3 Summary of Technical Barriers

In Table 6, a summary of the technical barriers found from the use cases and case study are presented. These are based on the quality measurements of open datasets described in Table 6.

Table 6: Summary of Encountered Technical Barriers of Swedish Open-Data Patforms

	Statistics Sweden	**The Swedish Public Employment Services API**	**Oppnadata**
Accessibility	Easy and intuitive user interface. Might have to repeat search process to find suitable data sets	Fairly complicated API which requires some time to get used to.	Easy and intuitive user interface. Might have to repeat search process to find suitable data set
Completeness	Sometimes years are missing	Complete data sets	N/A
Consistent representation	Lack of standardised variable names across different data sets. Some discrepancies regarding definitions of variables across data sets	Same name and definitions across data sets	N/A
Ease of operation	Data comes often pre-processed which makes the data handling easier but limits the possibility of more advanced calculations and connectivity to other data sets.	No limitation in terms of complete data sets and good possibilities to connect data to other data sets. Require knowledge about programming and data handling.	N/A
Interpretability	Definitions of variables and values dubious or non-existent	Definitions are either clear or possible to find explanations of.	N/A

6.3 Ethical Aspects

Open data increases transparency between government and the public. Moreover, it stimulates innovation and competition, which is valuable for society. The cost of open data is relatively affordable, which also levels the playing field for larger and smaller companies. One of the main technical barriers of open data was found to be lack of primary open data. However, in some cases the reasons for not releasing such data is for the sake of protecting GDPR and respondents in surveys. Even though this hinders the adoption of open data and innovation based on the resource, it is for valid reasons.

Another ethical aspect of open data has to do with who owns data made available by government. The data from the Swedish Public Employment Office is based on employment ad entries by Swedish companies. In the case study, the data was used to identify macro trends and no individual employment ad was examined in an extensive form. It is possible that users are unaware that their ads are becoming open data, free to use in products and services. Such aspects should be considered as more data is provided under open licenses. However, this applies more to the publishers of open data than those utilising it.

7 Conclusion

This chapter presents the conclusions to answer the two research questions asked in this thesis. Moreover, contributions, limitations and suggestions for future research are presented.

7.1 Answers to Research Questions

RQ1: How can open data be used for business development in the Swedish ICT-sector?
Open data has the potential to be used in the Swedish ICT-sector in three ways: 1) to improve market insights, 2) to improve existing processes, products and services, 3) to be used for the creation of new products and services, such as apps, data platforms, and consultancy services. The case study also showed that there is potential to leverage on open micro data that informs about a specific matter and aggregate the data to gain information about macro trends.

Overall, there is potential for ICT-companies to create revenue streams from open data. The open data market is growing in Europe, which means that the products and services based on, or complemented by, open data can be expected to improve significantly. Moreover, ICT-companies possess the capability and capacity to handle large supplies of data. By becoming conversant at handling new supplies of the resouce, these companies can expect increased market shares and revenue.

RQ2: What technical barriers are found from a user's perspective when trying to identify, access, and process open data?
The most prominent technical barrier of using open data was found in the trade-off between user friendliness when extracting open data, and the ability to then manipulate the data according to user needs. On the one hand, the datasets that were easier to obtain and needed no cleaning, had bigger limitations when it came to its applicability. These datasets consisted of

secondary data, which had already been handled and only presented single-value results, without explanatory raw data. On the other hand, datasets that offered deeper analytical opportunities, with access to the raw data, were harder to access and required technical knowledge on how to clean and process it.

Other encountered technical barriers were discontinuous and missing data points, dubious and non-existent variable definitions and inconsistency regarding variable names and categorisation.

7.2 Contributions

7.2.1 Academic Contributions

This master's thesis had the purpose to increase the knowledge of the business development potential and technical barriers of open data in the Swedish ICT-sector. This has been mainly done by answering the research questions and showcasing the potential use through an exploratory case study where open data is utilised. Second, the authors presented insights of the technical barriers from a user's perspective, regarding access points and extraction methods of open data, through the use cases.

The findings of this study contribute to existing literature by further support to the work of Magalheas and Roseira (2017) of the private sector application of open data, and Janssen et al., (2012) of the technical adoption barriers. Through the case study and use cases, the thesis also contributes to increasing the knowledge of visualisation and points of access of open data in Sweden, which was found to be an adoption barrier in Temiz's (2018) research.

7.2.2 Guidelines for Utilising Open Data

To summarise the thesis, both in technological and economical terms, the authors present a methodology for utilising open data for business development. These guidelines are not exhaustive but can perhaps encourage future research and business applications. Below follows a chronological order for gaining business insights from open data.

1. **Assess the needs**. A company begins by assessing the needs they want to resolve using open data. Perhaps they want to find new markets, produce a new product, hire new employees, gain some market information, etc.
2. **Assess the capabilities**. They also need to identify their capability of doing data analysis. Do they have internal competence to build complex data models? If not, will they hire external help? What is the time frame for the project? Do they need quick results or are they willing to wait for a more thorough analysis?
3. **Assess necessary data**. Once the needs and capabilities are assessed, the company should decide what type of data they think will generate insights related to their needs. If they plan to open a new store, data regarding infrastructure and the local demographics might be interesting; if they want to open a new factory they might be interested in data regarding rental prices, local education level and transportation.
4. **Find data sources**. The company could preferably begin its search for open data through oppnadata.se. The portal has a wide collection of open data sources and an intuitive categorisation system. The company can easily sort the sources by infrastructure, demographics, etc.
5. **Type of data**. A company that lacks data analyst competence or which need quick results from the project, should consider data sources with secondary pre-processed data (e.g. Statistikdatabasen from Statistics Sweden). This is the quickest and most convenient gathering of information. A company with higher computer science competence and with a wider time frame for the analysis should consider including primary data (e.g. such from the Swedish Public Employment Service). Though such data requires more time and knowledge to produce insights, it allows for more user control and hence, more complex computation.
6. **File format**. A company which do not wish any additional computation can chose whatever file format they wish. A company which would like to process the data should to (1) download the data in a file format acceptable for computation program (e.g. Excel, QlikSense) or (2) download the data in a format which is easily converted to the desired format.
7. **Insights generation**. Once the company has done all this, they can handle the data in whatever way they think will generate insights.

7.3 Discussion of Limitations and Future Research

This thesis used an exploratory research approach that included a sample from the open data available in Sweden. As typical for case studies, there is the risk for not covering all possible analysis dimensions. To maintain focus throughout the exploratory research approach, the analysis scope in this thesis was narrowed to a specific case study: a market analysis for an IT-consultancy company. However, the empiric and theoretic findings are potentially transferable to other contexts, with similar characteristics concerning the availability and extent of open data.

The same limitation applies to the use cases. It would be favourable to let several users interact with the systems, to ensure that all possible ways of interacting with them have been picked up on. While there is no obvious selection of people who are representatives for all users of these open-data systems, in future research one could test the systems on data analysts from the ICT-sectors.

For consideration of a representative sample of open data in Sweden, the analysis in this thesis included different types of open data sources, portals, and databases. Some of the open data was primary data, some other secondary data, and some data contained statistics, i.e. aggregated data. Because of this variation, the results can be considered as representative of what one could expect to utilise open data for.

We recommend further research on open data to confirm transferability of our methodology and findings. The conclusions of this thesis are in line with the existing literature about open data. Furthermore, the findings of this thesis confirmed what other researchers have found, in a case study scope and on a national level.

8 References

Bengtsson, L. (2020). *Supervisor at LTH*. Personal communication January-June 2020.

Bolagsverket. (2016). *API på Bolagsverket*. https://bolagsverket.se/om/oss/api-1.11642

Capgemini, & European Commission. (2015). *Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources*. https://doi.org/10.2759/328101

Capgemini, & European Commission. (2019a). *Open Data Maturity Report 2019*. https://doi.org/10.2830/073835

Capgemini, & European Commission. (2019b). *Open Data Maturity Report 2019 - Sweden*. https://www.europeandataportal.eu/sites/default/files/country-factsheet_sweden_2019.pdf

Charalabidis, Y., Alexopoulos, C., Lampoltshammer, T., Zuiderwijk, A., Janssen, M., & Ferro, E. (2018). *The world of open data: Concepts, methods, tools and experiences*. Springer. https://doi.org/10.1007/978-3-319-90850-2_1

DIGG Agency for Digital Government. (n.d.). *Nyttor och kostnader för öppna data*. Retrieved May 13, 2020, from https://oppnadata.se/nyttor-och-kostnader-for-oppna-data/

Dreier, D. (2020). *Systems Designer/Architect at Combitech*. Personal communication January-June 2020.

Dreier, D., & Howells, M. (2019). Osemosys-pulp: A stochastic modeling framework for long-term energy systems modeling. *Energies*, *12*(7). https://doi.org/10.3390/en12071382

European Data Portal. (2017). *Spotify*. [Use Case]. https://www.europeandataportal.eu/sites/default/files/use-cases/use_case_sweden_-_spotify.pdf

Forgard, J., Hellström, S., Nilsson, K., Olsson, E., & Törnström, E. (2020). Så blir klimatdata affärsnytta. *Medvind SMHI*, *1*, 2.

Guba, E. G. (1982). Criteria for Assessing the Trustworthiness of Naturalistic Inquiries. *Educational Communication and Technology*, *29*(4), 75–91. https://doi.org/10.1126/science.146.3642.347

Guetterman, T. C., & Fetters, M. D. (2018). Two Methodological Approaches to the Integration of Mixed Methods and Case Study Designs: A Systematic Review. *American Behavioral Scientist*, *62*(7), 900–918. https://doi.org/10.1177/0002764218772641

Jacobson, I. (1992). *Object Oriented Software Engineering: A Use Case Driven Approach* (1st ed.). Addison-Wesley.

Jacobson, I., Spence, I., & Kerr, B. (2011). Use-Case 2.0 ebook. *Communications of the ACM*, *December*, 55. https://doi.org/10.1145/2890778

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, *29*(4), 258–268. https://doi.org/10.1080/10580530.2012.716740

JobTech Development, & Swedish Public Employment Service. (n.d.). *Job Search API*. Retrieved February 20, 2020, from https://jobtechdev.se/docs/apis/jobsearch/

Johnson, G., Whittington, R., & Scholes, K. (2012). *Fundamentals of Strategy* (2nd ed.). Harlow : Pearson.

Jordbruksverkets statistikdatabas. (2009). *Swedish Board of Agriculture*. [Database]. http://statistik.sjv.se/PXWeb/pxweb/sv/Jordbruksverkets statistikdatabas/

Kalampokis, E., Zeginis, D., & Tarabanis, K. (2019). On modeling linked open statistical data. *Journal of Web Semantics*. https://doi.org/10.1016/j.websem.2018.11.002

Kitsios, F., Papachristos, N., & Kamariotou, M. (2017). Business models for open data ecosystem: Challenges and motivations for entrepreneurship and innovation. *Proceedings - 2017 IEEE 19th Conference on Business Informatics, CBI 2017*, *1*(September), 398–407. https://doi.org/10.1109/CBI.2017.51

Kubler, S., Robert, J., Neumaier, S., Umbrich, J., & Le Traon, Y. (2018). Comparison of metadata quality in open data portals using the Analytic Hierarchy Process. *Government Information Quarterly*, *35*(1), 13–29. https://doi.org/10.1016/j.giq.2017.11.003

Lakomaa, E., & Kallberg, J. (2013). Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs. *IEEE Access*, *1*, 558–563. https://doi.org/10.1109/ACCESS.2013.2279164

Lantmäteriet. (n.d.). *Öppna geodata*. Retrieved May 20, 2020, from https://www.lantmateriet.se/sv/Kartor-och-geografisk-information/oppna-data/

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management*, *40*(2), 133–146. https://doi.org/10.1016/S0378-7206(02)00043-5

Magalhaes, G., & Roseira, C. (2017). Open government data and the private sector: An

empirical view on business models and value creation. *Government Information Quarterly*, *September 2015*, 101248. https://doi.org/10.1016/j.giq.2017.08.004

Martin, C. (2014). Barriers to the open government data agenda: Taking a multi-level perspective. *Policy and Internet*, *6*(3), 217–240. https://doi.org/10.1002/1944-2866.POI367

Meldrum, M., & McDonald, M. (1995). *Key Marketing Concepts: The Ansoff Matrix* (pp. 121–122).

Open Knowledge Foundation. (n.d.). *Open Definition 2.1*. Retrieved March 10, 2020, from https://opendefinition.org/od/2.1/en/

Öppnadata. (n.d.). *Den nationella dataportalen för öppna data och PSI*. Retrieved February 20, 2020, from https://oppnadata.se/#noscroll

Pollack, S. (2012, March). What, Exactly, Is Business Development? *Forbes*. https://www.forbes.com/sites/scottpollack/2012/03/21/what-exactly-is-business-development/

Qlik. (2019). *QlikSense Desktop [Computer Software]*. https://www.qlik.com/us/trial/download-qlik-sense-desktop

Rojas, L. A. R., Lovelle, J. M. C., Bermúdez, G. M. T., Montenegro, C. E., de Ory, E. G., & Crespo, R. A. G. (2018). Metamodel to support decision-making from open government data. *Journal of Ambient Intelligence and Humanized Computing*, *9*(3), 553–563. https://doi.org/10.1007/s12652-016-0443-7

Schmarzo, B. (2015). *The challenge of macro analytics economics*. https://infocus.delltechnologies.com/william_schmarzo/the-challenge-of-macro-analytics-economics/

SMHI. (2014). *Meteorologisk prognosmodell (HIRLAM E11) - historiska analysdata*. [Dataset]. http://opendata-download-grid-archive.smhi.se/explore/?modeltype=3

Sorescu, A. (2017). Data-Driven Business Model Innovation. *Journal of Product Innovation Management*, *34*(5), 691–696. https://doi.org/10.1111/jpim.12398

Statistics Sweden. (n.d.). *Beställa mikrodata från SCB*. Retrieved May 13, 2020, from https://www.scb.se/vara-tjanster/bestalla-mikrodata/

Statistics Sweden Database. (1997). *Statistics Sweden*. [Database]. http://www.statistikdatabasen.scb.se/pxweb/en/ssd/

Temiz, S. (2018). *OPEN DATA AND INNOVATION ADOPTION: Lessons From Sweden*. KTH Royal Institute of Technology [Doctoral Thesis].

Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to Open

Government Data. *Government Information Quarterly*, *33*(2), 325–337. https://doi.org/10.1016/j.giq.2016.02.001

Zuiderwijk, A., & Janssen, M. (2014). Barriers and Development Directions for the Publication and Usage of Open Data: A Socio-Technical View. In M. Gascó-Hernández (Ed.), *Open Government. Public Administration and Information Technology* (4th ed.). Springer, New York, NY. https://doi.org/https://doi.org/10.1007/978-1-4614-9563-5_8