

UTILIZATION OF OPEN DATA IN BUSINESS IN HÄME REGION



Master's thesis

Hämeenlinna, Degree Programme in Business Management and Entrepreneurship

Spring 2020

Antti Viiman

Business Management and Entrepreneurship
Hämeenlinna

Author	Antti Viiman	Year 2020
Title	Avoimen datan hyödyntäminen liiketoiminnassa Kanta-Hämeessä	
Supervisor	Mikko Mäntyneva	

TIIVISTELMÄ

Tämän työn tarkoituksena oli kartoittaa avoimen datan hyödyntämisen nykytilanne liiketoiminnassa Kanta-Hämeessä, miten paikalliset yritykset voisivat hyödyntää sitä paremmin ja millaiselle datalle on kysyntää. Oletus oli, että avointa dataa ei alueella suuressa määrin hyödynnetä. Työn toimeksiantajana toimi AvoinHäme-hanke, jota toteuttaa HAMK Smart tutkimusyksikkö. Hankkeessa luodaan olosuhteita avoimen datan hyödyntämiseen kasvattamalla eri toimijoiden tietoa ja taitoja, sekä esittelemällä avoimen datan hyötyjä esimerkein, pilotein ja kokeiluina.

Teoreettisessa viitekehyksessä esitellään avoimen datan konsepti, perustellaan syitä datojen avaamiselle, kuvaillaan datan jakelukanavat, esitellään miten avointa dataa voi hyödyntää liiketoiminnassa, kuvaillaan sen rooli BI-toiminnassa, esitellään älykkäiden kaupunkien konsepti ja kuvaillaan avoimen datan nykytilanne Suomessa ja Kanta-Hämeessä.

Tutkimus tehtiin teemahaastatteluina kuudelle paikalliselle yritykselle eri toimialoilta, sekä yhdelle avoimen lähdekoodin asiantuntijalle. Tuloksena nähtiin, että vaikei avointa dataa hyödynnetä paljoa, esimerkkejä siitä on olemassa. Näyttää kuitenkin, että avoimen datan osuus näiden yritysten liiketoiminnassa on hyvin pieni. Suurin este avoimen datan läpimurtoon näyttäisi olevan yritysten tietämättömyys siitä, miten sitä voisi muuttaa kannattavaksi liiketoiminnaksi. Tärkeänä nähtiin myös teknisen osaamisen riittävyys. Yleensä aloitteet tulevat yksittäisiltä toimijoilta, eikä avoin data ole niinkään osa yritysten strategiaa. Datan saatavuus on kansallisella tasolla hyvä, mutta alueellisille data-avauksille on vielä kysyntää.

Avainsanat avoin data, datan hyödyntäminen liiketoiminnassa, business intelligence, älykäs kaupunki

Sivut 49 sivua, joista liitteitä 1 sivu

Business Management and Entrepreneurship
Hämeenlinna

Author	Antti Viiman	Year 2020
Subject	Utilization of Open Data in Business in Häme Region	
Supervisor	Mikko Mäntyneva	

ABSTRACT

The aim of this study was to find out the current situation of Open Data utilization in business in Häme region, how the local companies could utilize Open Data better and for what kind of data openings there is a demand. The base assumption was that Open Data is not exploited on a large scale within the region. The thesis was done for OpenHäme project conducted by the HAMK Smart research unit in Häme University of Applied Sciences. The project aims in creating conditions for utilizing Open Data in business by increasing the skills and knowledge of different actors and presenting the advantages of Open Data by examples, pilots and experiments.

The theoretical framework presents the concept of Open Data, gives reasonings for data openings, describes the distribution channels of Open Data, introduces how Open Data can be utilized in business, describes its role in Business Intelligence, presents the concept of evolving smart cities and describes the current situation of Open Data in Finland and in Häme.

The research was done by theme interviews with six local companies from various fields and one local open source expert. It was found that even if Open Data is not utilized a lot, some examples exist. It seems however that Open Data plays a rather minor part in these companies' business. The greatest barrier for Open Data breakthrough is the unawareness of companies how to turn it into profitable business. The required technical skills for utilizing data was also noted important. The initiatives come mostly from single individuals within the companies instead of its being an important part of company's strategy. The availability of Open Data is good nationally, but there is still some demand for more regional data openings.

Keywords open data, data utilization in business, business intelligence, smart city

Pages 49 pages including appendices 1 page

CONTENTS

1	INTRODUCTION	1
1.1	Background of the Thesis	1
1.2	Structure of the Thesis	3
1.3	Research Objective and Limitations.....	3
2	OPEN DATA.....	4
2.1	Open Data as a Concept.....	5
2.1.1	Reasoning Data Openings.....	6
2.1.2	Open Data Portals.....	9
2.2	Utilization of Open Data.....	13
2.2.1	Open Data Business Aspect.....	13
2.2.2	Business Intelligence	17
2.2.3	Smart Cities.....	18
2.3	Open Data in Finland.....	19
2.3.1	Open Government Partnership	23
2.3.2	The Six City Strategy	26
2.3.3	Open Data in Häme Region	27
3	METHODOLOGICAL CHOICES	29
3.1	Research Questions.....	29
3.2	Research Methodology	29
3.2.1	Qualitative Research.....	29
3.2.2	Data Collection Method: Thematic Interviews	30
3.3	Research Quality	31
3.4	Research Process.....	33
3.4.1	Selection of the Interviewees.....	34
3.4.2	Structure of the Interview	36
3.4.3	Implementation of the Interviews	38
4	RESEARCH FINDINGS	40
4.1	Open Data Utilization – Current Situation	40
4.2	New Open Data Utilization Possibilities.....	42
4.3	New Data Openings in Häme Region	43
5	CONCLUSION AND DISCUSSION	45
	REFERENCES.....	49

Appendices

Appendix 1 Thematic Interview Structure

1 INTRODUCTION

Open Data is a massive yet still largely untapped resource for business purposes. It can be used to build new companies and/or business models that deliver improved products or services, offer in-depth investment tools, provide better energy efficiency or lead to mobile apps offering consumer services. By utilizing Open Data companies can learn what their customers think about them, spot what is trending in the market and get support on choosing their business partners. Open Data is described as the world's greatest free resource. (Gurin, 2014, p. 1)

Public organizations have not yet in a very large scale nor consistently opened their data in Häme region. Open Data is a rather new phenomenon but has in recent years raised as an interesting topic in public discussion. A special interest is placed on how Open Data can be utilized in business.

1.1 Background of the Thesis

Häme region (Kanta-Häme in Finnish, sometimes translated in English as Tavastia proper) in this thesis refers to a region in Finland consisting of three cities, Forssa, Hämeenlinna and Riihimäki, and eight smaller municipalities. It is a 5 700 square kilometer wide region having approximately 175 000 inhabitants. The Helsinki metropolitan region locates about one-hour car drive to South from Häme. Housing, education, entrepreneurship and a versatile cultural life are valued in Häme. (Regional Council of Häme, n.d.)

Open Data has not been much utilized in local business yet. Few active actors have been discussing on Open Data and its possibilities for some years but the knowledge on the topic has not yet spread wide. Challenges in companies arise from the lack of knowledge and skills, active co-operation and resources. Some data openings have been made in the region but there is still not much Open Data available yet. The accessibility to Open Data is the basic element for utilizing it in business. (Kukkamäki, 2017, pp. 2–3)

In bigger cities in Finland like Helsinki and Tampere there are much longer traditions in the usage of Open Data. There is a lot to learn from them and actions to be transformed to Häme region as well. Open Data catalogues like Helsinki Region Infoshare and one in the Tampere City website have been created. Those services aim in creating, opening, sharing and utilizing data. They include examples of how the data is used in different applications, visualizations and services. (Kukkamäki, 2017, p. 4)

Open Data increases democracy and transparency in the public sector. Cities and other public organizations should recognize Open Data as an opportunity. It can however be difficult if no concrete actions are made for data openings. To get the best results there should be simultaneous actions within these organizations from the top management level and from the experts working in the field. In between these two operates the IT-department, which plays the key role in developing and utilizing the open data in the organization. (Kukkamäki, 2017, p. 5)

Today, many companies exploit some data in their business. Data, being it open or closed, is often very difficult and complex to read and interpret. Therefore, it is vital to make simplifying visualizations of the data to get it in a usable form. By the help of data visualization tools, it is possible to transform large quantitative data sets into useful information. Information gained from the visualized data can then be used to support decision making and business intelligence.

Not a lot of previous research has done on the business utilization aspect of Open Data in Finland. In his doctoral dissertation "Benefits from Open Data: Barriers to Supply and Demand of Open Data in Private Organizations", Antti Herala (2018, p. 54) states, that "While there is a definite interest towards open data in a wide variety of industries, the usage of open data is low, because of the usability, accessibility, and discoverability of open data." He conducted a survey with responses from 45 Finnish companies and found out that many of them have difficulties even in exploiting their internal data and the use of external data seems to be somewhere in distant future. Software companies were especially interested in utilizing this free resource for application development and data analytics, but even they had difficulties to transform it into major business.

This thesis was done for the OpenHäme-project conducted by HAMK Smart research unit in Häme University of Applied Sciences. It is a project funded by the European Regional Development Fund lasting from 1 May 2017 till 30 April 2020. The project aims in creating conditions for utilizing Open Data in business by increasing the skills and knowledge of different actors, creating an active Open Data community and presenting the advantages of Open Data by examples, pilots and experiments. Project is targeted for active and development-driven companies in the region, who could utilize Open Data in their business actions in the future. Cities, municipalities, business development companies and other public organizations are in the data opener role. The specific target sectors of the project are travelling and event production, traffic and logistics and bioeconomy. (AvoimHäme, n.d.)

1.2 Structure of the Thesis

Thesis starts with a literature review divided in three themes. First, the concept of Open Data is explained. This includes defining Open Data, reasoning why organizations should open their data and presenting the sharing channels for Open Data, the data portals. Second, Open Data is viewed from the business perspective: how Open Data can be utilized in business, how it can be used for Business Intelligence purposes and what is its role in the evolving Smart Cities. Finally, the current situation of Open Data in Finland and in Häme region is presented.

The methodological choices are presented next. This thesis is based on three research questions: what the current situation with Open Data utilization in business in Häme region is, how the local companies could utilize Open Data better and for what kind of Open Data there is a demand. The research was implemented as theme interviews with local companies.

The findings of the research are presented in chapter four. The research was divided in three main themes based on the research questions. First, the current situation in Open Data utilization of the companies involved is conducted. Next, their views on the future are presented. Finally, the data demand of these companies is examined.

The final chapter concludes and combines the theoretical framework and the findings from the empirical research. Recommendations and improvement ideas based on the research are also presented.

1.3 Research Objective and Limitations

This is a qualitative research. The aim of the thesis was to find out how the local companies utilize Open Data in their business currently and how they could do it better in the future. Data collection for the research was conducted as theme interviews with six companies and one open solutions expert. The results will be published and ideally the companies in the region will gain more knowledge on Open Data and data visualization. The increasing demand for Open Data may then eventually lead for more data openings as well. The research findings will also benefit HAMK Smart research unit and external Open Data actives in their future development actions.

This thesis is trying to find answers if and how Open Data is utilized in business in Häme region. The aim is not to get an overall coverage of the situation in the region. The general assumption based on discussions with local Open Data actives and prior researches is that companies utilizing Open Data are still a vast minority.

2 OPEN DATA

Data is in a central role in business today. Companies, individuals and internet-connected devices are generating huge amounts of data. These data masses are growing in an exponential rate. In terms of volume, variety and velocity the world of data is getting more and more competitive every day. With the digital revolution the ability to utilize available data masses has become essential for many businesses. These data masses requiring specific technology and analytical method for transforming it into value can be defined as Big Data. The capability for using, reusing and shaping data by various ways without diminishing its value is essential. To get the most out of data it must be made available for the party in a right form at a right time. By having this Open Data that is free and accessible for everyone, it is possible for the interested parties to exploit the data and add value to it. This will lead into new ways of utilizing the data and creating innovations. Opening essential public data for everyone will also save a lot of time and money and be of help for decision making in many sectors. Big Data and Open Data are the key concepts changing the world in the data revolution. (Monino & Sedkaoui, 2016, p. 23)

Open Data is a valuable resource for data-driven innovation. Without data, it is impossible to create information and without information, it is impossible to create new knowledge. This new knowledge is needed in supporting the development in the contemporary society and Open Data can be seen as a critical cornerstone in this development. There are already successful case examples on Open Data initiatives leading to competitive advantage and increased profitability in the private sector and improved efficiency and other gains in the private sector. Open Data is a great free resource and the continuous evolution of technological innovations is generating special conditions for the effective use of it. Yet, however, either the private or public organizations have not been able to fully utilize this potential. (Biedenbach & Boström, 2018, pp. 8–10)

People around the world exploit Open Data daily by using mobile apps using for example GPS information, weather data or data from private organizations. Business ventures based on Open Data are created to make our everyday life easier. The data revolution is rapidly leading us into new territory. Huge databases are becoming public as computers get more powerful, memory storage gets cheaper and there is an exponential growth in digital information. This leads to huge opportunities for Open Data to change the world. (Gurin, 2014, p. 2)

2.1 Open Data as a Concept

Open Knowledge International (n.d.) states that “Open Data and content can be freely used, modified, and shared by anyone for any purpose.” Open Data is digital data that is free and accessible for everyone. Reuse and redistribution must be permitted without restrictions. There are three criteria for data being Open Data:

- **Availability and access:** Data must be available in full. It must be free or for a reasonable reproduction cost. It must be available in a convenient and modifiable form and preferably downloadable from internet.
- **Re-use and Redistribution:** Re-use and redistribution must be permitted including the intermixing with other datasets.
- **Universal Participation:** There should be no restrictions for using the data.

(Dietrich, Gray, McNamara, Poikola, Tait & Zijlstra, n.d.)

Gurin (2014, pp. 9–13) describes Open Data as “accessible public data that people, companies, and organizations can use to launch new ventures, analyse patterns and trends, make data-driven decisions, and solve complex problems.” It is very different from Big Data, even if these two overlap. They differ especially in their philosophy, goals and practice. The mission of Open Data is to provide free, open and transparent data to transform the way of doing business, running a government or managing all kinds of transactions. The people behind the Open Data philosophy are a diverse group. Big Data, on the other hand, comes generally from passive data sources that generate data without a purpose, direction or meaning. It is often kept private for business or security reasons. Open Data is the opposite: public and purposeful.

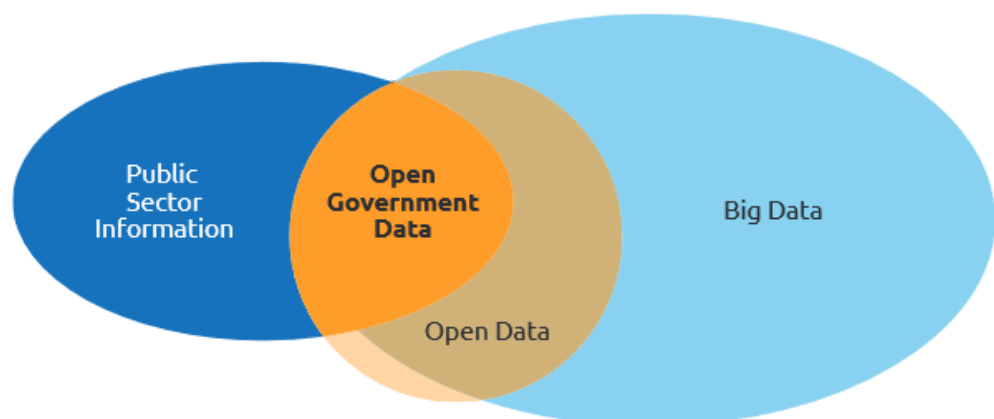


Figure 1. Overlapping of Open Data, Big Data and Public Sector Information. (Carrara, Chan, Fischer & van Steenberg, 2015, p. 22)

When Public Sector Information (PSI), Big Data and Open Data overlap, the results can be powerful. An example of this are the large public government data sets like weather or health data. Open Data may overlap with Big Data also without PSI. Large data sets from scientific research or social media for example fit in this category. Open Data not overlapping with Big Data is often open business data. (Carrara et al., 2015, p. 22)

Sir Tim Berners-Lee (2010) has introduced a following concept of five-star linked open data:

- ★ Available on the web (whatever format) but with an open license, to be Open Data
- ★★ Available as machine-readable structured data (e.g. Excel instead of image scan of a table)
- ★★★ as two stars plus non-proprietary format (e.g. CSV instead of Excel)
- ★★★★ All the above plus, use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
- ★★★★★ All the above, plus: Link your data to other people's data to provide context

Open Data can also be sorted in three categories depending on their sources: public data, scientific research data and private data. Public data includes all data and information collected by the public sector. Scientific research data comes from publicly funded research. Private data is the data from the private sector that can be made public with right incentives and privacy protections. Common for any Open Data is that it is made available for access, exploitation and reuse for any interested actor like companies, scientists and private citizens. (Monino & Sedkaoui, 2016, p. 24)

2.1.1 Reasoning Data Openings

Open Data is a largely untapped resource. Governments and other public organizations produce a great amount of data that is public by law and it could be made more open and available for others to use. It could be of great value and many people and organizations can benefit on it. Opening data will also increase democracy and transparency. (Dietrich et al., n.d.)

In addition to public transparency many other gains are expected on data openness. It can increase the citizen participation and respond to the crisis of confidence towards politicians and institutions. On the other hand, it

can create economic value through the development of new activities based on Open Data. (Chingard, 2013)

Governments might have different motivators for Open Data initiatives. Two key motivations can be identified. First, these initiatives create openness and increase transparency in the name of democracy and freedom of information. In the US for example, the central portal for datasets, data.gov portal, is based on “The Open Government Directive” which consist of three aspects: transparency, citizen participation and collaboration to strengthen democracy. The other motivator is based on the economic point of view. Governments hold a lot of non-sensitive data that could have great economic value for businesses and individuals. This data can be for example meteorological data, traffic data, geographical data, financial data or event calendars of a certain area. If this data is made public, individuals and businesses can create value-added products and services for either commercial purposes or public good. This can lead to improvement in government services, creation of jobs and opportunities, and eventually contribute to economic growth. Whatever the reasons for government data openings are, the focus should be based on the actual usage of the Open Data portal for its intended purposes. Main thing is to find out ways for the citizens and businesses to participate and collaborate in utilizing the data portal. (Chan, 2013, pp. 1890–1891)

Previously the public sector had the role of a data gatekeeper. In the Open Data philosophy, government should take the role of information provider. This way the cooperation between the public sector, companies and people can be developed by encouraging transparency, developing new applications and getting individuals involved in data enrichment. Internet and smart objects have also created a data universe in which data can be aggregated, cross-referenced, analyzed and reused. This will add more value for the data and will make it much more beneficial than it would be in isolated data storages. (Monino & Sedkaoui, 2016, pp. 26–27)

Governments receive a lot of requests for information from citizens, private sector and other agencies in the public sector. Handling these requests cost time and money. By opening the data for free these handling costs of information transfers can be eliminated and the different actor can benefit from one another. The public agencies can also benefit on the products and services created by the data enrichers. Many governments have witnessed a huge increase in the demand for information after opening their data. (Carrara et al., 2015, p. 100) Koski (2015, p. 14) presents that by opening public transport data, San Francisco has been able to reduce its number of incoming telephone queries by about 22% and generate annual savings of more than 1 million dollars.

By opening data as much as possible, it will become more available to be utilized within the interest of all. Monino & Sedkaoui (2016, p. 32) present that according to several studies the open exploitable data could generate

millions of dollars. In the Open Data movement open access and exploitation of data for all and by all is targeted. Open Data should be acknowledged as an opportunity and a long-term process for all actors.

Public sector is one of the most likely sectors to benefit from the use of Open Data resources. Higher data revenues will be followed by the establishment of new data-based companies. The meteorological data being open for long in many countries, the annual tax revenue from the meteorological sector has risen by hundreds of millions of euros as a result. The public sector's tax revenue is also expected to increase because of the opening of public administrations internal data resources across borders as international openness of taxation data can reduce international tax evasion. Education and research also highly benefit from data openings. It improves the preconditions for studying various social and economic phenomena. By sharing their data, major organizations can avoid duplicate work. In education the same impact is made by making existing data and teaching materials openly available. Benefits can be seen on a national level as well. The economic potential of Open Data is remarkable. Having enough effects on different sectors, the growth of new and existing companies may result as growth on a national or even global economic level. (Koski, 2015, pp. 13–15)

The Open Data movement is significant and potentially powerful emerging force. The general idea of Open Data is to make local, regional and national data available in a form where it can be directly manipulated. This means at least 3-star data presented in the Berners-Lee's model. Data may be edited using software tools for example for the purposes of cross-tabulation, mapping or visualization. (Gurstein, 2011)

Open Data must be published in a useful way for it to be valuable for reuse. In ideal situation the data is available in Open Data platforms that provide a central and responsive entry point where users can search for it. This way people will have an easier access to data and a possibility to find all the relevant information they need. Often all the required information does not lay in one file. As data is coming from many sources and organizations have their own publication policies the data can be very complex and heterogenous. By the help of software tools, the data can be then integrated. After the integration it is possible to explore the data, extract the relevant information and, if wanted, visualize the results. Currently two approaches in publishing styles can be identified in Open Data platforms: human-readable format and machine-readable format. In the first mentioned approach the complex raw data is transformed in a more comprehensive form by for example aggregated values, visualization or other simplifications. If this kind of report contains all the information, it can be very useful for the user. However, generally data is most beneficial if it is published in a machine-readable form without reprocessing it. In that case users can customize the data for their own needs by software tools.

This way the data reuse is also possible without extra processing. (Braunschweig, Eberius, Thiele & Lehner, 2012)

There are direct and indirect benefits in making governmental data available for free. The direct benefits can be measured in money in forms of revenues and Gross Value Added (GVA), cost savings and in the number of people needed in product or service creation process. The direct re-users, data enrichers, are mainly the ones gaining these benefits. The indirect benefits on the other hand are mostly enjoyed by the users of the products created by the direct re-users. These benefits can be divided into economic, political and social indirect benefits. Economic benefits may appear as new job potential, new products and services, knowledge economy growth, increased efficiency in public services and growth of related markets. Political benefits include the increased transparency and citizen participation, political awareness of people and access to information they are justified for. On social level this means social inclusion, empowerment and support to personal decision-making capabilities. (Carrara et al. 2015, pp. 35–36)

Opening data was a clear megatrend in the 2010s. To maximize the effects of data in gaining as many people as possible, it should be spread as wide as possible in an effective way. More and more data openings are done through Application Programming Interfaces (APIs). Even if building the APIs are often more expensive than traditional file sharing, it has many gains: real-time data streams, savings in telecommunicating costs, more precise information on data usage, publishing data with multiple metadata's and in many formats and the better possibilities for tailoring the data. More and more devices are connected to internet and exploit API's. Devices using Open Data through these API's may include phones, tablets, cars, light-systems and household devices. (Koski, Honkanen, Luukkonen, Pajarinen & Ropponen, 2017, pp. 40–41)

People increasingly often rely on Open Data, even if being unaware of it. Today's technology allows us to navigate, check the weather forecast or for example search for a restaurant through mobile applications built on Open Data. The easier access to Open Government Data bridges the gap between government and citizens in terms of information. The huge number of basic data of the governments can be transformed into economic and social value to the society. Open data strengthens the citizen right as people get easier access for free for the information they are legitimated for. (Carrara et al., 2015, p. 24)

2.1.2 Open Data Portals

For making the data accessible and easy to find it must be gathered in a common place. The number of Open Data portals is rising, particularly amongst public institutions. In 2017, from the EU18+ countries Liechtenstein was the only nation not having a national Open Data portal.

In the 2015-launched European Data Portal there is an open access for datasets from 34 different countries. The simple purpose of these portals is to enable the Open Data to be found easily. Typically, in the portals the user can use a search function to find some specific data on a certain topic or from a specific organization or browse through the dataset collection of the portal. For the data publishers, portals are excellent tools for facilitating data discovery to metadata and version management. They also provide an opportunity for the publishers to follow how many times their data has been accessed and to assess for what kind of data there is a demand. For governments, the portals are being increasingly used for communication purposes with the citizens. (Simperl & Walker, 2017, p. 5)

It is necessary to develop new ways of publishing, managing and using Open Data as it is getting more common and the volumes are rising rapidly. There is a risk called “virtuous data dumps” when just publishing the data to a portal. The portals have been a necessary step in the emergence of Open Data, but now more and more emphasis must be put in finding out ways how to get it in a broader and more impactful use. For citizens, it is important to know what can be done with the data. For data professionals the greatest challenges are in finding and making sense of the complex data. The portal owners must put focus on understanding these scenarios, listening to user experiences and developing the best tools for the users to utilize data. As Open Data utilizers are becoming a wider and more diverse group, changes in this traditional supply-led approach are required to better meet and provoke the demand. (Simperl & Walker, 2017, p. 6) A European Data Portal publication (Simperl & Walker, 2017, pp. 7–24) presents ten ways to do this:

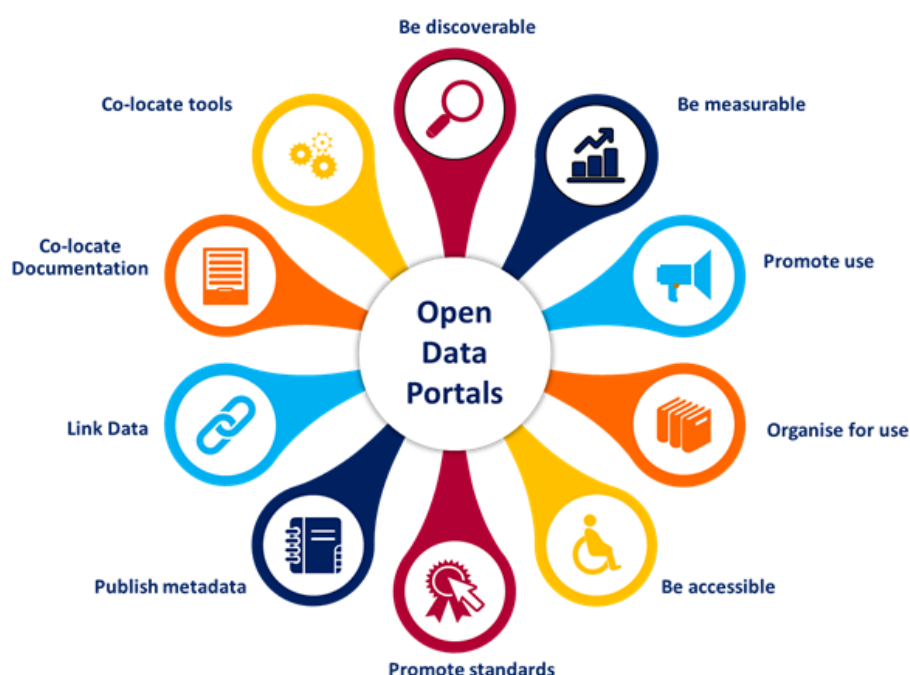


Figure 2. Ten Ways in which Open Data Portals Must Evolve for Sustainability and Added Value (Simperl & Walker, 2017, p. 7)

1. **Organize for use:** Open Data portals are not often designed with the user experience on mind. Portals are supposed to help people find and use data. Key factor in this development is the analyzing of the user behavior, offering sensible data previews for the users, recommending related datasets and providing comprehensive data descriptions, reviews and incentives.
2. **Promote use:** Impact stories and examples should be promoted for the data utilizers. Enabling and promoting opportunities for questions and discussions will help the creation of Open Data communities. The most critical thing could be the theming of datasets by their potential usage, rather than publication, categories.
3. **Be discoverable:** Often users do not know where to look for the data. Open Data portals must be easily discoverable and linked with each other. Data portals should be enabled or even optimized for search engines.
4. **Publish metadata.** For findability and cataloguing, accurate metadata is vital. There are three dimensions for the user in assessing the data: relevance, usability and quality.

Assess	Information needed about
<i>Relevance</i>	Context, coverage, original purpose, granularity, summary, time frame
<i>Usability</i>	Labeling, documentation, licence, access, machine-readability, language used, format, schema, ability to share
<i>Quality</i>	Collection methods, provenance, consistency of formatting/labeling, completeness, what has been excluded

Figure 3. The Metadata Users Consider Useful to Make Sense of Data (Simperl & Walker, 2017, p. 14)

5. **Promote standards:** To ensure interoperability, some standards must be adopted. A standard is an agreed way of doing something. Data portals need to define and promote standards in the way their data sets are named and their attributes and connections to other concepts.
6. **Co-locate documentation:** The length and technicality of supporting material can make it useless. Instead, it should be accessible from within the dataset and it should be context-sensitive for users to directly access the desired piece of information. This means there is no need to search for the documentation, which speeds up the access to the relevant material. Open government data publishers and portal hosts must re-think their ownership models to allow for a broader community to contribute in secondary data publishing. Detailed information of datasets helps the data practitioners to decide whether the data is relevant for their tasks. With co-locating tools a wider range of users and re-users can be engaged with.

7. **Link data:** Datasets are often used with each other so linking datasets enhance their value. The most popular combinations of datasets comprise population statistics, environmental datasets and regions and cities data. The EU and national governments are encouraging the use of this higher starred (3 stars or more) Open Data and Google is also indexing some types of Open Data. This will create demand for specialist services that can manage with the interlinking of repositories and repository collections accurately.
8. **Be measurable:** Assessing how data portals are meeting the user needs is crucial. They can be measured with at least two types of metrics: usage (for publishers) and quality (for users). For usage the most common metric is the number of downloads, which does not however tell for what purpose the Open Data is used for. In the future the increased use of digital object identifiers might give more detailed information on which data is considered of high value. Data quality metrics include completeness, correctness, timeliness and the fitness of use.
9. **Co-locate tools:** For users the standard process in utilizing Open Data is that they select their datasets from a portal and then use appropriate tools to process it further. This creates barriers as not all the potential users have the required skills to handle and understand the raw data. Careful co-locating of the documentation can have a huge impact on individuals' ability to explore a dataset and decide on its relevance. Many portals include visualization tools for meeting the user needs better. The Finnish Opendata.fi is a portal for both Open Data and interoperability tools.
10. **Be accessible:** Accessible data portals offer both options for data handling: Application Programme Interfaces (API) and options for manual processing, such as comma separated value files, to ensure all user needs are being met.

Open Data portals should maintain site analytics to get information on who uses the data and which data sets are downloaded. This kind of information is vital for example in calculating the effect of Open Data to cost savings. There should also be a feedback mechanism. It enables the interaction between supply and demand. Through a feedback mechanism it is also possible to make data requests and improvement ideas. This will lead to better quality public data. (Carrara et al., 2015, pp. 100–101)

There are thousands of data portals available world-wide and often cities and nations have one of their own. In Finland, majority of Open Data is gathered in the opendata.fi portal. It was published in 2014 as a part of the Open Knowledge Program of the Ministry of Finance. In the portal it is possible to search, publish and manage data. It also automatically harvests the metadata from Helsinki Region Infoshare (Open Data portal of the capital area in Finland) and the National Land Survey of Finland's spatial data. The European Data Portal gathers the Open Data from the national portals of its member countries. (Koski et al., 2017, pp. 52–53)

2.2 Utilization of Open Data

Open Data is used for various purposes. Investors utilize it to search for the best investment opportunities, established companies use it to understand their brands reputation and develop data-driven marketing strategies, start-ups create new services around weather, housing, transportation and more to benefit the public. Governments release Open Data to enhance trust in government and to foster business innovation. In medical research it is used for finding treatments and cures for diseases. The media uses Open Data for data journalism. (Gurin, 2014, pp. 2–3)

2.2.1 Open Data Business Aspect

Open Data is a great resource to support decision making in business. Decision making processes in companies are getting more and more knowledge based. New companies and products are born based on Open Data. In most cases Open Data is utilized as a part of a service or product, it supports decision making or it helps developing the automatization of processes. In the best case, data-based innovations lead to growth and productivity improvements. On the other hand, companies produce so much data that they only have resources to utilize a minor part of it. By opening that data companies can benefit the surrounding community. More and more Open Data is also used for teaching Artificial Intelligence. (Honkanen, 2019)

Due the policy changes around Open Data and the easier access to data sets more and more companies have started to use Open Data as one of their raw materials. This together with the rapid development of ICT brings huge potential for the re-use of Open Data. It is critical to recognize what kind of Open Data there is available, where it can create value and what kind of products and services can be produced out of it. The Data Value Chain opens the utilization process of the data. (Carrara et al., 2015, p. 28)

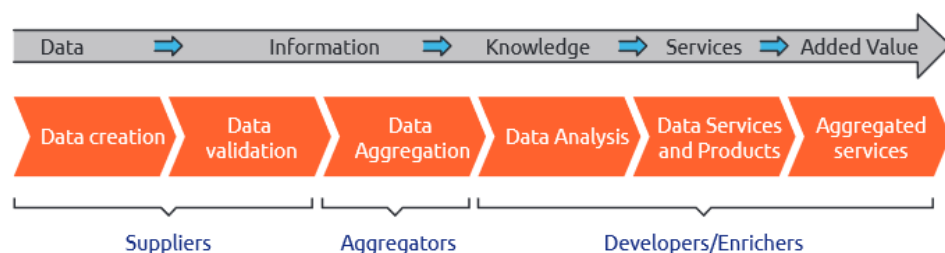


Figure 4. Data Value Chain Archetypes. (Carrara et al., 2015, p. 29)

Gurin (2014, p. 15) presents consulting company Deloitte's five Open Data business "archetypes":

- **Suppliers:** publish their data as Open Data to increase customer loyalty and improve their reputation. They create the data and validate it.

- **Aggregators:** collect and analyze Open Data and sell their insights or make profit on the data in other ways. They may combine data or make visualizations.
- **Developers:** use Open Data as a resource to design, build and sell the data to the customers in attractive and informative ways through for example web or mobile phone applications.
- **Enrichers:** improve their existing products and services by using Open Data to gain new or better insights to their products or services. They are typically large companies that for example use demographic data to understand their customers better.
- **Enablers:** make Open Data easy-to-use and charge for that.

Gurin (2014, pp. 15–16) also adds two other categories to the list:

- **Better Business through Open Data:** getting added value and improvements from Open Data in many aspects of consumer society like healthcare, energy, education, finance and transportation sectors. Service providers using Open Data “behind the scenes”.
- **Open Data Pure Plays:** companies or industries that would not exist without Open Data. Start-ups build based on Open Data.

Companies utilizing Open Data and Big Data in their innovation activities create significantly more product innovations compared to those who do not. In 2012-2014 the revenue of Finnish ICT-companies utilizing this data grew 17 per cent more than their competitors not exploiting it. (Koski et al., 2017, pp. 28–29)

Koski et al. (2017, p. 19) present some expected impacts and indicators of Open Data utilization for companies gathered in the table below.

Expected impact	Indicator
New products and services	The number of new products and services based fully or partially on Open Data. The share of revenue of these products and services.
New companies	The number of new companies exploiting Open Data.
Growth	Growth in revenue. Changes in the number of employees.
Productivity	Added value per employee. Working hours done.

Table 1. Expected impacts and indicators for companies on Open Data utilization.

Larger companies (250 employees or more) are more eager to utilize Open Data in their innovation activities than the smaller ones (excluding the newborn companies based on Open Data). The most common fields where

Open Data is used for product related innovations include science and technology, financing and insurance sector, energy and water supply and environmental management. Companies producing digital products or services exploit Open Data more than others. Often companies using Open Data also exploit some Big Data. One correlation is also that those companies having collaboration with public sector in their innovation activities are also the most eager ones to utilize Open Data. Radical innovations, which are new for the whole market, are also more common within companies using Open Data. (Koski et al., 2017, pp. 23–24)

To use Open Data effectively, it must be ensured that there are opportunities and resources available for translating the data into useful outcomes for the widest possible range of users. Things to consider include the cost and availability of internet access, language of the data, the technical or professional requirements for interpreting and making use of the data, and the availability of training in data use and visualization. The most likely parties getting benefits on Open Data are those who have the most resources to make use of the Open Data. Private sector has the means and interest in directly translating available data into new commercial products, or services or marketing strategies. (Gurstein, 2011)

Monino & Sedkaoui (2016, pp. 28–29) present that, when raw data has been made available for reuse and exploitation it can be transferred into useful information which can lead into new innovations and create new services. As there is a great interest for the public sector to increase transparency and modernization, the availability of Open Data is increasing. This will provide more opportunities for the data to be used, reused, reoriented and combined with other data. Utilization of data is a key asset that promotes innovation, growth and development. In this Open Data reuse process, there are different actors that can be divided into four categories:

- **Public Sector:** Governments and other public actors produce and receive a great amount of data.
- **Companies:** Utilizing data in their internal processes to achieve their goals. They produce data that they can relate to external data coming from different sources.
- **Scientific Research:** A lot of data is collected for scientific research and results are published.
- **Individuals:** Individuals can collect data and enrich it with their smart tools.

The greatest benefits on data reuse can be seen for the economic actors, such as big IT-companies, which have the means to utilize the public data to improve their production process. In the digital economy Open Data

utilization can also be seen a great opportunity for new innovative startup development. (Monino & Sedkaoui, 2016, pp. 28–29)

Governments are focusing in making their data available but not much on developing public-facing applications. It is the private sector that creates the added value by building business based on the data. The number of new start-ups is constantly raising, and new ideas and applications are born even on the weather and GPS datasets which have been available for decades. Open Data is booming in new sectors as well, like health and healthcare. (Gurin, 2014, pp. 24–25, 31)

New type of Open Data, Smart Disclosure, combines government data, information on companies' products and services and data on individuals own needs to help consumers make purchase decisions. As not many are database experts, people use Open Data through an intermediary, a "choice engine", to filter the relevant data. Choice engines consider individuals circumstances budget and priorities. It works like a travel booking site. A growing number of companies are turning travel or financial sites or for example price comparison sites into successful business. Smart Disclosure works the best when it combines data of the services a company has to offer with individual data of the customer, such as medical records, phone usage data or energy use patterns. This is special Open Data as it is open only to the individual whom the data is about. Companies dealing with this must be careful with data security. There are four major business models for Smart Disclosure: paid subscription, payment from institutional clients, payment for lead generation (vendors pay a commission for each new customer) or nonprofit (in areas like education or health). (Gurin, 2014, pp. 44–50)

The geographic information is recognized as the area of data with the highest commercial value and it can be combined with data sets from other categories. With geographical data it is possible to use maps to create data visualizations. Carrara et al. (2015, pp. 33–34) rank the Open Data areas in terms of commercial value as follows:

1. Geographic information (highest commercial value)
2. Meteorological and Environmental information
3. Economic and Business information
4. Social information
5. Traffic and Transport information
6. Tourist and Leisure information
7. Agricultural, Farming, Forestry and Fisheries information
8. Natural Resource information
9. Legal System information
10. Scientific information and Research Data
11. Educational content
12. Political content
13. Cultural content (lowest commercial value)

Even if there are a lot of possibilities for Open Data, the business aspect is often missing. Still, businesses in nearly all different fields could utilize Open Data in some way to improve their products or services. There is significant potential in public data for new innovations in different sectors. The possibilities of Open Data are often misunderstood. As the examples of collaborative consumption and sharing economy show, we are no longer what we own but what we share. The same philosophy can be adapted for data as well. With data sharing the tourism business and for example can collaborate with companies from other fields. In the area several Open Data has already been used: geographic data, even data, visitor statistics, supply statistics, survey data, supply information, transit data and governmental data. This data is mainly used for the development of mobile applications and websites for tourists. With tourism businesses opening their data like GPS-data, opening hours and product descriptions in a machine-readable format, innovative tourist services and applications can be created. (Pesonen & Lampi, 2016, pp. 2–4)

2.2.2 Business Intelligence

Business Intelligence (BI) is a practice of transforming data into meaningful and useful information via some form of visualization or report. The practice of BI is growing in business as more companies are looking ways in using their Big Data to gain valuable information. A good dashboard is useful, pleasurable to use and usable in a way that the users can accomplish their goals quickly and easy. Some qualities that make a data visualization excellent are: it should be visually appealing and scalable, it gives user the right information, it is accessible, and it allows rapid development and deployment. (Yuk & Diamond, 2014, pp. 8–11)

Government Open Data can also be used for BI. Companies can analyze their own and their competitor's business based on the massive amount of accessible government data like demographics, healthcare costs, wage and labour issues. Combining this with the company's own data in a visual form it is possible to gain new insights to support business strategy and planning. (Gurin, 2014, p. 42)

There are various of Business Intelligence and Analytics tools that transform raw data into meaningful information to support the decision making throughout an organization. This Visual Business Intelligence or commonly referred to Data Visualization helps users to understand the huge data amounts by showing them in a more visual and simplified way. As databases are growing in sense of their size and complexity, Data Visualization is an emerging field. Many organizations are depended on these tools in their decision making and BI. By the help of these tools the floating information can be transferred into business perspective. (Dupin-Bryant & Olsen, 2014)

As people are getting easier access to information and data, Data Visualization has become a hot topic. By the help of Data Visualization large data masses or complex data can be transferred in a readable and usable form. There are several data visualization tools available for everyone to create their own visual dashboards on their chosen data. Dashboard is a visual tool that helps people and organizations to track, monitor and analyze information to support their decision making and action planning. Dashboards are great evaluation tools and by the help of a dashboard it is easy to follow financial figures for example. Dashboards help people to view and digest information very quickly, which leaves more time for strategic planning. They are also a great tool for communication as by the help of a visual tool information can be transferred in a common format and make an impact. Dashboards are also for help in getting confidence on your own insights. To get useful dashboards it is important to be able to connect them with different data sources that may contain large amounts of data. Having access to multiple data sources is a great way to enrich data. (Stirrup 2014, pp. 6–21)

2.2.3 Smart Cities

As technology is developing and there are more opportunities for utilizing data, Smart Cities are emerging all over the world. There is no clear definition for a Smart City. A Brief description is that they are “those cities that utilize information and communication technologies (ICT) with the aim to increase the life quality of their inhabitants while providing sustainable development.” Policymaking in cities is affected by ICT as that is a new economic force for urban growth changing the way cities compete. In Smart Cities the use of ICT creates a basis for a strategy in several fields like economy, environment, mobility and governance. This is transforming the city infrastructure and services. Cities around the world are trying to find the best models for sustainable urban development. A Smart City generates smart ideas actively in an open environment. Open Data and living labs are utilized in the process which directly involves citizens in the co-creation of products or services. As ICT is implemented in municipal services, cities become more intelligent in their resource management. New business opportunities also arise by the help of these new technological innovations and applications. There are advanced infrastructures such as sensors, electronic devices and networks in a Smart City with an aim to stimulate sustainable economic growth and a high quality of life. (Bakıcı, Almira & Wareham, 2013, pp. 135–138)

There are three dimensions in summarizing the concept of a Smart City: technology dimension, human dimension and institutional dimension. The adoption of technology is critical for a Smart City. Technology and infrastructures are pre-requisites for a Smart City, but they are only the base. Cities adapting technology and investing in it attract talented individuals, companies, universities and research centers. Smart City is in consonance with the concept of smart citizens. This highlights the

importance of creativity, social integration, education and tolerance. The third dimension is the role of local governmental bodies. The work in between citizens and companies proposing their technological solutions. Smart Cities need smart governments and transparent governance. Effective cooperation between citizens, companies, universities and administration is essential. On solution supporting these goals is Open Data. By giving access to large quantities of datasets public organizations will allow different actors to solve common problems, improve services and create new innovations. (Capdevila & Zarlenga 2015, pp. 267–268)

In transportation development in Smart Cities one target is in automating all aspects of mobility. The solutions lay in self-driving Connected Automated Vehicles (CAVs) and sharing transportation resources. To make this possible the availability of large amounts of data and ubiquitous wireless network are a requirement. In Smart Cities the technological infrastructure is based on network sensors and actuators that interact with wireless mobile devices all over the city. A Cyber-Physical System like this collects data from for example traffic conditions, the occupancy of parking spaces, air/water quality information, the structural health of bridges, roads, or buildings, and the location and status of city resources. Developing the transportation is essential as it has a great effect in resource consumption, strain on the environment and citizen comfort. The development of CAVs offer an interesting opportunity for the future in terms of improving traffic safety and reducing pollution, energy consumption and traffic delays. (Cassandras 2017, pp. 1–2)

2.3 Open Data in Finland

Open Data celebrated its first decade in Finland in a seminar held in Helsinki 8.-10.10.2019. Finland can be considered as one of the forerunners in data opening. Roots lay already in the 1980s when digital geographical information was created for optimizing routes for milk truck drivers and inseminators. Contours were transformed in a digital form and a digital geographic model was created in the early phases of the 1G network. Data utilization was not however spread in a wider use as data was not available, at least not for a reasonable price. The Environmental Council, for example, could not afford the access to the data of the National Land Survey of Finland even if there was a need for it. (Manninen, Poikola, Rastas, Rainio, Reh binder, Tuominen, Kauhanen-Simanainen, Marttila, Suonsaari, Ahola, Saarinen, Koivusaari, Lahtinen, Kivekäs, Honkanen, Tjukanov, Nolvi, Haverinen, Kettunen & Lahti, 2019)

The data mass market was born in the 2000s when computers and mobile phones were spread into wide public use. A report from the Finnish Business and Policy Forum (EVA) in 2009 proposing the opening of data from the public sector for free use for everyone was a turnover effecting into more data being available for free. It was also stated in the government programme that time that public data resources will be

opened. The spearhead in the data openings was the geographical data from the National Land Survey. As it was opened for free use, in two days it was used more times than ever before. This was the first leap for Open Data spreading in Finland. (Manninen et al., 2019)

The Finnish “open data community” also began to form at that time. The first Open Data competition, Apps4Democracy was held in 2009. In the beginning the community did not yet have much understanding on what Open Data is. As a result, an Open Data guidebook was created in community collaboration in 2010. Helsinki Region Infoshare (HRI) was founded a year later. Four cities in the Helsinki Region (Helsinki, Espoo, Vantaa and Kauniainen) were all involved in the very beginning. HRI is a joint service for open data distribution in the region. It collaborates with other Finnish cities and international contacts. Open Knowledge Finland (OKF) was founded in 2012 as a result of Open Knowledge festival. OKF is an important institutional structure enabling the promotion of opening and utilization of data in Finland. (Manninen et al., 2019)

Digital experts are a significant group in the Open Data development. Through a 2013-began EU-project Code for Europe, involving also the city of Helsinki, Helsinki Regional Transport (HSL) and Forum Virium Helsinki, some of these experts were brought for this kind of data development in the public sector in Finland for the first time. Their mission was to find relevant data to be published and utilized. First, the decision-making data of the City of Helsinki was opened. Also, the first prototypes for HSL Route Search were developed. Today, Open Data is a natural part of the development of services for the public in Helsinki region. The focus is now in the creation of sustainable services exploiting the open source code. Helsinki has been the leading city in Open Data publishing and utilization in Finland and according to the chairman of the city council Otso Kivekäs, is aiming to become the number one in the world. (Manninen et al., 2019)

During the last decade, the Finnish Government has also done a lot for promoting Open Data in Finland and in the EU. Open Knowledge Programme, running in 2013-2015, boosted the opening of public data in a machine-readable format with open access for citizens, municipalities, agencies, companies, science and research, education and media. In the past even the governmental agencies had to pay for each other for the data. After the programme the promotion of Open Data has become a part of governmental actions. This has resulted in the opening of the national opendata.fi portal, the raise of Finland’s position in international indexes, data policy being involved in the preparation of government programmes and the revision of the Open Data and PSI Directive. (Manninen et al., 2019)

Today government, companies, educational institutes, scientists, media and people are exploiting Open Data in Finland. The digitalization of acquisitions has been one of the key projects in government programme.

Recently opened Tutkihankintoja.fi service allows the citizens, companies and interest groups to explore the state spending. The raw data from the site is also published at the 2014-established national data portal opendata.fi, as does also the HRI's data. Companies in Finland utilize Open Data in various ways. New businesses are born on the base of Open Data. Also new content can be added to existing applications by for example adding Open Data into geographical information. Media uses Open Data in data journalism. In the last few years a lot of interesting articles based on Open Data information has been published in Finland. (Manninen et al., 2019)

Majority of the Finnish Open Data is published in the opendata.fi (avoindata.fi in Finnish) portal. It is managed by the Digital and Population Data Services Agency of Finland. It is a national Open Data portal where people and organizations can publish and exploit Open Data. The portal is mainly a catalogue of public administration's Open Data. Utilizers of the portal include citizens, companies and public authorities. (Digital and Population Data Services Agency, n.d.) Most public organizations also publish their Open Data on their websites.

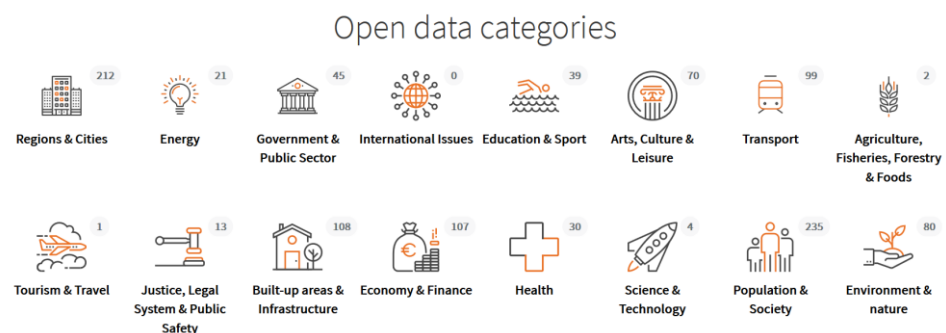


Figure 5. Available datasets by categories in the opendata.fi portal in January 2020

A report of the Prime Minister's Office Finland (Koski et al., 2017, pp. 10–12) presents some of the available Open Data in Finland:

- **Digital and Population Data Services Agency (VRK):** information on buildings (address, coordinates, building codes, the purpose of the building, voting area), catalogue on the available public services in Finland, citizen name data, municipality information
- **Institute for Health and Welfare (THL):** information from several registers, Open Data on access to treatment, specialized medical care, processing times in child protection, infertility treatments, abortions, sterilizations, childbirths and income supports
- **Meteorological Institute:** weather data and forecasts, climate observations
- **National Land Survey (NLS):** geographical data (roads, buildings, names of places, land usage, land height)
- **Natural Resources Institute (Luke):** statistics on agriculture, forestry, fishing and groceries

- **Patent and Register Office (PRH) / Tax Administration:** company information in the Business Information System (YTJ)
- **Statistics Finland:** Basic statistical information from Finland and abroad. Ready-made API's for data users.
- **Social Insurance Institution (KELA):**
- **Tax Administration:** public tax information on communities and joint benefits
- **Transport and Communications Agency (TRAFI):** information on vehicles and boats in use and/or in register, includes registration data and technical information
- **Transport Infrastructure Agency:** maps, statistics, publications and data sets related to roads and traffic, real-time traffic information, public transport schedules and routes, traffic predictions

The Finnish legislation restricts the opening of certain data because of privacy protection, trade secrets, copyrights and national security. The goal though, is to bring all the relevant data sources available in the opendata.fi portal. To increase the openness and transparency of Finnish public administration, international cooperation is done. Joining the Open Government Partnership, discussed more in detail in the next chapter, is an example of this. According to the Open Data Index by Open Knowledge Foundation, Finland is in the frontline in opening public data. (Koski, 2015, pp. 4–5) Finland lies the fifth in the recent ranking measuring different public Open Data areas in terms of if it is openly licensed, in an open and machine-readable format, downloadable at once, up-to-date, publicly available and available free of charge (Open Knowledge Foundation, n.d.). Even if the index is only indicative, it gives a good overview of the global situation.



Figure 6. Top 5 of the Open Data Index in 2017 (Open Knowledge Foundation, n.d.)

According to Koski et al. (2017, p. 16.) the biggest Open Data user group in Finland are the companies, with about a share of half of all the Open Data utilized. Other users are the other public sector actors, research institutes, universities and individuals.

6Aika is a project of the six biggest cities in Finland which aims in bringing Open Data as part of the normal operations of the cities and offering Open Data for all. A survey of the project (Muurinen, 2017, pp. 4–5) states that Finnish companies have found ways to exploit Open Data like company register data, financial statements, maps, geographical data, traffic data and genetics data as a resource for their products and services. Often the open resources are only used as a part of their product or service development. Common is to build digital services in forms of applications, cloud services, API's, accesses to data sources and Software as a Service- and Platform as a Service products. Companies are also selling analytics and information services and products, consulting and expert services, Mobility as a Service, conveying customer flows and different supportive services. Businesses based purely on Open Data are rare.

2.3.1 Open Government Partnership

Finland joined the Open Government Partnership (OGP) in 2013. It is “a multilateral initiative that aims to secure concrete commitments from governments to promote transparency, empower citizens, fight corruption, and harness new technologies to strengthen governance.” OGP was formally launched in 2011 by 8 founding nations and in the beginning of 2018 it had over 70 participating countries. To join OGP a country must “endorse a high-level Open Government Declaration, deliver a country action plan developed with public consultation, and commit to independent reporting on their progress going forward.” (Open Government Partnership, n.d.)

Finland committed to its first action plan in its year of joining, 2013, for years 2013-2015 (Ministry of Finance, 2013). Three more action plans have been published since, in 2015 for years 2015-2017, in 2017 for years 2017-2019 and in 2019 for years 2019-2023. The first two action plans had enhancing citizen engagement as their core target. Both of these consisted of four themes.



Figure 7. Finland Open Government Partnership Action Plans I and II (Ministry of Finance, 2017, p. 2)

In the first action plan drawn by the Finnish Ministry of Finance (2013, pp. 2–8) it is stated that in the Finnish government openness is an underlying value and it can be seen for example in the strong self-government of municipalities where citizens can be close to decision making. Basic principles in Finnish municipalities include transparency and citizens right to good governance. It has been acknowledged that government data being open enables the following the functioning of public administration and participation. Finland is committed in opening and publishing new data and changing existing Open Data into a machine-readable form.

The open government action plans cover the whole public sector in Finland. Openness and access to information are highly valued. Themes of the action plan for years 2017-2019 were: easy to understand, open region and access to information. (Ministry of Finance 2017, p. 2)

III National Action Plan 2017-2019

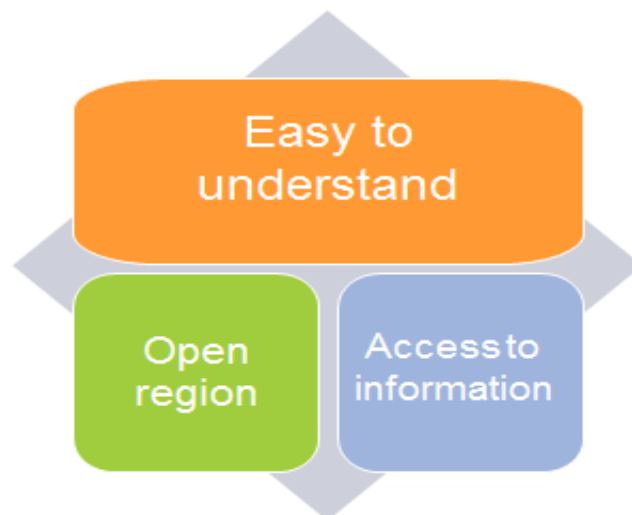


Figure 8. Finland Open Government Partnership Action Plan III
(Ministry of Finance 2017, p. 3)

Third open government action plan was to be implemented on all levels of public administration in Finland. This included the state administration as well as the municipalities and regions. Focus was in enhancing open government and citizen engagement in the regional level. Aim was in understandability of administration, openness of information and participation possibilities for citizens. (Ministry of Finance, 2017, p. 5)



Figure 9. The Eight Elements of Finland's Open Government Activities according to Action Plan IV (Ministry of Finance, 2019, p. 4)

Current action plan covers four years (9/2019-9/2023) as it is the best way to support the goals of the Finnish Government Programme throughout the whole government term. It was recognized that the previous action plan periods were too short to engage in long term activities and making an impact. The four starting points for promoting governmental openness are stated as trust, participation, transparency and combating corruption. The four commitments are lasting openness, open government strategy, transparency register and Open Data. (Ministry of Finance, 2019, pp. 3–13)

On Open Data the commitment and measures relate to OGP initiatives of transparency and technology & innovation. It is stated that “the public sector promotes government openness by opening public interfaces if there are no specific reasons to keep them restricted.” The most influential data sources are prioritized. For public data resources, easy-to-use and developer-friendly interfaces are implemented. Data is produced in machine-readable and -interpretable format so that it is easier to benefit from it. Guidance is provided for Open Data and interface creation. Aim is to increase the re-use of the data significantly. This will motivate the data producers to improve its quality. When the benefits of Open Data spread wide in the society, companies will hopefully become also producers of Open Data. Quality criteria intended to facilitate the utilization of data will also be prepared. Ethical guidelines related to AI and data policy is built. (Ministry of Finance, 2019, pp. 14–15)

2.3.2 The Six City Strategy

In Europe, Finland is one of the leading countries in developing smart cities. Smarter city platforms are developed to save peoples time in a sustainable way, improve their quality of life, create business opportunities and new jobs and to increase know-how. Citizens and app users' viewpoints are taken highly into count in Finland in developing smart solutions. Companies and start-ups are involved in the transformation process of smarter environment, mobility, governance, economy and living. The six largest cities of Finland (Helsinki, Espoo, Vantaa, Tampere, Oulu and Turku), covering 30% of the Finnish population, have together facilitated an open innovation platform called "The Six City Strategy". It is targeting in common development of smart solutions and implementation of experimental projects intending to tackle challenges related to urban environment. In these cities' companies have a great possibility to test their innovative solutions. Data, experiences and standards are shared between stakeholders. Thanks to Helsinki Region Infoshare, Helsinki has also been awarded as the best Open Data city in the world. (Laakso, 2017, p. 3)

In the capital city area, the cities began to open their public data in a large scale over ten years ago in cooperation. As a result of the 6Aika project (2014-2017) and The Six City Strategy the cooperation has spread with the national government, higher education and with other cities as well. Development of smart solutions is a natural part of operation in the big cities nowadays and many of the solutions and innovations are shared. In the recently published service map of Turku, the code has originated from Helsinki. The six cities involved in this cooperation are also members of a global Open & Agile Smart Cities network. (Manninen et al., 2019)

The smartest district in Helsinki is Kalasatama. In 2017 a total of 3 000 inhabitants lived at the area. The aim is, that in 2030 there will be 20 000 inhabitants and 8 000 jobs. Smart urban living and services can be experimented in the district and they are developed in collaboration with the residents, companies, city officials, researchers and other stakeholders. The smart solutions in the area include floating houses, adaptive and scalable technologies related to heating and lighting, digital health services, shared electric cars, future school, home appliances controlled by mobile devices, smart waste collection system, carbon neutral zoo, solar park, Smart Grid, IoT solutions testing and personal data usage for customized services. (Laakso, 2017, p. 4)

Smart Tampere program is aiming in building a smarter city and securing economic growth cooperative with local and international companies, universities, organizations, citizens and the city government. The city acts as a testing ground for smart solutions that can also be transferred as exportable products. Turku is strongly promoting in sustainable energy transition, circular economy and smart mobility. It is a pioneer city in

Mobility as a Service solutions. Oulu is the home base of 600 high tech companies and the fastest growing region in the Nordics. It hosts MAPGETS which utilizes Open Data to create a real-time and open 3D-environment where information is visualized utilizing 3D-models and location data. (Laakso, 2017, pp. 4–6)

2.3.3 Open Data in Häme Region

Häme region consists of three sub-regions surrounding the three cities within its borders: Hämeenlinna, Riihimäki and Forssa regions. A total of 175 000 inhabitants are living in Häme. Around Hämeenlinna there are a lot of work and education opportunities. Housing, education, entrepreneurship and culture are valued in the area. Riihimäki region is growing and the area attracts people and companies from the capital region to move in by considerably lower prices of sites, with just a half an hour car ride away. Forssa region is looking for ways for environmentally friendly and sustainable development. In Häme living varies a lot, from city life to farming. (Regional Council of Häme, n.d.).

The city of Hämeenlinna has published its purchase data on a yearly basis as Open Data since 2013. The data is open excluding the information restricted by data protection legislation. (City of Hämeenlinna, 2019a) It is available in avoindata.fi portal in .xls and .csv formats. The city of Riihimäki has published some maps in the portal and from Forssa there is no data available. The Regional Council of Häme has published the regional land use plan. Häme University of Applied Sciences offers its reservation, course unit, realization and curriculum data through API's (HAMK, n.d.). As there is not much local Open Data available, it is fruitful for organizations in the region to exploit national level Open Data.

Jussila, Kukkamäki, Mäntyneva & Heinisuo (2019, pp. 28–30) present two open source-based solutions developed in the region recently: Tavastia Events API and Hämeenlinna in Pocket smartphone application. Tavastia Events is the most important concrete development activity of the Open Häme project. It is based on an open source solution Linked Events, originally developed by the city of Helsinki. The Event API was created to respond for the demand pointed out by several stakeholders. It has two main components: an event database plus the API itself and a website tavastiaevents.fi. Through the website it is easy to search for events in the region. What separates this from the traditional web-based event calendars is that the API receives machine-readable messages in structured forms from several systems and saves them in the database. Anyone can fetch the event data programmatically and use it in a way one finds suitable. With this, the numerous event organizers and information users can automate their input and output processes. The advantages of this are that an event must only be submitted and possibly edited at one place and that for anyone using event data can easily collect and display their desired event data. Hämeenlinna in Pocket also uses this event data

and through the app it is possible to view them as a chronological list or visualized at the city map. The app also includes topical information and news, public transport information, digital library card, “my health platform”, schedules for free-time sports activities, care-time allocations for nurseries and a feedback channel.

3 METHODOLOGICAL CHOICES

Open Data is not yet in a very large scale utilized in Häme region. The aim of this research was to find out ways how to increase the knowledge of local organizations on Open Data, examples on how they can utilize Open Data and for what kind of data there is a demand.

3.1 Research Questions

The research problem defines the most suitable research frame for problem solving, what kind of data is collected and how (Eskola, Lätti & Vastamäki, 2018). Here it was interesting to find out the local companies view on Open Data. The research is based on three research questions:

1. What is the current situation with Open Data utilization in business in Häme region?
2. How could the local companies utilize Open Data better?
3. For what kind of Open Data there is a demand?

As a result, the OpenHäme project team will get an idea how local companies are utilizing Open Data now and what actions should be done to improve the situation. The research findings can be exploited by HAMK Smart research unit and the local open data actives from other organizations in their future development projects and actions. The companies participating the survey can also find out new ways of utilizing Open Data.

3.2 Research Methodology

3.2.1 Qualitative Research

This is a qualitative research. A qualitative research leads to increase in overall understanding of the quality, characteristics and meanings of the research object (Lähdesmäki, Hurme, Koskimaa, Mikkola & Himberg, 2009). It is an umbrella term for variety of research approaches and data collection methods. The information or data collected is mainly nonquantitative. The research material may consist of interview transcripts, fieldnotes, documents or visual materials like photographs or videos. The outcomes of qualitative research often compose of relevant representations and presentations of substantial findings from the analytic synthesis of data. There multiple genres, elements, and styles of qualitative research, like literary genre, ethnography, grounded theory, phenomenology, case study, content analysis, mixed methods research, poetic inquiry, arts-based research, autoethnography, action research, investigative journalism and critical inquiry. (Saldana, 2011, pp. 4–20)

Often qualitative research studies are based on interviews with participants. It is an effective way in seeking for target group's perspectives, feelings, opinions, values, attitudes and beliefs. The research topic and research questions create the basis of an interview, but new insights and unexpected areas may arise from improvised conversation during the interview. (Saldana, 2011, p. 32)

3.2.2 Data Collection Method: Thematic Interviews

Interview was chosen as a data collection method as it produces knowledge of opinions, ideas, observations, attitudes, values and experiences of living people. In interview the interviewer participates interactively in the production of the data. (Lähdesmäki et al., 2009) When it is wanted to know what a certain actor thinks of a certain issue, most efficient is to ask them directly. In this case this means the views of local companies on Open Data. Interview is a discussion where a researcher aims to find the answers related to the research topic from the interviewees. (Eskola et al., 2018)

More specifically the data collection method used is thematic interview. In thematic interview the respondent has the right to add comments one finds relevant. If the interviewer finds these new topics interesting, additional questions can be presented based on the new viewpoint. The interviewer, however, also has the possibility to direct the conversation back to the originally selected themes. (Routio, 2007)

Until 1980s the common set-up for interviews was that the interviewer tried to be as inconspicuous as possible and was asking structured questions. Today interview is a widely used method for research and other purposes and the view on the relationship and the interaction between the participants has changed. An active role for the interviewer is recommended. The interview situation can be made more natural and pleasant if the interviewer reacts to the responses and shows one is listening by for example nodding one's head. This encourages the interviewee to share one's experiences more freely. It is still important to remember that the researcher has the responsibility to take the initiatives as interview a goal-oriented data collection situation. Often the interviews are recorded for further analysis of the researcher. (Eskola et al., 2018)

Thematic interview lies somewhere in between a structured and an unstructured interview. The formation and order of the questions may vary between different interviews. The theme areas are however pre-defined, and all theme areas are being discussed in all the interviews. The interviewer may have a list of supportive questions under the themes, but one may freely decide which ones are relevant in each interview. The emphasis of the different themes may also vary between the interviews. (Eskola et al., 2018) In this research the variation of these issues in different interviews was dependent on the field of business, the maturity

of Open Data utilization in the company and the role of the interviewee in the company.

A successful interview may feel more like a natural conversation instead of a traditional interview. The basic structure is how ever important for the research results to be comparative. The approach in thematic, or semi-structured, interviews is both natural and structured. In the interview the discussion may jump from a pre-planned question to another, but if some answers are interesting and relevant the researcher may ask the respondent to tell more about something or give examples. (Gournelos, Hammonds & Wilson, 2019, pp. 105–106)

In this case the observations are made within the representatives of local companies in Häme Region. As Open Data is a rather new phenomenon and it is not much utilized within the region, it is important that the interviewer is able to introduce the topic and give enough background information for the respondent.

3.3 Research Quality

Sarah J. Tracy (2019, p. 269) has presented the eight “big tent” criteria for excellent qualitative research. In the model the end goals of a good qualitative research are differentiated from the researchers mean practices to get there. Appropriate research practices have been followed in this thesis process.

Criteria for Quality	Attributes
Worthy topic	Topic that is worthy, timely, significant, interesting.
Rich rigor	Theoretical constructs, data and time in the field, sample (s), context(s), data collection and analysis processes.
Sincerity	Self-reflexivity of researcher values and potential biases; transparency in methods and challenges.
Credibility	Thick description; triangulation or crystallisation; multivocality; member reflections.
Resonance	The research uses aesthetic, evocative representations, naturalistic generalisations, and/or transferable findings, to influence or move readers.
Significant contribution	Through advances conceptually/theoretically, practically, morally, methodologically, or heuristically.
Ethical	Procedural ethics, situational and culture specific ethics, relational ethics, exiting ethics.
Meaningful coherence	Methods and procedures fit stated goals, interconnect literature, research questions, findings and interpretations with each other, and achieve stated objectives.

Figure 10. The Eight Big Tent Criteria for Excellent Qualitative Research by Tracy (Rowe, 2013).

Worthiness of the topic is the first criterion for qualitative quality. Worthy topics may either strengthen the findings from earlier studies to show how they have remained relevant over time or then they can reveal new aspects that have been overlooked, misunderstood or mistaken. A good topic is worthy, timely, significant and interesting. (Tracy, 2019, p. 271) Not a lot of research has been done on Open Data business aspect, especially in Finland. The business potential of Open Data is widely recognized but not many examples of how companies are utilizing it are available. A lot of Open Data initiatives have been made in Helsinki area and other cities included in the Six City Strategy. Smaller cities and regions, like Häme, are still mostly in the early phases in opening data and recognizing the utilization possibilities.

The next goal is rich rigor. This means the care and effort made doing the research in an appropriate way. In quality research theoretical constructs, data collection and analysis are done thoroughly. Rigorous practices include identifying goals aligned with the context, spending enough time in the field, having appropriate procedures in collecting and analyzing data and having enough data to support the findings. (Tracy 2019, p. 271) This being a master's thesis, the author is not an expert in the field of Open Data. In fact, the phenomenon was rather new for the author. The research process took however over three years and the author had the opportunity to follow the changes happening in the field during the gathering of theoretical content. The theoretical part is based on literature and other recent sources. Data was collected as interviews with local companies. Interviews were recorded and transcribed for further analysis. In the analysis the findings were combined with each other and with the theoretical framework.

Third aspect is the sincerity of the research. A good qualitative research is authentic and vulnerable. The possible challenges are made visible and the methods used are transparent. The researcher must be self-reflective about one's subjective values, biases, and inclinations. (Tracy 2019, p. 272) The researcher himself does not have experience or much technical know-how on Open Data utilization. The research was guided by the OpenHäme project team which may have had an impact on the biases of the researcher. The research process took more time than originally planned caused by the humane factors of the researcher. Determining the interview questions and companies to be interviewed was found especially challenging. The implementation of the research is described transparently in later chapters.

Fourth marker for quality qualitative research is credibility. In research, it refers to dependability, trustworthiness and expressing a plausible reality. Readers of a credible report feel confident in using its data and findings to their purposes. It can be achieved through thick description, triangulation or crystallization, multivocality and member reflections with participants. (Tracy 2019, p. 275) On the findings of the theoretical framework and the

empirical research have been discussed with the OpenHäme project throughout the process and their comments guided the research going into the desired direction.

By resonance it is meant that the research influences, affects, or moves readers or a variety of audiences. This can be achieved by several different practices like transferability and naturalistic generalization or aesthetic merits. Aim is to make an impact to the readers actions. (Tracy 2019, pp. 279–280) One of the main targets of this thesis was to find out how companies are utilizing Open Data and what kind of potential there still is for further development. The findings presented are transferable and applicable for many businesses.

Sixth factor is if the research makes a significant contribution in some way. The findings of a significant research may “extend, transform, or complicate a body of knowledge, theory, or practice in new and important ways”. For a research to be significant the new insights or findings creating deeper understanding to a certain issue must not be necessarily huge, but they must somehow make an impact on the current landscape. (Tracy 2019, p. 281) This research is based on existing literature on Open Data and the examples of Open Data utilization in business. Aim was to extend the topic and give an insight on the situation in Häme region in this topic as not much further research has been made.

Ethical research practices are also considered in quality qualitative research. Tracy (2019, p. 283) divides this into procedural ethics, situational ethics and relational ethics. In this research the confidentiality of the interview participants has been ensured. Trust has been built by open and transparent communication with the participants. The permission to use their and their company’s names was asked. It was agreed that the analysis is made on a general level so that no one’s sayings can be identified, misunderstood or misused.

Meaningful coherence is the final characteristic of qualitative quality in the eight “big tent” criteria. The research should achieve its stated purpose, accomplish what is expected, use methods and representation practices that connect well with theory and accurately interconnect literature reviewed with research targets, methods, and findings. (Tracy, 2019, p. 286) The theoretical framework in this thesis is tightly connected to the research goals. The research findings are in line with the original purpose of the thesis.

3.4 Research Process

The research was done as thematic interviews with local companies and one open source expert. Target was to get companies from different field of business involved in the research.

3.4.1 Selection of the Interviewees

The selection of the interviewees takes often a lot of effort, but done well, can be very rewarding. It is important to reach those people whom it is believed to have the wanted knowledge or experience on the topic or it is known that one has been involved in certain actions or processes. The research mission defines the most important criteria for the interviewee selection. (Eskola et al., 2018)

The required number of interviews depend on the research and opinions on this also vary. A deep single case interview may be valuable for making as rich profile, but for many cases one individual is not enough to represent large enough population. With two interviews it is possible to compare the results and recognize possible contrast between them. For a broader spectrum it is recommended to have some three to six interviews. In certain researches ten to dozens or even hundreds of respondents are needed for credible and trustworthy findings. The general rule is to have interviews until the point they stop producing anything new. (Saldana, 2011, pp. 33–34) Here it was recognized a small group of respondents will give enough insight on Open Data utilization in the region. It was assumed that the research data will start to become recurrent after about five to six interviews.

To gain an overall insight of the situation and to get support for the actual company interviews it was decided to interview Timo Väliharju, the Executive Director of COSS – the Finnish Centre for Open Systems, first. Väliharju is leading the Smart Services ecosystem work in the Open Hämeenlinna ecosystem complex. The main themes of the Smart Services ecosystem are regional economy, smart digital services, open source and Open Data (City of Hämeenlinna, 2019b).

The target was to get company representatives from different fields of business. To cover the whole region, it was also seen important to get companies from all the sub-regions: Forssa, Hämeenlinna and Riihimäki. One decisive manner in the selection process of the companies was also that it was already known that they are utilizing Open Data, or the researcher was able to recognize Open Data utilization possibilities for the selected field of business.

The assumption based on discussions with open data actives (for example in AvoinHäme afterwork event 23 February 2018), internet sources and the insights of the OpenHäme project team and Timo Väliharju (interview 14 February 2020) was, that not many companies in Häme region are utilizing Open Data in their business currently. Companies with no clear utilization potential or technical resources for data handling were excluded from the interviewee selection. The aim of the research was to find out case examples of Open Data utilization and potentially new utilization possibilities for companies.

Based on this, a shortlist of possible companies and contact persons was created. Eventually eight companies were contacted by phone or e-mail. Six of these agreed for an interview. As these six companies included representatives from all the sub-regions of Häme and they were from heterogeneous fields of business, it was seen that there was enough take for the analysis.

The companies selected for interviews were

- Pohjantähti
- Hämeen Sanomat
- Ramboll Finland, Digital Mobility Lab
- Simosol
- Teleste, Display Competence Center
- Vitabalans

Pohjantähti is an insurance company having its headquarters in Hämeenlinna. It was an interesting case example as the assumption was that a lot of data must be analyzed in insurance product planning. The person chosen for the interview was a product development manager, who plans new insurance products and develops the existing ones in person and animal insurances for private and company sector (Wiitamäki, interview 19 February 2020).

Hämeen Sanomat is a local newspaper of the Häme Region, based in Hämeenlinna. It was chosen for the research to find out if the local media is utilizing Open Data for its data journalism purposes. The person interviewed was a news editor, who's duties include a role of a supervisor, editing the news for the print and web version of the newspaper, print design, developing the web page from a journalistic perspective and writing news (Uhari, interview 24 February 2020).

Ramboll is a global planning and consulting company. In Finland they employ over 2 500 experts and operate nationwide. (Ramboll, n.d.) The chosen interviewee was a data scientist from Ramboll Finland's Digital Mobility Lab, which designs traffic models, analyzes traffic data and creates new digital solutions for cities and regions in Finland (Piipponen, interview 27 February 2020). Traffic planning sector was compelling for the research due its obvious Open Data utilization potential. It is also one of the specific target sectors of the OpenHäme project.

Simosol is offering digital solutions for the fields of forestry and agriculture. They describe themselves as the "global leaders in solutions for sustainability and optimization" (Simosol, n.d.). The company has its office in Riihimäki and the CEO of the company agreed for an interview. Simosol operates in Finnish forest harvesting planning and does land surveys internationally. Open Data is utilized in both. (Rasinmäki, interview 6 March 2020)

The Display Competence Center of Teleste delivers passenger information systems for trains, trams and other rolling stock. Information displays are also set up to stations and bus stops. They also produce commercial displays. The team of the interviewee does mechanics, equipment, electronics and software planning. The unit is based in Forssa. (Sulonen, interview 13 March 2020). It was expected that in the field strongly connected to public transport, Open Data would be used as an important piece of information.

Vitabalans is a pharmaceutical company that develops, manufactures and markets medicines, food supplements, medical devices and special animal feeds. All of their production happens in Hämeenlinna but they are exporting their products in 13 countries. (Taipale, interview 7 April 2020) The researcher was eager to learn what kind of data may be utilized in medicine production, development and marketing.

3.4.2 Structure of the Interview

A basic principle in an interview is that it is better not to go straight to the point but to make some easy questions first. The theme of the research might not be actively on the mind of the person interviewed and so it is worthwhile to give some orientation to the topic first. (Eskola et al., 2018) To make the interview situation comfortable a couple of warm-up questions are asked before going to the topic. The interviewee will be given an opportunity to tell about one's company and one's own role in the company first. Then the discussion will go into how the company is utilizing data in general (not necessarily Open Data) in their business. This will be valuable information for the interviewer before guiding the discussion into the actual themes.

The questions in an interview may vary from highly open to highly closed questions. Open questions give freedom for the respondent to choose to which direction one wants to go. Questions being too open may lead the discussion to side-roads. Highly closed questions however are great for the comparativeness of the results but may affect the richness of details in the answers. Some viewpoints may be left out if the interviewee is only giving short answers to closed questions. In a qualitative in-depth interview, it is recommended to use mostly moderately open questions. (Gournelos et al., 2019, pp. 107–108)

Highly open questions	Moderately open questions	Highly closed questions
Strengths *Gives participant completely freedom to respond in any way.	Your happy place *Participant has both direction and freedom in their response *Allows for participants' response comparison *Scripted questions prevent interviewer bias	Strengths *Saves time *Easier to compare participant responses
Weaknesses *Takes more time *Prone to leading questions and response bias		Weaknesses *Limits the freedom of the participants' response and prone to selection bias
Example <i>What is health?</i>	Example <i>What was your experience like at your last visit to the doctor?</i>	Example <i>Were you satisfied or dissatisfied with your last physician visit?</i>

Figure 11. Open and Closed Questions in an Interview.
(Gournelos et al., 2019, p. 108)

The structure of the thematic interview was build based on the theoretical framework and research questions of this thesis, the insights from the OpenHäme project team and the interview of Timo Väliharju. The final base structure (appendix 1) consisted of three main themes around Open Data in the companies: current situation, future utilization potential and wishes on new data openings. Supportive moderately open interview questions were planned under each theme. The thematic interview method allowed the researcher to make alterations on the questions and tailor them for each interview considering the company's field of business and the familiarity of the theme for the person interviewed.

The first theme focuses on how the company is utilizing Open Data in their business currently – if they are. It is also mapped how familiar the interviewee is personally with the concept of Open Data and its utilization possibilities and barriers. Finally, it is found out whether the company is utilizing Data Visualizations in their business.

The second theme is about the future of Open Data in the company. Interesting are the possible utilization plans, the attitude of the company management towards Open Data and the examples of other companies in the field. It is also asked what kind of support the company would hope for to utilize Open Data better. The interviewer may also give some tips and ask opinions on the possible new ways of utilizing Open Data.

In the third theme it is harvested how much the person knows about currently open data sets and if the company has ideas on further data that should be opened. The data wishes may concern whole new possible publishers or certain datasets. It is also found out whether the company has done any actions or data requests to improve the situation.

3.4.3 Implementation of the Interviews

The interviews were implemented in February-April 2020. Some of the interviews took place in the company offices or other public places, some were implemented via telephone or internet call to avoid unnecessary travelling. Detailed information the interviews is presented next.

Organization	Industry	Representative	Method	Date
COSS - the Finnish Centre for Open Systems	Non-profit organization	Timo Väliharju, Executive Director	Telephone	14.2.2020
Pohjantähti	Insurance	Päivi Wiitamäki, Product Development Manager	Face-to-face	19.2.2020
Hämeen Sanomat	Media	Markku Uhari, News Editor	Face-to-face	24.2.2020
Ramboll Finland (Digital Mobility Lab)	Traffic Planning	Johanna Piipponen, Data Scientist	Telephone	27.2.2020
Simosol	Programming (Forestry and Agriculture)	Jussi Rasinmäki, CEO	Face-to-face	6.3.2020
Teleste (Display Competence Center)	Information Solutions	Piia Sulonen, Head of Display Competence Center	Microsoft Teams	13.3.2020
Vitabalans	Pharmaceutics	Juho Taipale, Full Stack Developer	Microsoft Teams	7.4.2020

Table 2. Thematic Interviews on Open Data Utilization Implemented in February-April 2020.

All the interviews were implemented in Finnish as it was the native language of all the participants. By this it was ensured that no language barriers did not interfere the research. All the interviews (excluding the one with Timo Väliharju) were recorded. The main points of each interview were transcribed in detail in Finnish and reported for the OpenHäme project. The analysis of the research, in more general level, was conducted in English and is presented in the next chapter.

4 RESEARCH FINDINGS

All the companies in the research were utilizing a lot of data in their business. This was no surprise as the potential for data utilization was obvious within the companies selected to the research. Majority of them were also utilizing Open Data, some more, some less. Most often Open Data was combined with some other data, being it company's own data or closed data from public organizations. Data was used to support decision making, predict the future trends, monitor change and develop new products and innovations.

Common for all was that statistics from public organizations were exploited a lot. Some of this data was open and some required an information request to the publisher. Majority of the information behind an information request was closed data for a well-grounded reason like information security or other legislation related issues.

Some organizations in the research did not have the technical capability or time resources to handle massive sets of raw data. Instead they preferred processed data and statistics from the publishers. Some pointed out that their bigger competitors in the field are exploiting data a lot more due to their comparatively massive resources. The input-output ratio must always be carefully thought when planning to start processing data. In many cases however, there were plans or at least interest to hire new people or use more resources in other way for data processing in the near future. The ongoing data revolution was recognized in the organizations involved.

In some of the organizations, on the other hand, utilization of Open Data was already a natural part of their business and they had people, or even a unit, capable of handling the data. Both national and international Open Data was used in the companies interviewed. Some of the business models presented could not have existed without Open Data.

4.1 Open Data Utilization – Current Situation

The concept of Open Data was familiar for all the respondents. Quite many of them had a personal interest towards the phenomena and followed the discussion within Open Data communities even on their spare time. The technical skills to actually process and exploit raw data however varied. For some the data processing was in the core of the job, but others needed support from their IT-department or elsewhere. Respondents with lighter IT-skills, were basically able to handle data that can be handled with MS Excel. It was recognized that from knowing the basic principles of Open Data there is still a long way for being able to utilize it for business purposes. First, one needs to find out a bright idea to use some specific data and second to be able to execute that vision.

The process for searching Open Data was quite similar in all the companies. Most often the desired data was downloaded from the publisher's web site directly. Based on their experience, the respondents knew the relevant data publishers for their field of business. Open data portals like avoindata.fi or HRI were also familiar for most of the respondents, but usually the case was that they went to search for some specific data from a specific publisher. Some stated that they could spend some more time going through the portals as well as they could find some potential data-sets they have not yet figured out. Others felt they were in no need for new data sets or publishers as there was already a lot more they can sort out from the data sets they are currently using.

There was a lot of variation in how much the respondents were utilizing Open Data. Some used it daily, some weekly, some occasionally and some basically not at all. This depended a lot of the nature of business and the role of the person in the company. Majority thought, that if they had more time resources, they might be able to utilize Open Data more. The role of Open Data was often only supportive and the main data sources were company's internal data or public data subject to authorization. Yet still, there were some business models that would not exist if there were no Open Data.

The Open Data utilized also varied within the example cases and included Open Street Map data, geographic data from NLS, General Transit Feed Specification (GFTS) data, data from the Finnish Transport Infrastructure Agency, purchase data of municipalities, voting data, campaign finance data, tax data, PRH data, goods export data, financial data, forest information, international satellite data and weather data. Sometimes after using a specific set of Open Data, additional data was requested from the publisher. As presented by Carrara et al. (2015, pp. 33-34), geographic information is recognized as the area of data with the highest commercial value and it can be combined with other data. This was prominent also in the case examples as open maps were widely used in different business sectors.

Even if Open Data has often been used as supportive information, a couple of the respondents have also created business solutions that are based on Open Data. This included mobile applications and interactive maps. These were mainly free and open services. From one of them it was possible for the customer to buy desired reports directly from the map. The business aspect was though mainly missing, or it was minor in these solutions.

The two most pointed out barriers for Open Data utilization among the respondents were the unawareness of its' possibilities and the technological incapability in the company to process data. Many were hoping for case-examples of successful Open Data utilization to gain insights themselves. However, if there is no or not enough expertise in coding in the company, the possibilities to utilize Open Data are limited.

To some point, training the personnel on data processing could be helpful. The technical worries are somewhat related to the problem that data processing takes time. Under tight schedules it is sometimes impossible to start harvesting some interesting data. It is, therefore, important that the published Open Data should be well-formatted and documented. Notable was that the barriers are higher in smaller companies with less supportive resources from the IT-department to data-processing. The lack of case examples on Open Data utilization was also recognized as a barrier. If there were transparent case descriptions available on the utilization possibilities, the companies might be able to get new insights themselves.

The general opinion was that in a national level in Finland the availability of Open Data is not a problem. There is some data that was not open some years ago, but today the availability is significantly better. The availability-related barriers occur more the local scale as there is less Open Data available outside the cities involved in the six-city strategy in Finland. In an international level it was recognized that there is not that much Open Data than in Finland. For example, the access to geographical data is very limited in some countries or continents. If there was a pay for some data, it might have not been possible to use it profitably.

Data Visualizations were a daily routine in processing the data in many of the interviewed companies. Often, however, this was based on the competence of just few individuals within the company. Many saw, that as the technology has developed and applications have become more easy-to-use, the role of Business Intelligence and Data Visualizations will increase in their business in the future. The visualizations mentioned in the research included graphs and visualized maps. They were mainly used for producing easy-to-understand information for the end customer.

4.2 New Open Data Utilization Possibilities

First, it would be important to gain enough knowledge on Open Data and its' utilization possibilities. Some of the companies were not too familiar with this and felt it is hard to find relevant data and figure out what to do with it. It was also not evident if there would be some data the companies could publish themselves and somehow benefit on it.

In general, the company management in the case examples had a positive attitude towards Open Data. Free information, exploiting it and making predictions based on it were seen extremely important. Finding new trends and creating new solutions is the base for the business.

Even if the attitude was positive, no concrete data utilization plans had been made. Everything is much based on the personal interest of separate individuals in the companies. In most cases the company management had the trust that their employees will find the way to do their job best. New

data utilization possibilities are born by individuals' getting interested and familiarized with some data and experimenting it.

Despite the lack of systematical Open Data utilization development plans, the respondents were generally optimistic that the situation is getting better. Digitalization is affecting business more and more and tools for data processing are getting easier to use. Some were also hopeful for more data openings and improvements in data quality in the coming years. Even if it was difficult to name any new Open Data utilization possibilities, the companies felt that they will be able to utilize the data they are now using even better in the future. Some had ideas of including also new data types in their existing products to make them more comprehensive. The usage of local or areal data more felt compelling for many.

In one of the case company new people had recently been hired for data analysis related duties. Their main task is in processing company's own data, but it is also expected they can utilize external data in new ways. Some scattered plans or at least ideas on the next Open Data exploitations were also presented by some respondents.

The companies were also closely monitoring their competitors and what data are they using in their business. In some business fields it was very difficult to get this information whereas in others it was openly shared in the conferences in their field for example. Many pointed out the fact that the biggest players in the field have noticeably better resources and that they are not themselves able to utilize data on the same level.

Companies with lower IT-capabilities thought that data processing education for their employees might be helpful as it has not been a part of their traditional expertise. The digitalization of the business world is slightly altering the demands. Some of the respondents also personally hoped they would have more time for learning about data handling. The more digitally nature companies felt there is no need for external support on data utilization. They felt that improvements in data usability and quality are most important factor that can improve their chances of utilizing Open Data.

4.3 New Data Openings in Häme Region

Most respondents were quite familiar with the Open Data available now. Some felt there is already much data available they have not been able to use due for example their time resources. Others felt new openings are necessary. Better documentation of data was mentioned as it would make finding relevant data easier. Improvements were also hoped for data quality. Also, the updates for some data was found problematic if it was not clearly stated what has been changed.

The respondents had quite a solid view that on a national level in Finland the availability of Open Data is good. New Open Data on the level of regions and/or municipalities was hoped by many. Some knew that in Helsinki area and some other big cities the situation on data openings is much different to other parts of Finland. For many, it would be interesting to get statistics in different geographical areas to be able to make comparisons and plan their business operations.

Wishes for new data openings included:

- data from competitors
- accidents (regionally)
- healthcare (what services people use regionally)
- crimes
- court decisions
- emergency releases
- more financial data
- research measurement information
- information on public transport (already open in big cities)
- some closed data from authorities (questioned why it is closed)
- comprehensive international geographic data

Most of the companies, within the knowledge of their representative persons, had not made any efforts to get some data to become Open Data. Either they did not have any desired wishes, or they did not know who to contact or believe they can make an effect. One had in the past been involved in discussions with one public organization and eventually got their desired data open.

Some felt that even if there is an active Open Data community in Finland, the number of people belonging to it is small. In the big picture it was seen that, the data sets that can with a reasonable effort be made open, should be made open. Companies believed that more data is going to become open, but the new utilization possibilities remain still unclear.

5 CONCLUSION AND DISCUSSION

Open Data has been often referred as the world's greatest free resource. One might assume that a free resource would attract companies to widely exploit it in any way possible. Well, some do, but on a large scale the potential of Open Data remain untapped. There are still deficiencies in the awareness in companies on Open Data, their capabilities in handling it and transforming it into business and also partially in the availability of data. In a small region like Häme these deficiencies are being emphasized.

In Finland the availability of Open Data is rather good. Finland is one of the leading countries in Europe in data publishing and transparency. A lot of data has been made available in recent years, especially in the big cities. Open Data is easily reachable through public organizations websites and data portals like opendata.fi and Helsinki Region Infoshare. Open Data on a European level is gathered also in European Data Portal.

Smart Cities are evolving world-wide which increases the demand and creates new opportunities for Open Data. Smart solutions increasing people's life quality are being developed in transportation, engineering, sustainability, healthcare, education, IoT and many other sectors. Open Data is a significant raw material supporting many of these solutions. In Finland the Smart City development can be mostly seen in Helsinki and other cities involved in the Six City strategy, but likely the smaller cities and regions will eventually follow the development.

Finland is committed in opening and publishing new data and changing existing Open Data into a machine-readable form through the membership in Open Government Partnership. The initiatives aim in increasing governmental transparency and new innovations and technologies born based on growing re-use of public data. Ideally the data quality will also improve as time passes. The conditions for Open Data usage have become quite propitious. The question is how the companies can respond to the challenge.

The business possibilities on Open Data are diverse. As presented in this thesis, it can be exploited in various ways and various business sectors. It can be used as a tool for decision making, it may give added value to a product or service or a business may even be built on Open Data. Like Gurin (2014, pp. 15–16) presents, there are many types of utilizers in the process: suppliers, aggregators, developers, enrichers, enablers, value-adders and new start-ups. The highest commercial value is in geographical information (Carrara et al., 2015, pp. 33–34), but there is also a lot data of which utilization possibilities are highly sector dependent. On the other hand, in some fields there are options for multiple fields of Open Data to be utilized. The media for example can use data for many purposes.

Still, there are many challenges for Open Data breakthrough. The availability of data is still under construction. The most technically, legally and politically easiest to open data sets are in many places opened, but more sensitive or impactful data remain closed. Another issue besides data-availability is its interoperability. This means that even if data is available in many different territories, it is not comparable from city to another. It would be beneficial to create common reference sources within certain fields. (Chingard, 2013)

The recognized risks of opening data include issues related to constitutional right of citizens, privacy protection and the legal restrictions of data openings. The insufficient capability of data utilization of some public organizations is also a barrier for Open Data. Large unorganized data sets in unsuitable formats are difficult to utilize. Quality maintenance and keeping the data up to date are essential for keeping the data relevant. Another obstacle comes from the data-user side. Lack of knowhow may result in false interpretations of the data and lead to wrong decisions for companies or private users. There is also a risk that only the digital elite can exploit the data as not every citizen has the capabilities for it. This may lead to increasing social exclusion. Educating citizens and companies to respond the data revolution is vital for breaking these barriers. (Koski, 2015, pp. 16–17.)

The major barrier for Open Data utilization seems to be the companies unawareness of its' possibilities and the incapability to transform complex data into business. Even if the market value of Open Data is often estimated to be in thousands of billions, Herala (2018, p. 60) states that "it is not seen as a viable resource for companies, because it is difficult to find and use."

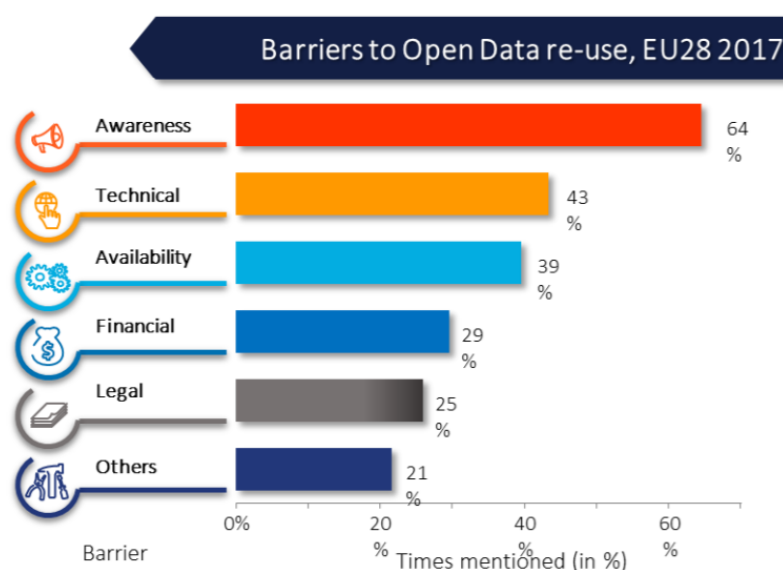


Figure 12. Barriers to Open Data re-use in EU28-countries in 2017 (Carrara, Radu & Vollers, 2017, p. 92)

European Commission (Carrara, et al., 2017, p. 92) has investigated the barriers for Open Data in EU28 countries. The lack of awareness on the benefits of Open Data is the most frequently mentioned barrier, technical barriers being next. The technical barrier referred mainly to the lack of machine-readable datasets preventing the immediate re-use of the data. Measures are taken in forms of awareness campaigns in some countries. In some countries the availability was also poor, and the respondents did not know where to find Open Data. The other barriers, like financial and legal barriers, were often linked to the poor awareness and lack of machine-readable Open Data. The findings in EU are much in line with the results of this research.

The barriers related to awareness are also braking Open Data breakthrough in the region of Häme. Local companies are not utilizing Open Data on a large scale. This is quite unquestionable assumption based on discussion with local Open Data actives. There is a Facebook group called Open Data Häme, founded already in 2013, for people interested in the topic. The group has currently (28 March 2020) 94 members. Nearly three years ago (3 March 2017) the group had 78 members (Kukkamäki, 2017, p. 1), which indicates there has been no significant change in the number of these actives in the region.

Timo Väliharju (interview 14 February 2020) stated that there are not a lot of local data sources available in Häme region and they are being exploited poorly. Public and private organizations are still taking their early footsteps in Open Data initiatives and are way behind the bigger cities in Finland. The interest of the companies in the theme is at a weak level and they are not attending the organized events or trainings. There is a little awareness in turning Open Data into profitable business. The cities in the region should also get more involved in opening data and promoting the issue. Applications based on Open Data are in the interest of these cities too. Open Data is mainly utilized in the ICT sector and transport and logistics.

This research did however reveal some local companies from different sectors that are utilizing Open Data. Some were utilizing it much, some less. The sectors involved in the research included insurance sector, media, traffic planning, programming and information solutions. The respondents were familiar with Open Data distribution channels and had developed their own models for utilizing it. Geographical Open Data was used in several sectors.

From business perspective, Open Data was not seen as a gold mine. The solutions created on Open Data gave added value for the products and services of the case companies, but not a lot of high-profit applications were being developed. The main barrier for Open Data utilization popping out in the interviews with these companies was the lack of awareness on Open Data possibilities and the lack of time and technical capability within

the company exploit data. In many cases the Open Data utilization was based on a personal interest of some individuals in the company on the topic rather than it being an important part of company's strategy.

Based on the research there seems to be a positive attitude towards Open Data in the companies in Häme region, but not a clear view or plans to utilize it on a large scale. If the companies had more time and resources, they could be able to get more out of Open Data. The companies with lower IT skills felt that trainings related to data analytics and coding, or even "advanced Excel skills", would help their personnel to utilize data better. Companies with higher IT skills underlined the importance of data quality and machine-readability for them to get interested in starting to use it.

The situation in data openings is considerably better in Helsinki and other big cities in Finland compared to Häme region. An undeniable demand for regional data exists within the local companies interested in utilizing Open Data. On a national level companies are happy with the level of Open Data availability, even if they might have some scattered ideas on new data openings.

The state of Open Data has taken giant leaps in Finland during the past decade. Some business solutions have been created on Open Data, but the business potential of it has not been released in a scale many might have thought some years back. The current decade will show if the development in technology and data quality will make the utilization easier to companies and lead to more profitable applications. The Smart City development forms a promising soil for this in bigger cities. Hopefully innovative solutions and success stories will be created by the companies in the smaller Häme region as well.

REFERENCES

- AvoinHäme (n.d.). Hankekuvaus. Retrieved 21 January 2018 from <https://xn--avoinhme-5za.fi/avoinhame-hanke-kuvaus/>
- Bakıcı, T., Almira, E. & Wareham, J. (2013). A Smart City Initiative: The Case of Barcelona. *Journal of the Knowledge Economy* 4(2), 135-148.
- Berners-Lee, T. (2010). Linked Data. Retrieved 29 October 2018 from <https://www.w3.org/DesignIssues/LinkedData.html>
- Biedenbach, G. & Boström, G.-O. (2018). *A Data-driven Lab in the Context of Open Data: Opportunities and Challenges for a Sustainable Business Model*. Umeå School of Business, Economics and Statistics. Retrieved 28 March 2020 from <http://umu.diva-portal.org/smash/get/diva2:1262264/FULLTEXT01.pdf>
- Braunschweig, K., Eberius, J., Thiele, M. & Lehner, W. (2012). The State of Open Data: Limits of Current Open Data Platforms. Semantic Scholar. Retrieved 24 April 2018 from <https://www.semanticscholar.org/paper/The-State-of-Open-Data-Limits-of-Current-Open-Data-Braunschweig-Eberius/f56700355734c88ecb04005b84eb80362e35803f>
- Capdevila, I. & Zarlenga, M.I. (2015). Smart City or smart citizens? The Barcelona Case. *Journal of Strategy and Management* 8(3), 266-282.
- Carrara, W., Chan, W., Fischer, S. & van Steenberg, E. (2015). Creating Value through Open Data. European Commission. Retrieved 20 January 2020 from https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf
- Carrara, W., Radu, C. & Vollers, H. (2017). Open Data Maturity in Europe 2017: Open Data for a European Data Economy. European Commission. Retrieved 27 March 2020 from https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n3_2017.pdf
- Cassandras, C.G. (2017). Automating Mobility in Smart Cities. *Annual Reviews in Control* 44, 1-8. Retrieved 8 October 2018 from <https://www.sciencedirect.com.ezproxy.hamk.fi/science/article/pii/S1367578817301281?via%3Dihub>

Chan, C.M.L. (2013). From Open Data to Open Innovation Strategies: Creating e-Services Using Open Government Data. *46th Hawaii International Conference on System Sciences*, 1890-1899. Retrieved 17 February 2018 from

https://www.researchgate.net/publication/251880036_From_Open_Data_to_Open_Innovation_Strategies_Creating_E-Services_Using_Open_Government_Data

Chingard, S. (2013). A Brief History of Open Data. *Paris Innovation Review*. Retrieved 29 October 2017 from

<http://parisinnovationreview.com/articles-en/a-brief-history-of-open-data>

City of Hämeenlinna (2019a). Avoin data. Retrieved 9 January 2020 from <https://www.hameenlinna.fi/hallinto-ja-talous/talous-ja-hankinnat/avoin-data/>

City of Hämeenlinna (2019b). Älykkäät palvelut. Retrieved 25 February 2020 from

<https://www.hameenlinna.fi/tyo-ja-elinkeino/openhameenlinna-fi/alykkaat-palvelut/>

Digital and Population Data Services Agency. Palvelukuvaus. Retrieved 9 January 2020 from <https://www.avoindata.fi/fi/opas/palvelukuvaus>

Dietrich, D., Gray, J., McNamara, T., Poikola, A., Pollock, R., Tait, J. & Zijlstra, T. (n.d.). Open Data Handbook. Retrieved 29 October 2017 from <http://opendatahandbook.org/guide/en/>

Dupin-Bryant, P. & Olsen, D. (2014). Business Intelligence, Analytics and Data Visualization: A Heat Map Project Tutorial. *International Journal of Management & Information Systems* 3, 185-200.

Eskola, J., Lähti J. & Vastamäki, J. (2018). Teemahaastattelu: Lyhyt selviytymisopas. In Valli R. (ed.) *Ikkunoita tutkimusmetodeihin 1. Metodien valinta ja aineistonkeruu: virikkeitä aloittelevalle tutkijalle*. Retrieved 20 January 2020 from <https://ellibslibrary.com/hamk/978-952-451-516-0>

Gournelos, T., Hammonds, J. R. & Wilson, M. A. (2019). *Doing Academic Research: A Practical Guide to Research Methods and Analysis*. Retrieved 21 January 2020 from

<https://ebookcentral-proquest-com.ezproxy.hamk.fi/lib/hamk-ebooks/reader.action?docID=5773107>

Gurin, J. (2014). *Open Data Now: The Secret to Hot Startups, Smart Investing, Savvy marketing, and Fast Innovation*. New York: McGraw Hill Education.

Gurstein, M. (2011). Open Data: Empowering the Empowered or Effective Data Use for Everyone?. First Monday. Retrieved 30 October 2017 from <http://firstmonday.org/ojs/index.php/fm/article/viewArticle/3316/2764> "

HAMK (n.d.). HAMKin avoin data. Retrieved 29 January 2020 from <https://www.hamk.fi/tutkimus/avoin-data/>

Herala, A. (2018). *Benefits from Open Data: Barriers to Supply and Demand of Open Data in Private Organizations*. Dissertation. Acta Universitatis Lappeenrantaensis 809. Retrieved 27 March 2020 from <https://lutpub.lut.fi/bitstream/handle/10024/158560/Antti%20Herala%20A4.pdf?sequence=1&isAllowed=y>

Honkanen, M. (2019). Avoin data ja liiketoiminta. Avoimen datan opas. Retrieved 28 October 2019 from <https://www.avoindata.fi/fi/opas/avoin-data-ja-liiketoiminta>

Jussila, J., Kukkamäki, J., Mäntyneva, M. & Heinisuo, J. (2019). Open Data and Open Source Enabling Smart City Development. *Technology Innovation Management Review* 9(9)/2019, 25-34. Retrieved 29 January 2020 from: https://timreview.ca/sites/default/files/article_PDF/TIMReview_September2019%20-%20final-C.pdf

Koski, H. (2015). The Impact of Open Data – a Preliminary Study. *Ministry of Finance Publications* 15b/2015. Retrieved 25 January 2020 from <https://vm.fi/documents/10623/1107406/The+Impact+of+Open+Data/1c432b3a-a5e8-41ea-a5ea-135280a69ea3?version=1.0>

Koski, H., Honkanen, M., Luukkonen, J., Pajarinen, M. & Ropponen, T. (2017). Avoimen datan hyödyntäminen ja vaikuttavuus. *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisu* 40/2017. Retrieved 9 January 2020 from https://vnk.fi/documents/10616/3866814/40_avoimen+datan+16032017.pdf/0444467d-5400-4f0c-8728-2447cef039ad/40_avoimen+datan+16032017.pdf?version=1.0

Kukkamäki, J. (2017). Avoimen datan tila Kanta-Hämeessä. Häme University of Applied Sciences. Retrieved 14 September 2019 from <https://docplayer.fi/47935705-Avoimen-datan-tila-kanta-hameessa-raportti-joni-kukkamaki.html>

Laakso, M. (2017). Smart Cities in Finland - an Introduction. Flanders Investment & Trade. Retrieved 29 September 2019 from https://www.flandersinvestmentandtrade.com/export/sites/trade/files/market_studies/Smart%20Cities_Finland_2017.pdf

Lähdesmäki, T., Hurme, P., Koskimaa, R., Mikkola, L. & Himberg, T. (2009). *Methods Paths for Humanists*. University of Jyväskylä, Faculty of Humanities. Retrieved 2 February 2018 from <https://koppa.jyu.fi/avoimet/hum/menetelmapolkuja/en>

Manninen, A., Poikola A., Rastas, T., Rainio, A., Rehbinder, M., Tuominen, T., Kauhanen-Simanainen, A., Marttila, S., Suonsaari, J., Ahola, A., Saarinen, J., Koivusaari, K., Lahtinen, S., Kivekäs, O., Honkanen, M., Tjukanov, T., Nolvi, T., Haverinen, K., Kettunen, E. & Lahti, T. (2019). Avoin data 10 vuotta Suomessa 10.10.2019. HelsinkiKanava. Retrieved 26 October 2019 from <https://www.helsinkiKanava.fi/fi/web//helsinkiKanava/player/vod?assetId=39472438>

Ministry of Finance (2013). Finland's Action Plan on Open Government. Open Government Partnership. Retrieved 27 January 2020 from <https://www.opengovpartnership.org/wp-content/uploads/2019/06/20130314-OGP-Action-Plan-Finland.pdf>

Ministry of Finance (2017). Open Government III Action Plan 2017-2019 Finland. Avoin Hallinto. Retrieved 27 January 2020 from <https://avoinhallinto.fi/assets/files/2019/09/Open-Government-III-Action-Plan.pdf>

Ministry of Finance (2019). Open Government National Action Plan for 2019-2023 Finland. Avoin Hallinto. Retrieved 27 January 2020 from https://avoinhallinto.fi/assets/files/2019/09/ENGLANTI_Avoinhallinto_IV_toimintaohjelma_FINAL_240919.pdf

Monino, J-L. & Sedkaoui, S. (2016). *Big Data, Open Data and Data Development*. Wiley ISTE.

Muurinen, R. (2017). Liiketoimintaa avoimesta datasta. 6Aika avoin data ja rajapinnat project. Retrieved 25 January 2020 from https://6aika.fi/wp-content/uploads/2017/05/Liiketoimintaa_avoimesta_datasta_2017.pdf

Open Government Partnership (n.d.). About OGP. Retrieved 11 February 2018 from <https://www.opengovpartnership.org/about/about-ogp>

Open Knowledge Foundation (n.d.). Place Overview. Retrieved 25 January 2020 from <https://index.okfn.org/place/>

Open Knowledge International (n.d.). Open Definition. Retrieved 27 October 2018 from <http://opendefinition.org/>

Pesonen, J. & Lampi, M. (2016). Utilizing Open Data in Tourism. Texas A&M AgriLife. Retrieved 24 January 2020 from http://agrilife.org/ertr/files/2016/01/ENTER2016_submission_122.pdf

Picon, A. (2015). *Smart cities: A Spatialized Intelligence*. John Wiley & Sons Ltd.

Ramboll Finland. (n.d.). Yritys. Retrieved 28 February 2020 from https://fi.ramboll.com/ramboll_finland_oy

Regional Council of Häme (n.d.). The Häme Region. Retrieved 29 January 2020 from: <https://www.hameenliitto.fi/en/hame-region>

Routio, P. (2007). Thematic Interview. Arteology, the science of products and professions. The Aalto University School of Art and Design. Retrieved 2 February 2018 from <http://www2.uiah.fi/projects/metodi/164.htm#teemahaas>

Rowe, K. (2013). Eight “Big Tent” Criteria for Excellent Qualitative Research. *Semantic Scholar*. Retrieved 26 January 2020 from <https://www.semanticscholar.org/paper/Women%27s-cycling-participation-in-Australia%3A-and-the-Rowe/99a0d11bd1af2a14d3addf23a25ea4256f973f9f/figure/17>

Saldana, J. (2011). *Fundamentals of qualitative research*. Retrieved 21 January 2020 from <https://ebookcentral-proquest-com.ezproxy.hamk.fi/lib/hamk-ebooks/reader.action?docID=665394>

Simosol. (n.d.). Global Leaders in Solutions for Sustainability and Optimization. Retrieved 28 February 2020 from <https://www.simosol.fi/>

Simperl, E. & Walker, J. (2017). Analytical Report 8: The Future of Open Data Portals. European Data Portal. Retrieved 5 November 2018 from https://www.europeandataportal.eu/sites/default/files/edp_analyticalreport_n8.pdf

Stirrup, J. (2014). *Tableau Dashboard Cookbook*. Birmingham, UK: Packt Publishing Ltd.

Tracy, S. J. (2019). *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact*. Retrieved 23 January 2020 from <https://ebookcentral-proquest-com.ezproxy.hamk.fi/lib/hamk-ebooks/reader.action?docID=5847435>

Yuk, M. & Diamond, S. (2014). *Data Visualization for Dummies*. Hoboken, NJ: John Wiley & Sons, Inc.

INTERVIEWS:

Piipponen, J. (2020). Data Scientist, Ramboll Finland. Interview 27 February 2020.

Rasinmäki, J. (2020). CEO, Simosol. Interview 6 March 2020.

Sulonen, P. (2020). Head of Display Competence Center, Teleste. Interview 13 March 2020.

Taipale, J. (2020). Full Stack Developer, Vitabalans. Interview 7 April 2020.

Uhari, M. (2020). News Editor, Hämeen Sanomat. Interview 24 February 2020.

Väliharju, T. (2020). Executive Director, COSS. Interview 14 February 2020.

Wiitamäki, P. Product Development Manager, Pohjantähti. Interview 19 February 2020.

Appendix 1

APPENDIX HEADING

Thematic Interview Base Structure

Warm-up questions:

- a. Can you tell me shortly about the company you are working for and what are your responsibilities?
 - b. How is your company utilizing data in general in your business?
1. Open Data utilization in the company – current situation
 - a. How familiar are you with the concept of Open Data? What do you know about it?
 - b. Do you know where search for Open Data?
 - c. How much do you utilize Open Data in your business now?
 - d. What kind of data sets do you utilize?
 - e. What business solutions / innovations have you created with the help of Open Data?
 - f. What are the possible barriers for your Open Data utilization?
 - g. How familiar are you with the concept of Business Intelligence?
 - h. Do you utilize Data Visualization in your business?
 - i. Are you familiar with Data Visualization tools? Does your company have the expertise to use them?
 2. New Open Data utilization possibilities for the company
 - a. How well do you know the possibilities of Open Data utilization?
 - b. What does your company management think about Open Data?
 - c. Have you thought of some new Open Data utilization possibilities for your company?
 - d. Do you know how others have utilized Open Data in your field of business?
 - e. What kind of Open Data could you still utilize in your business (GPS/weather/economic/etc.)?
 - f. What kind of support would you like to get to utilize Open Data better?
 3. Open Data demand
 - a. How well do you know what kind of data sets there are available?
 - b. For what kind of data sets would you still like to get an access to?
 - c. Which are the organizations you would like to get more data from?
 - d. Have you done some efforts to make some desired data to become Open Data?